Water Research X 13 (2021) 100123

Contents lists available at ScienceDirect

Water Research X

journal homepage: www.sciencedirect.com/journal/water-research-x

Spatial-temporal targeted and non-targeted surveys to assess microbiological composition of drinking water in Puerto Rico following Hurricane Maria

Maria Sevillano [a], Solize Vosloo [i], Irmarie Cotto [a], Zihan Dai [b,c], Tao Jiang [a], Jose M. Santiago Santana [d], Ingrid Y. Padilla [e], Zaira Rosario-Pabon [f], Carmen Velez Vega [f], José F. Cordero [g], Akram Alshawabkeh [a], April Gu [h], Ameet J. Pinto [i,*]

[a] Department of Civil and Environmental Engineering, Northeastern University, Boston, MA, United States of America
[b] Key Laboratory of Drinking Water Science and Technology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China
[c] University of Chinese Academy of Sciences, Beijing, China
[d] Department of Natural Sciences, University of Puerto Rico, Carolina, PR, United States of America
[e] Department of Civil Engineering and Surveying, University of Puerto Rico, Mayagüez, PR, United States of America
[f] University of Puerto Rico—Medical Sciences Campus, San Juan, PR, United States of America
[g] Department of Epidemiology and Biostatistics, University of Georgia, Athens, Georgia, United States of America
[h] School of Civil and Environmental Engineering, Cornell University, Ithaca, NY, United States of America
[i] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, United States of America

ARTICLE INFO

Keywords:
Drinking water quality
Hurricane Maria
Metagenomics
qPCR
Genome-resolved metagenomics

ABSTRACT

Loss of basic utilities, such as drinking water and electricity distribution, were sustained for months in the aftermath of Hurricane Maria's (HM) landfall in Puerto Rico (PR) in September 2017. The goal of this study was to assess if there was deterioration in biological quality of drinking water due to these disruptions. This study characterized the microbial composition of drinking water following HM across nine drinking water systems (DWSs) in PR and utilized an extended temporal sampling campaign to determine if changes in the drinking water microbiome were indicative of HM associated disturbance followed by recovery. In addition to monitoring water chemistry, the samples were subjected to culture independent targeted and non-targeted microbial analysis including quantitative PCR (qPCR) and genome-resolved metagenomics. The qPCR results showed that residual disinfectant was the major driver of bacterial concentrations in tap water with marked decrease in concentrations from early to late sampling timepoints. While *Mycobacterium avium* and *Pseudomonas aeruginosa* were not detected in any sampling locations and timepoints, genetic material from *Leptospira* and *Legionella pneumophila* were transiently detected in a few sampling locations. The majority of metagenome assembled genomes (MAGs) recovered from these samples were not associated with pathogens and were consistent with bacterial community members routinely detected in DWSs. Further, whole metagenome-level comparisons between drinking water samples collected in this study with samples from other full-scale DWS indicated no significant deviation from expected community membership of the drinking water microbiome. Overall, our results suggest that disruptions due to HM did not result in significant and sustained deterioration of biological quality of drinking water at our study sites.

Introduction

A 2015 report on the Safe Drinking Water Act violations in Puerto Rico (PR) indicated high levels of contaminants such as volatile organic compounds (VOC), total coliform bacteria, and disinfection by products (DBPs) impacted around 70% of the islands population (NRDC, 2017). This report recommended investment in drinking water systems (DWSs), including treatment, distribution system upgrade and maintenance, and source water protection. Such investments are also important across the US, as the water infrastructure continues to age (ASCE, 2017) and water

* Corresponding author at: Georgia Institute of Technology, 311 Ferst Drive, Atlanta, Georgia 30332-0512, United States.
E-mail address: ameet.pinto@ce.gatech.edu (A.J. Pinto).

https://doi.org/10.1016/j.wroa.2021.100123
Received 12 May 2021; Received in revised form 28 September 2021; Accepted 30 September 2021
Available online 12 October 2021

quality violations are being increasingly reported (Allaire et al., 2018). Further complicating the issue of providing regulation compliant water while relying on an aging water infrastructure is the increasing frequency and intensity of extreme weather events (Estrada et al., 2015; Goodess, 2012). In the year 2017 alone, Hurricanes Harvey, Irma, and Maria (HM) caused widespread damages and were categorized as historic billion-dollar disasters in the US (NOAA NCEI, 2020). The resiliency of DWSs during these extreme events is particularly important, as lack of access to safe drinking water may result in further detrimental health impacts. Natural disasters can contaminate source waters, impacting proper treatment, distribution, and ultimately affect consumer health (Ashbolt, 2015; Exum et al., 2018) Previous studies have highlighted water quality degradation associated with extreme weather events like hurricanes. Schwab et al. measured concentrations of fecal coliforms, E. coli, and enterococci in tap and surface waters following Hurricane Katrina and did not recover any of the bacterial indicators in tap water samples irrespective of chlorine residual concentration (Schwab et al., 2007). A recent study on the impacts of Hurricane Harvey on water quality from two DWSs in Texas highlighted that source water quality and water demand and their relationship to water age strongly impacted microbial communities and influenced the time for recovery (Landsman et al., 2019). Similarly, a amplicon sequencing-based study carried out in St. Thomas, post Hurricane Irma and HM revealed that the microbial community structure in rain cisterns, coastal stations, and surface runoff waters was dramatically different between sampling sites with fecal indicator bacteria (FIB) detected in cisterns used as household water supply (Jiang et al., 2020).

HM, classified as a category 4 hurricane, impacted 3 million people in PR. Loss of basic utilities (i.e., water, cellular coverage, and electricity) was associated with remoteness category (Kishore et al., 2018). Water services generally recovered quickly in densely populated areas, while remote areas either recovered quickly or months later. However, electricity services took longer to recover irrespective of remoteness category. Even when water services were restored, intermittent water supply was common due to unreliable electrical supply; this could potentially degrade water quality via stagnation and loss of disinfectant residual, intrusion, and backflows (Bautista-de los Santos et al., 2019). Boil advisories and point of use chlorination were in place after water services resumed and were reported by the Puerto Rico Aqueducts and Sewers Authority (PRASA) through mid-January 2018 (Exum et al., 2018). Previously, Lin et al. (Lin et al., 2020) provided insights into metals, micropollutants, and molecular toxicity of pre- and post-HM drinking water samples in PR and showed the impact of HM on chemical water quality, suggesting that trace metals were potential drivers of cumulative risk from drinking water. Additionally, Keenum et al. (Keenum et al., 2021) characterized five unregulated small scale DWS and one large PRASA DWS in PR six months after HM. The authors used targeted culture and molecular based analyses (i.e., quantitative PCR (qPCR), 16S rRNA amplicon sequencing) and demonstrated similar microbial communities and concentrations of opportunistic premises plumbing pathogens (OPPPs) compared to those reported in the continental US. In our study, we also aim to evaluate the microbial water quality in the aftermath of HM. However, unlike Keenum et al. (Keenum et al., 2021), we conducted a recurrent sampling campaign beginning in December 2017 spanning nine locations across PR for a duration of a year. Despite the magnitude of HM in PR, there hasn't been a large effort to characterize microbial water quality. To date, there have been two reports focused on chemical contamination (Lin et al., 2020; Warren, 2019) and two (including this one) on microbial composition (Keenum et al., 2021) of DWSs on the island. These studies are essential to establish relationships, sampling infrastructure, and methodologies needed to respond to future storms, as well as to communicate risk and execute corrective actions to decrease exposure risk and unwanted health outcomes. Thus, our goals were (1) to utilize an extended spatial-temporal sampling campaign to determine if changes in drinking water microbiome were indicative of disturbance followed by recovery, (2) if this disturbance-recovery dynamic was associated with presence of potential pathogens, (3) whether potential pathogen presence was persistent or transient, and finally (4) whether microbial composition of PR drinking water was consistent with or deviated significantly from other drinking water systems.

2. Materials and methods

2.1. Drinking water sampling and water quality analyses

Nine sampling locations were chosen across different geographic locations in PR (Fig. 1). The choice of the sampling locations/drinking water systems was based on (1) accessibility and (2) geographic distribution. The sites represent locations that were consistently accessible throughout the sampling timeframe. In the immediate aftermath of HM's landfall, the island's entire power grid was inoperable, resulting in potable water distribution losses in all sampling sites and most of the island. Despite varying degree of damages at the sites (e.g., electricity, water distribution, road accessibility, flooding), the sites returned to operation in the order of weeks, owing to their commercial/service capabilities. The sites' abstraction source was surface waters for SJU, CAR, HAT, AGU, MAY, and HUM, and groundwater for MAN, GUA, and CAY. Tap water was flushed for 20–25 min during which time temperature and total chlorine were monitored to ensure stabilization indicating elimination of stagnant water from the premises plumbing. Subsequently, it was filtered in triplicate on site through 0.2 μm Sterivex filters (EMD Millipore™, Cat. no. SVGP01050) using a field peristaltic pump (Geotech, Cat. no. 91352123) until the filter clogged or up to a 20 L volume for each filter. Sterivex filters for samples as well as field blanks were transported on ice during sample transportation, frozen at −20 °C at the end of the sampling day, and then stored at −80 °C until DNA extraction. Water quality parameters (i.e., temperature, pH, conductivity, and dissolved oxygen) were recorded on site with an Orion Star probe (Thermo Scientific, Cat. no. 13645571). A portable spectrophotometer (HACH, Cat. no. DR1900–01H) was used to measure total chlorine (HACH, Cat. no. 2105669) and phosphate (HACH, Cat. no. 2106069) on site. Nitrogen species (i.e., ammonia, nitrate, nitrite) were measured in the laboratory with a HACH spectrophotometer using HACH test and tube format (HACH, Cat. no. 2606945, 2605345, 2608345, respectively). Total Organic Carbon (TOC) was measured with a Shimadzu TOC- LCPH Analyzer (Shimadzu, Kyoto, Japan). Additional details about the 54 samples can be found in Table S1.

2.2. DNA extraction and total bacteria qPCR

DNA extractions were performed using a modified version of the DNeasy PowerWater Kit (QIAGEN, Cat no. 14900–50-NF) protocol (Vosloo et al., 2019). Briefly, the polyethersulfone (PES) membrane from the Sterivex filter was processed by aseptically cutting it into smaller pieces and transferring to a Lysing Matrix E tube (MP Biomedical, Cat. no. MP116914100). Subsequently, 294 μL of 10X Tris-EDTA buffer pH 8 (G-Biosciences, Cat. no. 501035446) was added to the Lysing Matrix E tube and supplemented with 6 μL of lysozyme (50 mg mL^{-1}, Thermo Fisher Scientific, Cat. no. 90082), followed by a 60 min incubation at 37 °C with mixing at 300 rpm. Subsequently, 300 μL of PW1 solution from DNeasy PowerWater Kit was mixed in and 30 μL of Proteinase K (20 mg mL^{-1}, Fisher Scientific, Cat. no. AM2546) was added. An incubation period of 30 min at 56 °C with mixing at 300 rpm followed. Previously removed spheres from the corresponding Lysing E matrix tube were replenished and 630 μl chloroform: isoamyl alcohol (Fisher Scientific, Cat. no. AC327155000) was added. Bead beating was performed at setting 6 for 40 s using a FastPrep – 24™ (MP Biomedical, Cat. no. 116004500). The resulting homogenized mixture was centrifuged for 10 min at 14,000 x g at 4 °C and the upper aqueous phase was transferred to a clean 1.5 mL tube. A supplement of 6 μL carrier RNA (prepared by mixing 310 μl of Buffer EB from DNeasy PowerWater Kit

Site	Source	Site	Source	Site	Source
SJU	S	HAT	S	GUA	G
CAR	S	AGU	S	CAY	G
MAN	G	MAY	S	HUM	S

Fig. 1. Site nomenclature by geographic location. Colors correspond to sampling sites and are consistent in all figures. Abstraction sources, surface water (S) or groundwater (G), are depicted in table.

with 310 μg lyophilized carrier RNA (QIAGEN, Cat. no. 1,068,337)) was mixed with 600 μL of recovered supernatant. This was then purified using the automated DNA purification protocol with DNeasy PowerWater Kit on a QIACube system (QIAGEN, Cat. no. 9001292). In addition to the samples, controls were processed identically and consisted of unused transported Sterivex filter membranes (filter blank), no input material (reagent blank), and sterilized deionized water filtered through Sterivex (water blank). This set of three controls were included with each sampling campaign ($n = 6$) and extraction run.

The 16S rRNA gene was targeted to quantify bacterial concentrations using a previously published primer pair (Caporaso et al., 2011, Table S2). qPCR was performed on a QuantStudio™ 3 Real-Time PCR System (ThermoFisher Scientific Cat. no. A28567). PCR reactions were carried out in a 20 μl volume containing Luna Universal qPCR Master Mix (New England Biolabs, Inc., Cat. no. NC1276266), primer pair (IDTDNA), DNAse/RNAse-Free water (Fisher Scientific, Cat. no. 10977015), and 5 μL of 10X diluted DNA template. Reactions were prepared by an epMotion M5073 liquid handling system (Eppendorf, Cat. no. 5073000205D) in triplicate. The cycling conditions were as follows, initial denaturing at 95 °C for 1 min followed by 40 cycles of denaturing at 95 °C for 15 s, annealing at 50 °C for 15 s, and extension 72 °C for 1 min. Melting curve analyses was performed by ramping from 72 °C to 95 °C for 15 s, and 60 °C for 1 min, 95 °C for 15 s. A negative control (NTC) and a standard curve consisting of 7 points ranging from 10^1 to 10^7 copies of 16S rRNA gene were included in every qPCR run.

2.3. qPCR analyses for waterborne pathogens

Previously published primers targeting *Legionella* spp. (Nazarian et al., 2008; Yáñez et al., 2005), *Mycobacterium* spp. (Chern et al., 2015; Radomski et al., 2010), pathogenic Leptospira (Stoddard et al., 2009), and *Pseudomonas aeruginosa* (Anuj et al., 2009) were used for qPCR assays. Reactions were set up by an epMotion M5073 liquid handling system in triplicate. The assays consisted of 2X PrimeTime Gene Expression Master Mix (IDTDNA, Cat no. 290479057) with low reference ROX dye, target primers and probe (IDTDNA), 5 μL of 10X diluted DNA template and water (UltraPure™ DNase/RNase-Free Distilled Water, Thermo Fisher Scientific, Cat. no. 10977015). Single target reactions were conducted in a total volume of 20 μL, whereas duplex qPCR (i.e., *Pseudomonas aeruginosa* assay) were conducted in 25 μL. Primer and probe sequences and cycling conditions are described in Table S2. Standard cycling conditions for all reactions were programmed on a QuantStudio™ 3 Real-Time PCR System. Target gene copy numbers were determined by comparing threshold cycle with standard curve generated using gblocks gene fragments as standards (Table S2).

The qPCR efficiency, R2, Limit of detection (LOD), and Limit of quantitation (LOQ) for all qPCR assays are provided in Table S2. Negative controls (i.e., filter blank, reagent blank, and water blank) for each sampling campaign were included in every assay.

2.4. Shotgun sequencing and metagenomic reads pre-processing

Genomic DNA from selected samples were sent to University of Illinois Roy J. Carver Biotechnology Center (UI-RJCBC) for library preparation using a low input DNA kit (NuGEN, Cat. no 0344NB). Libraries were loaded into two SP lanes on a NovaSeq 6000 instrument with an output of 2 × 150 nt reads. The prepared libraries included 33 samples and 3 pooled blanks (i.e., filter blank, reagent blank, and water blank) from all locations. The raw reads obtained from UI-RJCBC were processed with fastp (Chen et al., 2018) v0.19.7 to remove homopolymer stretches using the following flags '–trim_poly_g –trim_poly_x'. Further, trimmed reads were mapped with BWA-MEM (Li, 2013) v0.7.12 against the UniVec database build 10.0 (National Center for Biotechnology Information 2016) to perform vector screening and retained unmapped paired reads. Subsequently, Nonpareil (Rodriguez-R et al., 2018) v3.303 was used on the quality filtered reads to estimate average community coverage and metagenomic dataset diversity using kmer algorithm with a kmer size of 20 and default parameters.

2.5. Metagenome assembly and mapping

Sample reads were co-assembled based on each sampling location using metaSPAdes (Nurk et al., 2017) v 3.11.1 with the following flags '-meta -t 16 –phred-offset 33 -m 500 -k 21,33,55,77,99,119′ and further filtered to a minimum scaffold length of 500 bp. Reads from samples and controls (i.e., extraction blank, filter blank, DI water blank) were mapped to co-assemblies using BWA-MEM. An approach similar to Dai et al. (Dai et al., 2020) was used to remove potential contaminant scaffolds. Briefly, BWA-MEM with flag -F4 and -f2 was used to map sample and control reads against co-assemblies. The BEDtools (Quinlan and Hall, 2010) genomecov using flags -g, and -d was used to calculate coverage and per base coverage using generated BAM files. Relative abundances (RA) and normalized coverage deviation (NCD) were calculated for each scaffold with coverage and per base coverage information, respectively.

Scaffolds that were not detected in the blanks or for which sample RA was greater than the blank RA and the sample NCD was less than the blank NCD were considered true scaffolds and were used in downstream analyses. All scaffolds that did not meet these criteria were considered contaminant scaffolds and removed from further analyses. Assembly statistics were obtained from QUAST (Gurevich et al., 2013) 5.0.2. To contextualize the metagenome assemblies with respect to other distribution systems, assemblies from other drinking water systems (considered unperturbed systems because samples were not associated with any natural disaster or water quality issues) were compared against our co-assemblies. MASH (Ondov et al., 2016) v2.1.1 was used to estimate the dissimilarity between assemblies using '-r' and '-m 2′ flags, and a sketch size of 100000.

2.6. Taxonomic classification of metagenomic assemblies

The taxonomic classification of scaffolds was performed with a contig annotation tool (Von Meijenfeldt et al., 2019) (CAT v5.0.4) program in which open reading frames (ORF) are predicted with Prodigal (Hyatt et al., 2010) and used as alignment queries by DIAMOND (Buchfink et al., 2014) against the NCBI non-redundant (nr) protein database (downloaded ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/, 2020–03–04). Selected genera known to contain pathogenic species, as well as non-pathogenic species, that are relevant to drinking water systems were further examined. The CAT annotations of true scaffolds were evaluated against annotated controls and only scaffolds with rpoB normalized coverage above controls and additional annotation support were considered to avoid false positives at the genus level. Specifically, additional support consisted of classification with kaiju (Menzel et al., 2016) v1.7.2 using reference indexes containing NCBI BLAST nr database and microbial eukaryotes with default parameters (downloaded http://kaiju.binf.ku.dk/server, nr_euk 2019–06–25) and/or by kraken2 (Wood et al., 2019) v2.0.9-beta against RefSeq database (downloaded https://lomanlab.github.io/mockcommunity/mc_databases.html). Additionally, we examined the scaffolds of eukaryotic origin with metaEuk (Levy Karin et al., 2020) v3.8dc7e0b to assign taxonomy using a publicly available MMseqs2 database containing protein profiles from the marine eukaryotic reference catalog (MERC), Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP), and Uniclust50 (downloaded http://wwwuser.gwdg.de/~compbiol/metaeuk/2020_TAX_DB/). The relative abundance of each taxonomic unit is presented by using the relative abundance of scaffolds (see Section 2.5) within each unit.

2.7. Metagenomic assembled genomes

Co-assemblies were binned with CONCOCT (Alneberg et al., 2014) within anvi'o (Eren et al., 2015) v5.1 by clustering scaffolds 2500 bp or longer into metagenome assembled genomes (MAGs) and manually refining them within the anvi'o platform. Furthermore, dRep (Olm et al., 2017) v2.3.2 was used to dereplicate MAGs and obtain representative genomes with flags '-comp 50 -con 10′ and default values. GTDB-tk (Parks et al., 2018) v0.1.3 was used to assign taxonomy to MAGs with the flag 'classify_wf'. Sample reads were mapped to corresponding single pseudo contig MAGs. Pseudo contigs were generated with the union command in EMBOSS (Rice et al., 2000) utility. Mapping was performed with BBMap (Bushnell, 2015) v38.24 using a 90% identity threshold and setting flags 'ambiguous=best', 'mappedonly=t', and 'pairedonly=t'. Detection of a MAG in a sample was established when $\geq$ 25% of its bases were covered by at least one read from the corresponding sample. Coverage was determined with samtools (Li et al., 2009) v1.10 'coverage'. The abundance of a MAG in a sample was calculated as sample reads mapped per million reads per genome length in kbp (RPKM). Further information about MAGs, such as number of 5S rRNA, 16S rRNA, 23S rRNA, and tRNA counts was obtained by annotating the MAGs using DRAM (Shaffer et al., 2020) v1.0.6. The databases used with DRAM were downloaded with the following flags 'DRAM-setup.py prepare_databases –output_dir DRAM_data –skip_uniref'. MAGs from this study were compared to 52,515 MAGs recovered from environmentally diverse metagenomes by Nayfach et al. (Nayfach et al., 2020) (downloaded https://portal.nersc.gov/GEM/genomes/fna.tar, 2020–11–10) using FastANI (Jain et al., 2018) v2.3.2 with default parameters. Metadata linked with this genomic catalog of earth microbiomes, hereafter referred to as JGI MAGs (downloaded https://portal.nersc.gov/GEM/genomes/genome_metadata.tsv, 2020–11–30) was used to address niche association. Further, SEARCH-SRA (Stewart et al., 2015; Torres et al., 2017; Towns et al., 2014) online portal was used to interrogate the SRA database (246,329 records) by aligning metagenomic datasets to our MAGs. Only records that mapped 10 or more reads from the SRA collection were further inspected. The metadata associated with SRA accession numbers (downloaded https://s3.amazonaws.com/starbuck1/sradb/SRAmetadb.sqlite.gz, 2021–04–08) was incorporated through custom scripts in R software (R Development Core Team, 2016) that rely on the dbplyr (Wickham et al., 2021) package. Records that were classified as retrieved from metagenomic library sources and with a whole genome sequence strategy were retained for analysis. Considering that SRA metadata is user provided, manual curation to ensure consistency and retrieve the same ecosystem categories as in JGI MAGs metadata was performed. Data with missing context (lacking information in title or description) were considered as "others" and removed from analyses. The association between a MAG and an ecosystem category was determined by multiplying the total number of reads from the ecosystem category mapping to the MAG by the ratio of the number of unique SRA records associated with a MAG and ecosystem category and the number of unique SRA records within the entire dataset assigned to the ecosystem category. For a schematic representation of bioinformatic approaches performed in this study, refer to Figure S1 and Figure S2.

2.8. Data analyses and statistics

Statistical analyses were conducted in R and visualizations generated with ggplot2 (Wickham, 2011) package. PCA analyses of water quality parameters were performed with centered and scaled data in R base prcomp(). Linear regression models of log10 (volume normalized 16S rRNA gene copies) against water quality parameters were fit using base R lm(). Pearson correlation between log10(volume normalized 16S rRNA gene copies) and chlorine concentrations (mg/L) was obtained with base R cor() and exponential decay curve with nls(). Euclidian distances between pairwise Mash distance of DW metagenomes was calculated with vegan (Oksanen et al., 2015) function vegdist() and clustered with complete linkage method with base R hclust(). PCoA ordination of Mash distances was performed with ape (Paradis et al., 2004) function pcoa(). Group-wise non parametric testing was performed with R base statistic packages using function kruskal.test() or wilcox.test(), and adjusted p values were obtained based on Benjamini Hochberg correction. Permutational hypothesis testing (n iterations=10,000) of differences in group means between two groups was performed after up-sampling group data to balance observations using upsample() from the groupdata2 (Olsen, 2021) package. Analyses of variance (ANOVA) was performed with aov() and followed up with post hoc Tukey-Kramer testing using TukeyHSD() in base R.

3. Results and discussion

3.1. Bacterial concentrations were associated with water quality parameters, particularly total chlorine concentrations

Water quality parameters were recorded for each sampling location and timepoint (Fig. 2A, Table S1). PCA analyses was conducted to assess whether water chemistry varied spatially and/or temporally. Nitrogen species were excluded from PCA analyses as their concentrations were below detection limit at several locations/timepoints and nitrate

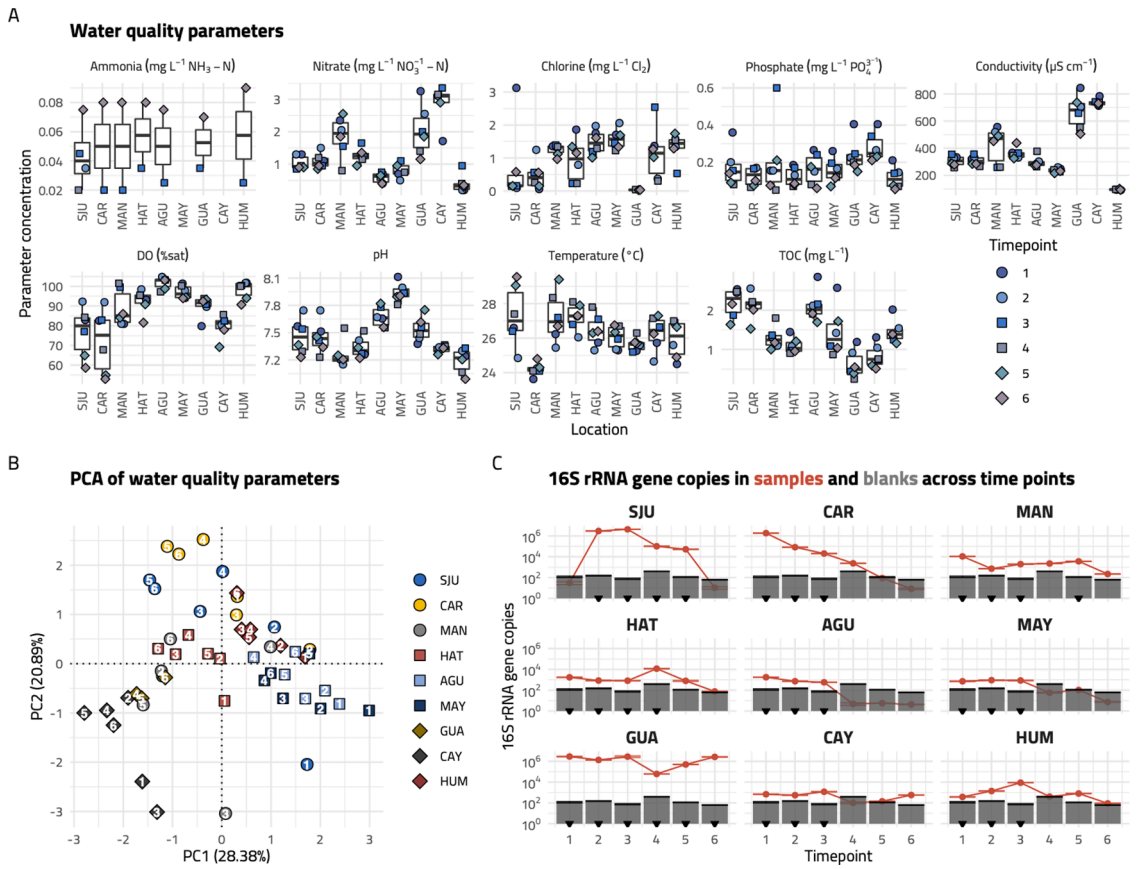

Fig. 2. (A) Concentrations of water quality parameters for nine locations from December 2017 to October 2018. Shapes and colors correspond to timepoints. All nitrite measurements were below detection limit (BDL, not shown), for ammonia 36 measurements were BDL, and for chlorine 3 samples were BDL. (B) Principal component analyses of water quality parameters. Colors and shapes correspond to sampling location. Direct labels of timepoints within sampling location, in white. The PCA does not show a clear clustering between samples or timepoints. (C) Gene copies of 16S rRNA gene for samples and blanks. Red lines correspond to samples and gray bars correspond to blanks. Comparison between samples and blanks guided sample selection for sequencing. Inverted black triangles denote samples that were sequenced.

concentrations also strongly correlated with conductivity (Pearson's $R = 0.88$, $p<0.001$) and phosphate concentrations (Pearson's $R = 0.47$, $p<0.001$). No clear clustering of samples by location or timepoint was observed despite drinking water samples being obtained from variable source waters (Fig. 2B). For instance, CAY, GUA, and MAN had higher mean and variable conductivities compared to other locations, which is consistent with source water type of these locations being groundwater. Phosphate, on the other hand, was relatively narrowly distributed across samples, and in lower concentration compared to other DW systems (Gooddy et al., 2015), with CAY location consistently higher than other PR locations.

We quantified the abundance of 16S rRNA genes in all samples as a measure of bacterial concentrations (Fig. 2C). Volume normalized 16S rRNA gene copies (16S rRNA gene copies mL^{-1}) ranged from 6.4×10^{-2}–9.1×10^{4} copies mL^{-1}. Independently regressing variables in the PCA as descriptors of log10(16S rRNA gene copies mL^{-1}) for each location resulted in six significant ($p<0.05$) linear models. The goodness of fit for all models was relatively high with an average adjusted R^2 of 0.704 ± 0.062, and significant associations between log10(16S rRNA mL^{-1}) and DO in CAR, pH in HUM, temperature in CAR, MAY and GUA, and TOC in AGU. However, regressing parameters against log10(16S rRNA mL^{-1}) for all locations only resulted in statistically significant associations with temperature (adj $R^2 = 0.067$, $p < 0.05$) and total chlorine (adj $R^2 = 0.227$, $p < 0.001$). A multiple linear regression model with all water quality parameters as descriptors of log10(16S rRNA mL^{-1}) ($n = 51$), indicated that total chlorine was the major driver associated with decreasing 16S rRNA gene concentrations (Adj $R^2 = 0.28$, $p<0.001$, Fig. 3A, Table S3). Chlorine concentrations measured in samples were comparable to those reported in other US studies (Stanish et al., 2016), except for GUA where it was either below detection limit (BDL) or very low for all timepoints. All samples, except timepoint 1 in SJU, had total chlorine concentrations below 3 mg L^{-1} (Fig. 3B, Table S1). The negative relationship between bacterial load and chlorine concentration is particularly evident for SJU and CAR, where decreasing total chlorine concentration were associated with increased bacterial concentrations, relatively stable chlorine concentrations correspond to stable bacterial concentrations, and absence of chlorine shows high bacterial concentration, respectively (Fig. 3B). Not surprisingly, these results suggest that maintaining chlorine residual is critically important for ensuring low bacterial concentrations which could be particularly challenging due to infrastructure damage from natural disasters. Some locations did exhibit significant variation in chlorine concentrations between December 2017 and February 2018 (e.g., SJU, CAR, HAT), with some of these variations associated with water main breaks (e.g., CAR in December 2017).

3.2. *Microbial communities and metagenomes of PR samples were similar to those seen in other drinking water systems*

Based on total bacteria qPCR results and comparison to blanks, a select number of samples per location were subjected to metagenomic sequencing ($n = 33$, Fig. 2C, Table S4). The initial three sampling points for all locations were sequenced, unless their 16S rRNA gene copy numbers were below or equal to their matched controls. Further, any

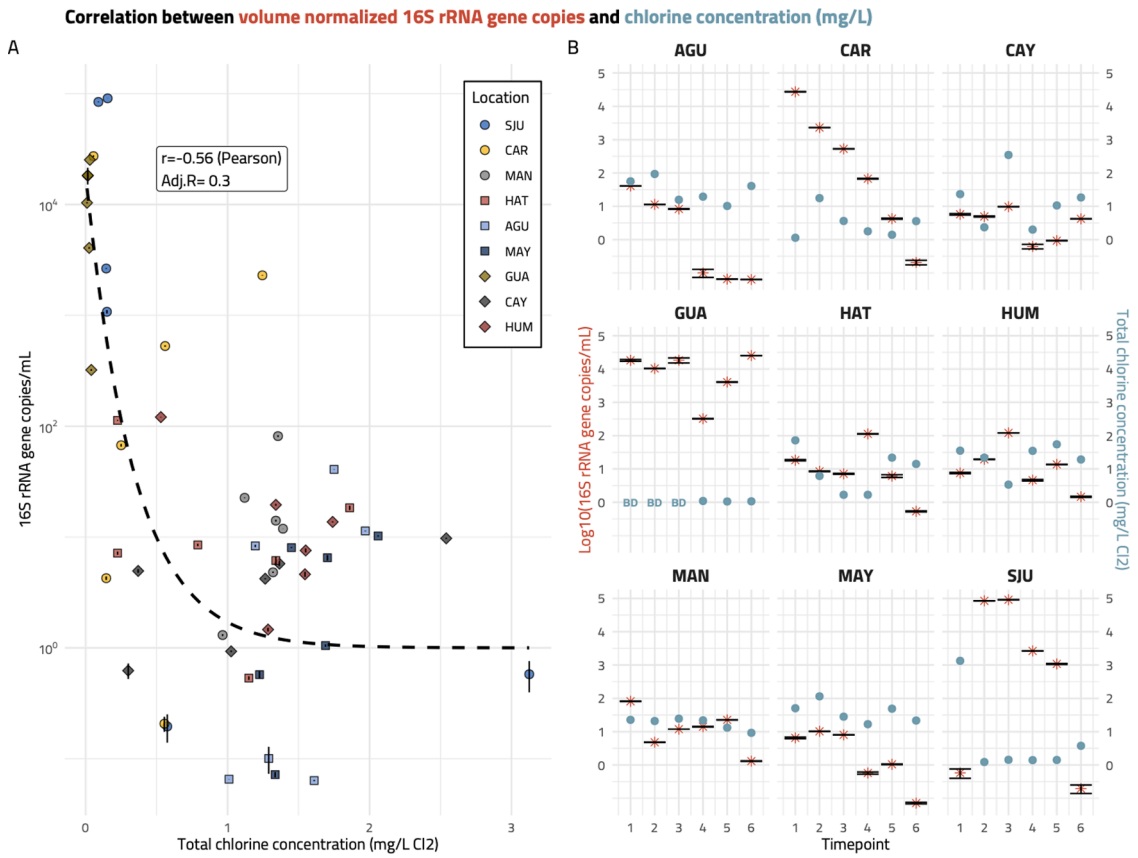

Fig. 3. (A) 16S rRNA gene copies normalized by volume correlated with total chlorine concentration. Colors and shapes correspond to sampling location. Total chlorine concentration was the major driver among all recorded water quality parameters when all samples were aggregated in a multiple linear regression model. The dashed line corresponds to an exponential decay fit characteristic to inactivation of bacteria. (B) Double y- axis plot with the concentration of log 16S rRNA gene copies normalized by volume (left, red) and total chlorine concentration (right, blue) by time point faceted by sampling location. Chlorine concentrations were compliant with EPA regulation and bacterial concentrations were within range of typical drinking water systems. Total chlorine concentration was below detection (BD) at GUA first three timepoints. On a per location basis, clear examples of the negative effect of total chlorine can be observed (i.e., SJU, MAN, and GUA), however other parameters may drive bacterial concentrations in the remaining locations.

other sampling point 10-fold or higher 16S rRNA gene copy numbers than the highest observed in the controls for corresponding timepoint was sequenced as well. A total of 1.18 Gb raw reads were generated after quality filtering and 1.16 Gb reads were not mapped against UniVec, resulting in less than 2.6% of the reads being discarded for the majority of samples ($n = 31$), only 2 samples (i.e. HUM_1, HUM_2) retained less than 92.5% of the raw reads (Table S4). Within sample diversity (Nd), as assessed by Nonpareil curves, ranged from 15.33 to 19.16 (Fig. 4A, Table S5). These indices rely on redundancy of reads in a metagenome to estimate diversity of metagenome, with higher Nd corresponding to more diverse communities. The Nd values observed here are consistent with those seen in other chlorinated drinking water systems (Dai et al., 2020). The spread of observed Nd values within location was larger for GUA, HUM, and HAT, indicating higher temporal variation in diversity, while low variability in Nd values at CAR indicative of low temporal differences. Kruskal-Wallis test of Nd by location reveal significant differences in the median of at least one of the groups ($p<0.05$), however multiple hypothesis correction with Dunn test did not identify any significant pairwise differences. Significant and positive correlations were observed between Nd and pH at AGU ($p<0.05$), and Nd and DO at HUM ($p<0.01$), and significant negative correlations between HAT diversity and log10(16S rRNA mL^{-1}) and CAY diversity and nitrate ($p<0.05$ for both locations). The median nonpareil estimated sample coverage attributed to sequencing effort across all samples was 90% and the range from the samples 70% to 97.9%. There were no time or location specific trends in variable coverage across samples.

The reads were subsequently assembled and scaffolds identified as potential contamination were removed as outlined in the materials and methods section. A summary of statistics for the 9 co-assemblies that were generated can be found in Table S6. CAT was used to annotate true scaffolds (i.e., scaffolds retained post contamination analysis) and coverage information allowed us to obtain per sample profiles (Fig. 4B, Table S7). Bacteria generally constituted the largest portion in the samples from SJU, CAR, GUA, and CAY, with a mean relative abundance (RA) of 96.6 ± 1.72%. In contrast, scaffolds of eukaryotic origin (18.5 ± 19.5%) and unclassified scaffolds (5.62±2.85%) constituted a significant proportion of the community in MAN, HAT, AGU, MAY, and HUM. The RA of eukaryotic contigs is not consistent within locations, suggesting high temporal variation of the eukaryotic fraction in the systems. Despite highest eukaryotic contigs RA at timepoint 1 for AGU and MAY and timepoint 2 at HUM, water quality parameters at these sites had relatively low temporal variation. On the other hand, MAN had its highest RA at timepoint three when DO and phosphate concentrations were higher than mean values. In contrast with other sites with high contribution of eukaryotic contigs, MAN water source is groundwater. Phosphate concentration is associated with source water type and treatment processes (i.e., corrosion control). Douterelo et al. (Douterelo et al., 2018) compared metagenomic samples from sites with different source waters and saw dominance of bacteria and no significant differences in the RA of eukaryotes. Moreover Inkinen et al. (Inkinen et al., 2019) correlated the presence of phosphate concentrations with active eukaryotes in a chlorinated groundwater DWS in contrast with surface waters. At the HAT site, the largest eukaryotic RA was seen at timepoint 1 consistent with highest chlorine concentration. Disinfection is a driver

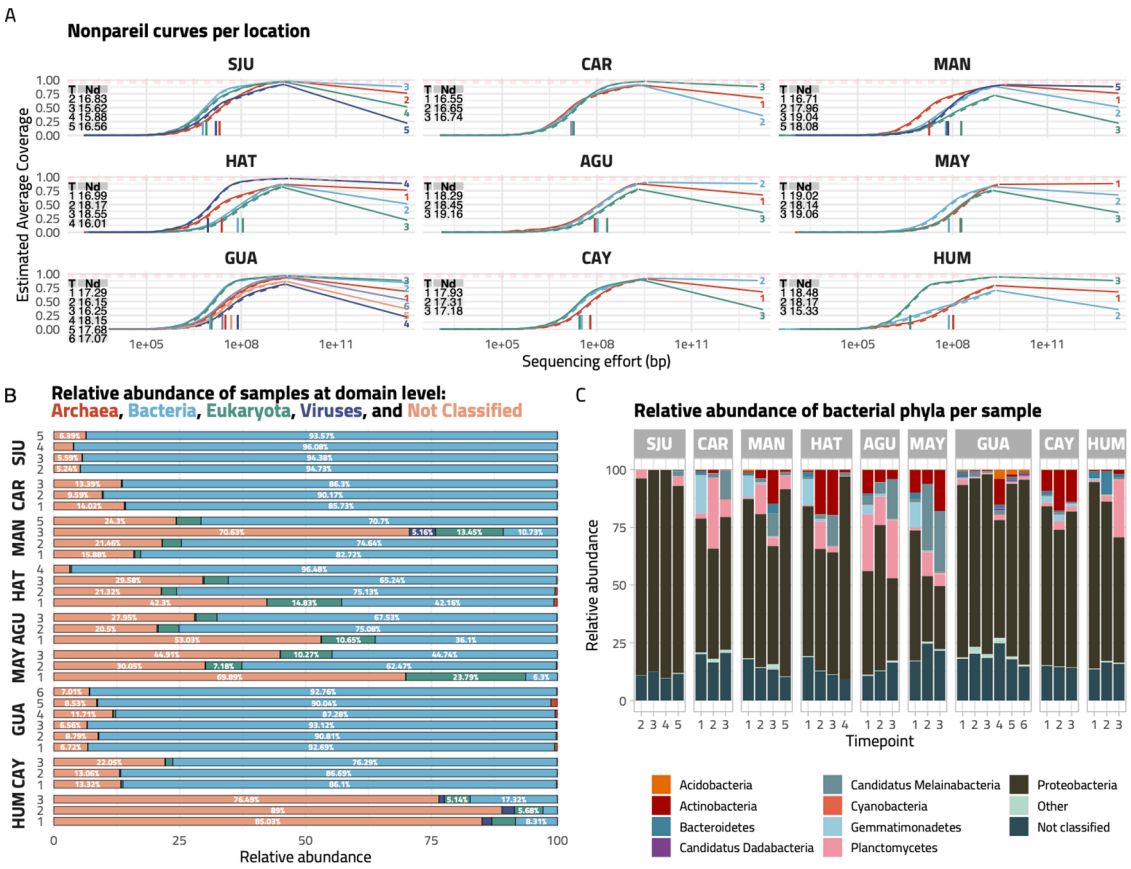

Fig. 4. Diversity of microbial communities in metagenomes. (A) Nonpareil curves faceted by location. Each curve within facet denotes a sample directly labelled by timepoint. Respective Nonpareil diversity (Nd) is shown in table on the top left corner of each facet. The bars correspond to the sample estimated sequencing effort. Nd value indicates within sample community complexity in the sequence space. (B) Domain level relative abundances (RA) calculated with sample coverage information and CAT annotations. RAs greater than 5% are directly labelled. In some locations a large proportion of the samples were not classified, but Bacteria makes up the largest portion of most samples. (C) Relative abundances at Phylum level. Scaffolds classified to Bacterial domain were designated as 100%. Proteobacteria was the dominant phyla in the majority of samples.

of microbial composition in DWS and it likely reduced the contribution of bacterial community in these samples, resulting in an observed increase in the RA of eukaryotes. We further investigate the eukaryotic component of the microbial communities identified through CAT with MetaEuk (Figure S3, Table S8 and S9). MetaEuk based classification indicated that free living amoeba (FLA) capable of supporting intracellular growth of opportunistic pathogens (e.g., *Vermamoeba, Acanthamoeba*, etc.) were transiently detected at low RAs at CAR, MAN, HAT, AGU, MAY, CAY, and HUM (Figure S4, Table S9). Waterborne parasites like Giardia and Cryptosporidium were not detected.

The coverage of scaffolds classified as bacteria was normalized by rpoB gene coverage in the respective samples to assess the bacterial community (Fig. 4C, Table S10). The proportion of bacterial scaffolds not classified beyond the domain level ranged from 9.20 to 24.80% and patterns were consistent within location. Similar to previous studies, Proteobacteria was the dominant phylum in the majority of samples, ranging from 27.31 to 90.07%, with a mean of 64.2 ± 16.6% for all samples. Actinobacteria and Planctomycetes were also detected in all samples. Actinobacteria is another group that is regularly detected in tap water (Hull et al., 2017). The Actinobacterial composition of MAN, HAT, AGU, MAY, and CAY tend to be higher relative to other locations, with an average RA of 9.10±6.95% in these samples, a mean 1.21±2.77% in other samples, and a global 5.27±6.61% RA. Planctomycetes is present at a RA greater than 1% in 72.72% of samples, but is predominant in CAR and AGU, with a mean RA of 13.41±15.08% and 20.58±7.24%, respectively, compared to a global average of 6.25±8.4%. On average, 81.2 ± 9.61% of sample cumulative RA was not classified up to genus level. The dominant classified genera were Bradyrhizobium, Gemmata, Gemmatimonas, Hyphomicrobium, Methylobacterium, Mycobacterium, Novosphingobium, Pseudorhodoplanes, and Sphingomonas. We further compared the metagenomic assemblies recovered from the nine PR samples to other DWS (not impacted by natural disasters, i.e., undisturbed, Table S11), to assess if there were indications of significant deviation that could be attributed to HM. There was no clear clustering of metagenomes as shown by PCoA ordination of pairwise Mash distances including the nine co-assemblies from PR and 52 co-assemblies from other DWS (Fig. 5A). Furthermore, there was no statistical difference between pairwise Mash distances grouped as PR vs other DWS and other DWS vs other DWS using permutational t tests (p>0.3, Fig. 5B). This suggests that the differences in metagenomes between HM impacted and other DWSs are similar to those observed between other DWSs. Additionally, complete linkage clustering indicated that SJU, MAN, HAT, and AGU, and GUA and CAY clustered closely, both within and between each other; CAR did not cluster directly with another location and MAY and HUM were similar, but separate from the rest of the PR locations (Figure S5A). The respective Mash distances of early and late samples clustered identically as the previous analyses when leveraging coverage data and CAT classification of scaffolds to subset scaffolds pertinent to these categories (Figure S5B-C). This indicates that the metagenomes from the samples collected in PR were largely consistent with what would be expected from drinking water samples, irrespective of time of collection (i.e., December 2017 or October 2018).

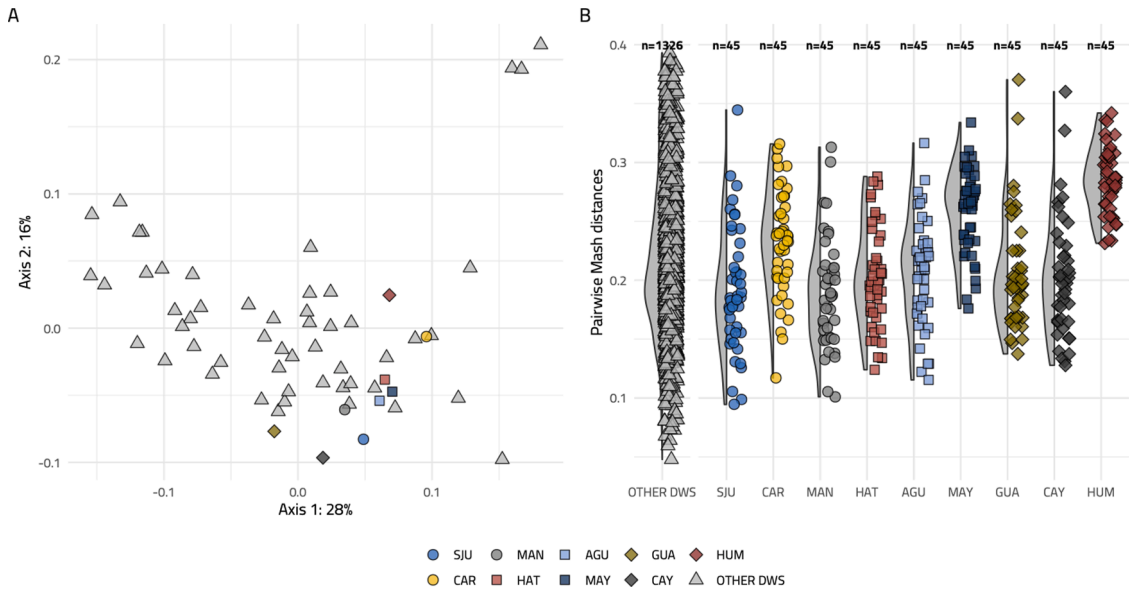

Fig. 5. (A) PCoA ordination of Mash distances between DWS, including PR co-assemblies and reference DWSs. (B) Distribution of Mash distances prior to up sampling used for permutational t-test with group 1: other vs other and group 2: PR vs others. No significant differences were observed ($p>0.3$) between groupings. Colors and shapes correspond to PR locations or reference DWS.

3.3. Opportunistic premise plumbing pathogens were ubiquitous and detected at low concentrations

Genera that contain pathogenic species (i.e., Legionella, Leptospira, Mycobacterium, and Pseudomonas) were further investigated (Fig. 6A) using CAT annotations with additional support from Kraken and/or Kaiju. While monitoring of indicator organisms and residual chlorine is part of the emergency response in the aftermath of hurricanes

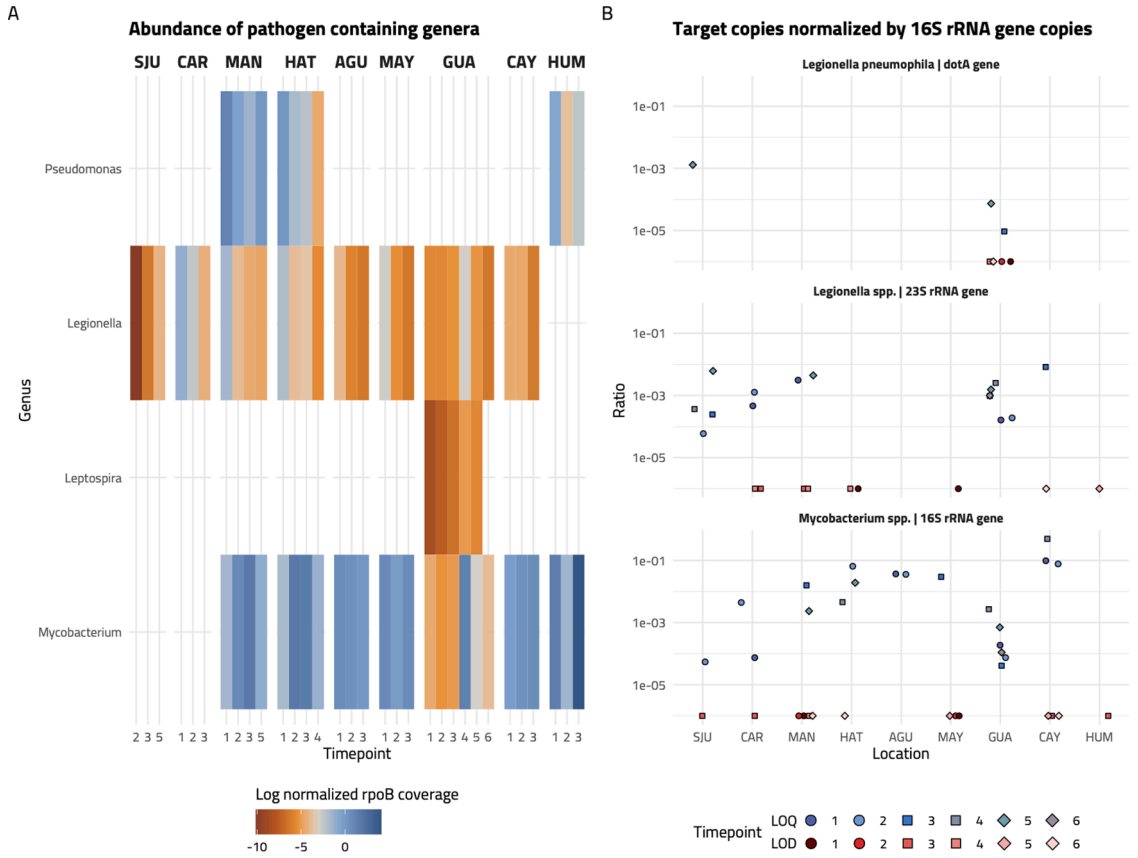

Fig. 6. (A) rpoB normalized coverage of selected genera relevant to drinking water systems. Annotations are based on CAT classification and support from Kaiju or/and Kraken. (B) Ratio of copy numbers of targets (i.e. *Legionella pneumophila, Legionella* spp., and *Mycobacterium* spp.) to 16S rRNA gene copies above LOQ. Colors and shapes correspond to timepoints. *Legionella* spp. and *Mycobacterium* spp. are ubiquitous and more abundant than other targets. Notice the color change for detected, but not quantifiable (LOD, red hues), and quantifiable (LOQ, blue hues) Pseudomonas aeruginosa or Mycobacterium avium were not detected in any samples.

(Patterson and Adams, 2011), challenges with regulatory compliance were common in PR prior to HM and testing laboratories remained non-operational months after the hurricane. Further there was no systematic effort to monitor the prevalence of OPPPs (e.g., Legionella, Pseudomonas, and NTM). Heavy rain and flooding can severely impact water sources and as a result, distribution systems may increase the prevalence of pathogens in drinking water systems, leading to potential health risks.

Previous studies have reported the incidence of waterborne illnesses post hurricanes, including diseases with Legionella, NTM and Leptospira as causative agents (Maness, 2019; Shukla et al., 2018; Sutter and Sosa Pascual, 2018; Walker, 2018). Of the potential OPPP genera, Legionella and Mycobacterium were detected in most locations, while Pseudomonas was consistently detected in MAN, HAT, and HUM. Pathogenic Leptospira was only detected in GUA. There were statistically significant (Wilcoxon test, $p<0.05$) temporal differences in relative abundances of OPPPs for Legionella in MAN, Mycobacterium in HAT, MAY, GUA, and HUM, and Pseudomonas in HAT.

We used qPCR to quantify the abundance of *Legionella spp*. and *Mycobacterium spp*., while also conducting more targeted assays to detect and quantify the abundance of *Legionella pneumophila, Mycobacterium avium*, and *Pseudomonas aeruginosa*, and pathogenic species of the Leptospira genus (Fig. 6B). *Mycobacterium avium* and *Pseudomonas aeruginosa* were not detected in any of the samples. The mean concentration for *Legionella pneumophila, Legionella* spp., and *Mycobacterium* spp. in the samples was 0.71, 7.03, and 1.65 copies mL^{-1}, respectively.

Mycobacterium spp. were observed in all locations with a general frequency of detection of 61.11%. *Mycobacterium* spp. were detected at every sampling timepoint in MAN, GUA, and CAY at very low concentrations (1.17±1.19 copies mL^{-1}), while their concentrations were as high as 10 copies mL^{-1} at CAR and only detected in the first three timepoints. Consistent with metagenomic results, *Legionella* spp. was widely observed across all sampling locations at a 44.44% frequency of detection, with highest concentrations observed in SJU, CAR, and GUA. *Legionella* spp. concentrations decreased from 5 and 12 copies mL^{-1} at SJU and CAR, respectively to non-detects from December 2017 to October 2018. *Legionella* spp. thrive in warmer temperatures (Lesnik et al., 2016), such as those in PR. Interestingly, concentrations *Legionella* spp. and *Mycobacterium* spp. are several orders of magnitude lower that what has been published in literature (Huang et al., 2021; Isaac and Sherchan, 2020; Ley et al., 2020; Liu et al., 2019) (i.e., $1–10^4$ copies mL^{-1}). A potential reason for this could be over-chlorination in the systems, which had been reported in the aftermath of HM (Brown et al., 2018), including in early phase of sampling as this study showed.

The dotA gene assay to target *Legionella pneumophila* revealed low concentrations (i.e., 0.72±0.61 copies mL^{-1} for SJU and GUA locations. However, in SJU, L. *pneumophila* was detected only in timepoint 5, while being detected at GUA at all timepoints. This is consistent with observations from other DWS where *Legionella pneumophila* was detected at low frequency and low concentrations (Lu et al., 2016; Wang et al., 2012).

The LipL32 gene of pathogenic Leptospira was not quantifiable, but detected only in GUA and at timepoints 1, and 4 (i.e., 3.7% frequency). The detection of Leptospira at this location was consistent with the detection of the genus Leptospira using metagenomics. Leptospira is not routinely reported in DWS, apart from the recent study by Keenum et al. (Keenum et al., 2021), possibly due to the efficacy of routine disinfection practices in the elimination of this pathogen (Wynwood et al., 2014). However, its importance has been highlighted in rivers and creeks when used as drinking water without proper treatment, particularly in situations of water scarcity, such as hurricanes (Keenum et al., 2021; Truitt et al., 2020). The presence of Leptospira in GUA is possibly exacerbated by the absence of residual chlorine at this location.

3.4. A small fraction of recovered metagenome assembled genomes were associated with pathogens

Metagenomes were co-assembled by location, binned, and manually refined with anvi'o. 105 bacterial MAGs were recovered after dereplication with dRep and quality filtering for completeness greater than 50% and percent redundancy lower than 10%. We identified one or more 16S rRNA in 39% of the MAGs. Further, we compared the differences in abundances between samples by accounting for MAGs read recruitment. Of these, 37% of the MAGs were detected in a quarter or more of the samples (Fig. 7, Table S12). The PR MAGs were shared homology with 4.55% of a recently published JGI MAG collection based on ANI values ranging from 74.65 to 99.49%. JGI MAGs ecosystem categories represent aquatic (36.75%), human (31.31%), terrestrial (6.5%), built environment (5.03%), and wastewater (5%) environments. In contrast, the ecosystem distribution of the PR MAGs pairwise comparisons with JGI MAGs was comprised of aquatic (32.49%), terrestrial (20.57%), built environment (13.64%), plants (12.5%), and lab enrichment (4.95%) habitats. The aquatic ecosystem had the largest number of same species representation (20% of PR MAGs, Fig. 7) with species boundaries level set at 83% cutoff threshold (Jain et al., 2018), and included 2 JGI MAGs also recovered from DWS.

Despite this observation, the JGI MAG dataset suffers from lack of representation of MAGs assembled from DWS habitats ($n = 7$). Therefore, a complimentary approach was used by mapping metagenomic reads from diverse ecosystems against the MAGs assembled in this study using the SearchSRA tool. This analysis indicated that the aquatic ecosystem was found to be the top environmental association for 63.8% of our recovered MAGs, the other top ecosystems were terrestrial (21.9%), human (13.3%), and mammal (1.0%) associated environments (Fig. 7). However, if we consider the top four environments, aquatic ecosystem category is represented in all of our MAGs. There were statistically significant differences (ANOVA, $p<0.001$) between the proportion of reads mapping from each ecosystem category mapping to the PR MAGs, the only pairwise comparisons that were not statistically significant were terrestrial vs aquatic (Tukey's, $p = 0.96$) and mammals vs human environments (Tukey's, $p = 0.06$). Additionally, more than 13% of the aquatic ecosystem metagenomes were associated with DWS. Altogether, these analyses show that our MAGs are widely distributed in the environment, but are largely associated with aquatic and DWS associated environments.

The classification of resulting representative MAGs consisted of 64.8% Proteobacteria, followed by 14.3% Cyanobacteria, 12.4% Planctomycetota, and 3.81% Actinobacteriota. All of the Actinobacteria MAGs were classified as Mycobacterium. More than 50% of the MAGs were not classified to genus level. The most abundant genera among the MAGs included Hyphomicrobium ($n = 6$), Bradyrhizobium ($n = 4$), Gemmata ($n = 4$), Mycobacterium ($n = 4$), and Porphyrobacter ($n = 4$). There was no relationship between environmental parameters and MAG abundance as assessed by Mantel statistic ($r = 0.133$, $p > 0.05$) or constrained redundancy analyses. Three of the four Mycobacterium MAGs were classified up to species level and correspond to *Mycobacterium gordonae, Mycobacterium paragordonae*, and *Mycobacterium phocaicum,* all of which have been recovered from drinking water systems previously and are associated with infections in immunocompromised individuals (Shachor-Meyouhas et al., 2014). M. *gordonae* was more abundant (RPKM=2.27±2.77) and frequently detected (58%) than M. *paragordonae* (RPKM=0.77±0.41, 42% detection). Nevertheless, their abundance and frequency of detection was higher than for M. *phocaicum* (RPKM=1.8 ± 1.9, 15% detection). Mycobacterium MAGs were not detected from SJU, and infrequently detected at CAR, GUA, and HUM. A single Pseudomonas MAG was recovered and classified as *Pseudomonas alcaligenes*. This MAG was only detected once and at the first timepoint at CAR, HAT, AGU, and MAY. At the MAN location, it was detected in the initial and final timepoint and at higher abundance at the final timepoint. *Pseudomonas alcaligenes* carries multiple antibiotic resistance

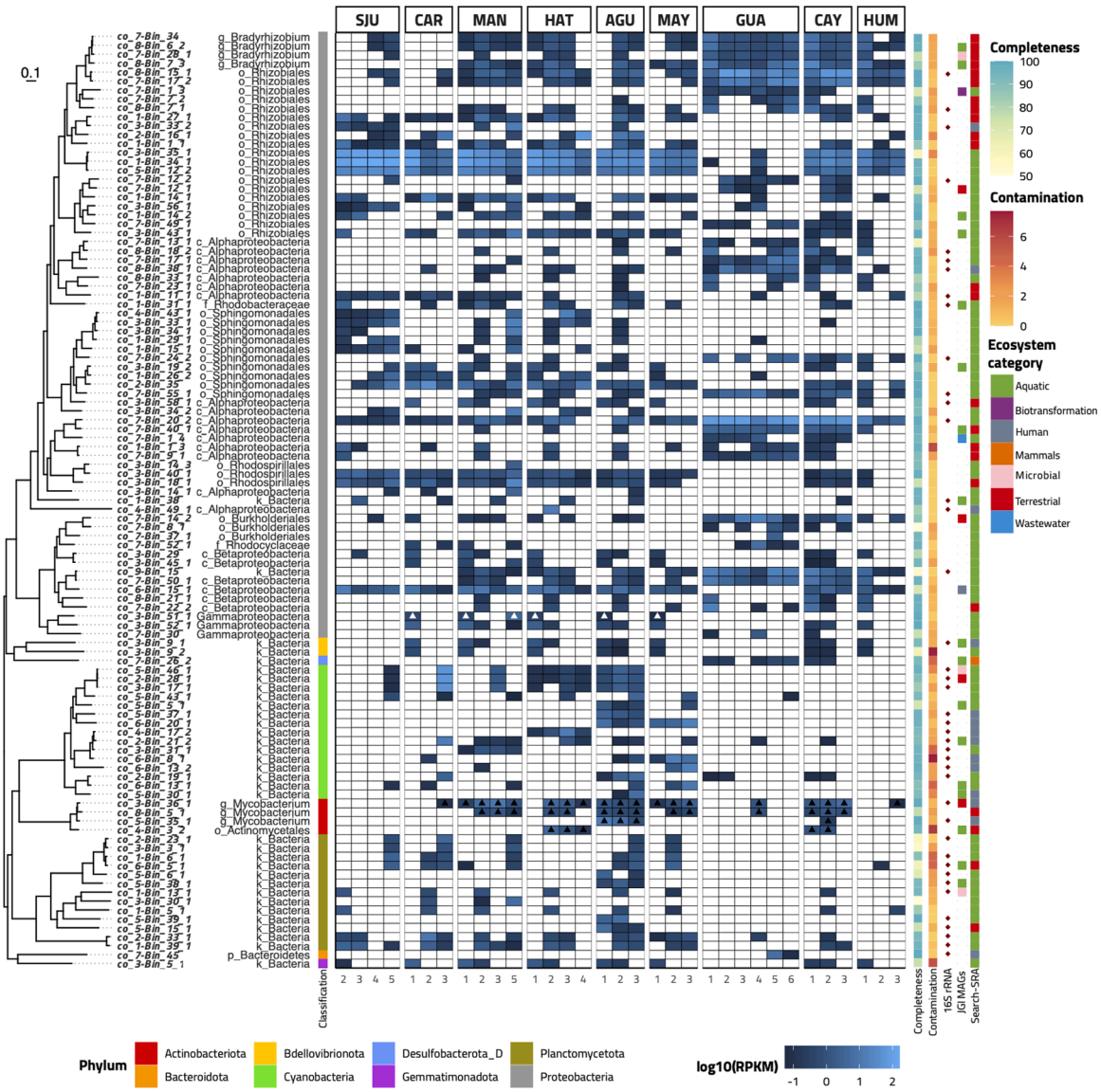

Fig. 7. Metagenomic assembled genomes (MAG) information. From left to right: Phylogenomic tree of 105 recovered MAGs. Each branch is labelled by MAG code (italics) and marker lineage of MAG classified by GTDB-tk and annotated by color of phylum. Heatmap rows correspond to respective MAG and color gradient denotes log10(RPKM) values of MAGs detected in samples. A MAG was considered detected if $\geq$ 25% of its bases were covered by at least one read from the corresponding sample. Presence of color denotes detection of MAG and the blue gradient becomes lighter as abundance (i.e., RPKM) increases. Triangles in the heatmap highlight *Pseudomonas* spp. (white) and *Mycobacterium* spp. (black). MAG completeness ranges from 50 to 100% with color gradient from pale green to blue increasing value and contamination ranges from 0 to 10% with intense red corresponding to higher values. Red diamonds correspond to MAGs with one or more 16S rRNA gene detected. The highest ANI value between JGI MAGs and corresponding MAG from this study is colored according to ecosystem categories from JGI MAGs, if blank, no ANI above 83% was observed for that particular MAG. SearchSRA top environmental niche is depicted using the same color legend.

genes, are considered opportunistic human pathogens and have been identified in previous literature characterizing drinking water systems, particularly in chlorinated systems (Jia et al., 2019; Ma et al., 2019).

4. Conclusion

This study characterized the microbial communities of nine locations in the aftermath of severe hurricanes (i.e., Irma and Maria) in a spatial-temporal yearlong survey using targeted and non-targeted molecular methods. Our results highlight that maintaining a disinfectant residual helps manage microbial concentration at the taps, yet sampling locations showed significant variation in the earlier timepoints. The estimated bacterial concentrations based on 16S rRNA gene abundance at the sampling locations were consistent with literature established values characterizing DWSs and generally decreased over time. Additionally, members of the microbial community were comparable to those found in other DWSs which were not impacted by natural disasters. Regardless of the ubiquity of some targeted OPPPs, such as *Legionella spp*. and *Mycobacterium spp*., they were present at low concentrations. Interestingly, pathogenic Leptospira was only detected at a single location and its presence could be associated with a lack of disinfectant residual at that site. A small fraction of metagenome assembled genomes were associated with potential pathogens, and other recovered MAGs represent previously reported taxa routinely found in drinking water systems. Altogether, the water disruptions (i.e., no water or intermittent supply) that were sustained after HM did not have a significant impact on the microbiological quality of drinking water in our study sites.

Declaration of Competing Interest

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

Data availability

Metagenomic data is available on NCBI at Bioproject number: PRJNA718649 and the co-assemblies and metagenome assembled genomes are available on figshare at: https://doi.org/10.6084/m9.figshare.c.5414964.

Acknowledgements

This study is supported by the United States National Science Foundation (NSF, CBET-1829754, CBET-1832756, IIS-1546428), the National Institute of Environmental Health Sciences (NIEHS) grants P42ES017198 and P50ES026049, and U.S. Environmental Protection Agency (EPA) grant R83615501. The authors acknowledge and extend their sincerest gratitude towards Lilliana Gonzalez, Jesus Lee-Borges, Perla Torres, Vibha Bansal, and Ezio Fasoli for their collaboration in sampling logistics.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.wroa.2021.100123.

References

Allaire, M., Wu, H., Lall, U., 2018. National trends in drinking water quality violations 115, 1–6. https://doi.org/10.1073/pnas.1719805115 https://doi.org/.

Alneberg, J., Bjarnason, B.S., Bruijn, I.De, Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F., Quince, C., 2014. Binning metagenomic contigs by coverage and composition 11, 1144–1146. https://doi.org/10.1038/nmeth.3103 https://doi.org/.

Anuj, S.N., Whiley, D.M., Kidd, T.J., Bell, S.C., Wainwright, C.E., Nissen, M.D., Sloots, T. P., 2009. Identification of Pseudomonas aeruginosa by a duplex real-time polymerase chain reaction assay targeting the ecfX and the gyrB genes. Diagn. Microbiol. Infect. Dis. 63, 127–131. https://doi.org/10.1016/j.diagmicrobio.2008.09.018 https://doi.org/.

ASCE, 2017. Infrastructure Report Card: A comprehensive assesment of America's infrastructure.

Ashbolt, N.J., 2015. Microbial Contamination of Drinking Water and Human Health from Community Water Systems 95–106. https://doi.org/10.1007/s40572-014-0037-5 https://doi.org/.

Bautista-de los Santos, Q.M., Chavarria, K.A., Nelson, K.L., 2019. Understanding the impacts of intermittent supply on the drinking water microbiome. Curr. Opin. Biotechnol. 57, 167–174. https://doi.org/10.1016/j.copbio.2019.04.003 https://doi.org/.

Brown, P., Ve, C.M., Murphy, C.B., Welton, M., Torres, H., Rosario, Z., Alshawabkeh, A., Cordero, F., Padilla, I.Y., Meeker, J.D., 2018. Hurricanes and the Environmental Justice Island : irma and Maria in Puerto Rico 11, 148–153. https://doi.org/10.1089/env.2018.0003.

Buchfink, B., Xie, C., Huson, D.H., 2014. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60. https://doi.org/10.1038/nmeth.3176 https://doi.org/.

Bushnell, B., 2015. BBMap [WWW Document]. https://sourceforge.net/projects/bbmap/.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight, R., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. U. S. A. 108, 4516–4522. https://doi.org/10.1073/pnas.1000080107 https://doi.org/.

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560 https://doi.org/.

Chern, E.C., King, D., Haugland, R., Pfaller, S., 2015. Evaluation of quantitative polymerase chain reaction assays targeting Mycobacterium avium, M. intracellulare, and M. avium subspecies paratuberculosis in drinking water biofilms. J. Water Health 13, 131–139. https://doi.org/10.2166/wh.2014.060 https://doi.org/.

Dai, Z., Sevillano-Rivera, M., Calus, S., Bautista-de los Santos, Q.M., Eren, A.M., van der Wielen, P.W.J.J., Ijaz, U.Z., Pinto, A., 2020. Disinfection exhibits systematic impacts on the drinking water microbiome. Microbiome.

Douterelo, I., Calero-Preciado, C., Soria-Carrasco, V., Boxall, J.B., 2018. Whole metagenome sequencing of chlorinated drinking water distribution systems. Environ. Sci. Water Res. Technol. 4, 2080–2091. https://doi.org/10.1039/c8ew00395e https://doi.org/.

Eren, A.M., Esen, C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., Delmont, T.O., 2015. Anvi ' o : an advanced analysis and visualization platform for ' omics data 1–29. https://doi.org/10.7717/peerj.1319 https://doi.org/.

Estrada, F., Botzen, W.J.W., Tol, R.S.J., 2015. with an influence from climate change 8, 6–11. https://doi.org/10.1038/NGEO2560 https://doi.org/.

Exum, N.G., Betanzo, E., Schwab, K.J., Chen, T.Y.J., Guikema, S., Harvey, D.E., Harvey, D.E., 2018. Extreme Precipitation, Public Health Emergencies, and Safe Drinking Water in the USA 305–315.

Gooddy, D.C., Lapworth, D.J., Ascott, M.J., Bennett, S.A., Heaton, T.H.E., Surridge, B.W. J., 2015. Isotopic Fingerprint for Phosphorus in Drinking Water Supplies. Environ. Sci. Technol. 49, 9020–9028. https://doi.org/10.1021/acs.est.5b01137 https://doi.org/.

Goodess, C.M., 2012. How is the frequency, location and severity of extreme events likely to change up to 2060 ? Environ. Sci. Policy 27, S4–S14. https://doi.org/10.1016/j.envsci.2012.04.001 https://doi.org/.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086 https://doi.org/.

Huang, J., Chen, S., Ma, X., Yu, P., Zuo, P., Shi, B., Wang, H., Alvarez, P.J.J., 2021. Opportunistic pathogens and their health risk in four full-scale drinking water treatment and distribution systems. Ecol. Eng. 160 https://doi.org/10.1016/j.ecoleng.2020.106134 https://doi.org/.

Hull, N.M., Holinger, E.P., Ross, K.A., Robertson, C.E., Harris, J.K., Stevens, M.J., Pace, N.R., 2017. Longitudinal and Source-to-Tap New Orleans, LA, U.S.A. Drinking Water Microbiology. Environ. Sci. Technol. 51, 4220–4229. https://doi.org/10.1021/acs.est.6b06064 https://doi.org/.

Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11. https://doi.org/10.1186/1471-2105-11-119 https://doi.org/.

Inkinen, J., Jayaprakash, B., Siponen, S., Hokajärvi, A.M., Pursiainen, A., Ikonen, J., Ryzhikov, I., Täubel, M., Kauppinen, A., Paananen, J., Miettinen, I.T., Torvinen, E., Kolehmainen, M., Pitkänen, T., 2019. Active eukaryotes in drinking water distribution systems of ground and surface waterworks. Microbiome 7, 1–17. https://doi.org/10.1186/s40168-019-0715-5 https://doi.org/.

Isaac, T.S., Sherchan, S.P., 2020. Molecular detection of opportunistic premise plumbing pathogens in rural Louisiana's drinking water distribution system. Environ. Res. 181, 108847 https://doi.org/10.1016/j.envres.2019.108847 https://doi.org/.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., Aluru, S., 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9, 1–8. https://doi.org/10.1038/s41467-018-07641-9 https://doi.org/.

Jia, S., Wu, J., Ye, L., Zhao, F., Li, T., Zhang, X.X., 2019. Metagenomic assembly provides a deep insight into the antibiotic resistome alteration induced by drinking water chlorination and its correlations with bacterial host changes. J. Hazard. Mater. 379, 120841 https://doi.org/10.1016/j.jhazmat.2019.120841 https://doi.org/.

Jiang, S.C., Han, M., Chandrasekaran, S., Fang, Y., Kellogg, C.A., 2020. Assessing the water quality impacts of two Category-5 hurricanes on St. Thomas, Virgin Islands. Water Res. 171, 115440 https://doi.org/10.1016/j.watres.2019.115440 https://doi.org/.

Keenum, I., Medina, M.C., Garner, E., Pieper, K.J., Blair, M.F., Milligan, E., Pruden, A., Ramirez-Toro, G., Rhoads, W.J., 2021. Source-to-tap assessment of microbiological water quality in small rural drinking water systems in puerto rico six months after hurricane maria. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.0c08814 https://doi.org/.

Kishore, N., Marqués, D., Mahmud, A., Kiang, M.V., Rodriguez, I., Fuller, A., Ebner, P., Sorensen, C., Racy, F., Lemery, J., Maas, L., Leaning, J., Irizarry, R.A., Balsari, S., Buckee, C.O., 2018. Mortality in Puerto Rico after Hurricane Maria. N. Engl. J. Med. 379, 162–170. https://doi.org/10.1056/NEJMsa1803972 https://doi.org/.

Landsman, M.R., Rowles, L.S., Brodfuehrer, S.H., Maestre, J.P., Kinney, K.A., Kirisits, M. J., Lawler, D.F., Katz, L.E., 2019. Impacts of hurricane harvey on drinking water quality in two Texas cities. Environ. Res. Lett. 14 https://doi.org/10.1088/1748-9326/ab56fb https://doi.org/.

Lesnik, R., Brettar, I., Höfle, M.G., 2016. Legionella species diversity and dynamics from surface reservoir to tap water: from cold adaptation to thermophily. ISME J. 10, 1064–1080. https://doi.org/10.1038/ismej.2015.199 https://doi.org/.

Levy Karin, E., Mirdita, M., Söding, J., 2020. MetaEuk-sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. Microbiome 8, 1–15. https://doi.org/10.1186/s40168-020-00808-x https://doi.org/.

Ley, C.J., Proctor, C.R., Singh, G., Ra, K., Noh, Y., Odimayomi, T., Salehi, M., Julien, R., Mitchell, J., Nejadhashemi, A.P., Whelton, A.J., Aw, T.G., 2020. Drinking water microbiology in a water-efficient building: stagnation, seasonality, and physicochemical effects on opportunistic pathogen and total bacteria proliferation. Environ. Sci. Water Res. Technol. 6, 2902–2913. https://doi.org/10.1039/d0ew00334d https://doi.org/.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM 00 1–3.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 https://doi.org/.

Lin, Y., Sevillano-Rivera, M., Jiang, T., Li, G., Cotto, I., Vosloo, S., Carpenter, C.M.G., Larese-Casanova, P., Giese, R.W., Helbling, D.E., Padilla, I.Y., Rosario-Pabón, Z., Vélez Vega, C., Cordero, J.F., Alshawabkeh, A.N., Pinto, A., Gu, A.Z., 2020. Impact of

Hurricane Maria on Drinking Water Quality in Puerto Rico. Environ. Sci. Technol. 54, 9495–9509. https://doi.org/10.1021/acs.est.0c01655 https://doi.org/.
Liu, L., Xing, X., Hu, C., Wang, H., 2019. One-year survey of opportunistic premise plumbing pathogens and free-living amoebae in the tap-water of one northern city of China. J. Environ. Sci. (China) 77, 20–31. https://doi.org/10.1016/j.jes.2018.04.020 https://doi.org/.
Lu, J., Struewing, I., Vereen, E., Kirby, A.E., Levy, K., Moe, C., Ashbolt, N., 2016. Molecular Detection of Legionella spp. and their associations with Mycobacterium spp., Pseudomonas aeruginosa and amoeba hosts in a drinking water distribution system. J. Appl. Microbiol. 120, 509–521. https://doi.org/10.1111/jam.12996 https://doi.org/.
Ma, L., Li, B., Zhang, T., 2019. New insights into antibiotic resistome in drinking water and management perspectives: a metagenomic based study of small-sized microbes. Water Res. 152, 191–201. https://doi.org/10.1016/j.watres.2018.12.069 https://doi.org/.
Maness, L.R., 2019. The Effect of Hurricanes on Pathogenic Diseases. J. Environ. Health 81, 16–20.
Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat. Commun. 7, 1–9. https://doi.org/10.1038/ncomms11257 https://doi.org/.
Nayfach, S., Roux, S., Seshadri, R., Udwary, D., Varghese, N., Schulz, F., Wu, D., Paez-Espino, D., Chen, I.M., Huntemann, M., Palaniappan, K., Ladau, J., Mukherjee, S., Reddy, T.B.K., Nielsen, T., Kirton, E., Faria, J.P., Edirisinghe, J.N., Henry, C.S., Jungbluth, S.P., Chivian, D., Dehal, P., Wood-Charlson, E.M., Arkin, A.P., Tringe, S. G., Visel, A., Abreu, H., Acinas, S.G., Allen, E., Allen, M.A., Andersen, G., Anesio, A. M., Attwood, G., Avila-Magaña, V., Badis, Y., Bailey, J., Baker, B., Baldrian, P., Barton, H.A., Beck, D.A.C., Becraft, E.D., Beller, H.R., Beman, J.M., Bernier-Latmani, R., Berry, T.D., Bertagnolli, A., Bertilsson, S., Bhatnagar, J.M., Bird, J.T., Blumer-Schuette, S.E., Bohannan, B., Borton, M.A., Brady, A., Brawley, S.H., Brodie, J., Brown, S., Brum, J.R., Brune, A., Bryant, D.A., Buchan, A., Buckley, D.H., Buongiorno, J., Cadillo-Quiroz, H., Caffrey, S.M., Campbell, A.N., Campbell, B., Carr, S., Carroll, J.L., Cary, S.C., Cates, A.M., Cattolico, R.A., Cavicchioli, R., Chistoserdova, L., Coleman, M.L., Constant, P., Conway, J.M., Mac Cormack, W.P., Crowe, S., Crump, B., Currie, C., Daly, R., Denef, V., Denman, S.E., Desta, A., Dionisi, H., Dodsworth, J., Dombrowski, N., Donohue, T., Dopson, M., Driscoll, T., Dunfield, P., Dupont, C.L., Dynarski, K.A., Edgcomb, V., Edwards, E.A., Elshahed, M. S., Figueroa, I., Flood, B., Fortney, N., Fortunato, C.S., Francis, C., Gachon, C.M.M., Garcia, S.L., Gazitua, M.C., Gentry, T., Gerwick, L., Gharechahi, J., Girguis, P., Gladden, J., Gradoville, M., Grasby, S.E., Gravuer, K., Grettenberger, C.L., Gruninger, R.J., Guo, J., Habteselassie, M.Y., Hallam, S.J., Hatzenpichler, R., Hausmann, B., Hazen, T.C., Hedlund, B., Henny, C., Herfort, L., Hernandez, M., Hershey, O.S., Hess, M., Hollister, E.B., Hug, L.A., Hunt, D., Jansson, J., Jarett, J., Kadnikov, V.V., Kelly, C., Kelly, R., Kelly, W., Kerfeld, C.A., Kimbrel, J., Klassen, J.L., Konstantinidis, K.T., Lee, L.L., Li, W.J., Loder, A.J., Loy, A., Lozada, M., MacGregor, B., Magnabosco, C., Maria da Silva, A., McKay, R.M., McMahon, K., McSweeney, C.S., Medina, M., Meredith, L., Mizzi, J., Mock, T., Momper, L., Moran, M.A., Morgan-Lang, C., Moser, D., Muyzer, G., Myrold, D., Nash, M., Nesbø, C.L., Neumann, A.P., Neumann, R.B., Noguera, D., Northen, T., Norton, J., Nowinski, B., Nüsslein, K., O'Malley, M.A., Oliveira, R.S., Maia de Oliveira, V., Onstott, T., Osvatic, J., Ouyang, Y., Pachiadaki, M., Parnell, J., Partida-Martinez, L. P., Peay, K.G., Pelletier, D., Peng, X., Pester, M., Pett-Ridge, J., Peura, S., Pjevac, P., Plominsky, A.M., Poehlein, A., Pope, P.B., Ravin, N., Redmond, M.C., Reiss, R., Rich, V., Rinke, C., Rodrigues, J.L.M., Rossmassler, K., Sackett, J., Salekdeh, G.H., Saleska, S., Scarborough, M., Schachtman, D., Schadt, C.W., Schrenk, M., Sczyrba, A., Sengupta, A., Setubal, J.C., Shade, A., Sharp, C., Sherman, D.H., Shubenkova, O.V., Sierra-Garcia, I.N., Simister, R., Simon, H., Sjöling, S., Slonczewski, J., Correa de Souza, R.S., Spear, J.R., Stegen, J.C., Stepanauskas, R., Stewart, F., Suen, G., Sullivan, M., Sumner, D., Swan, B.K., Swingley, W., Tarn, J., Taylor, G.T., Teeling, H., Tekere, M., Teske, A., Thomas, T., Thrash, C., Tiedje, J., Ting, C.S., Tully, B., Tyson, G., Ulloa, O., Valentine, D.L., Van Goethem, M.W., VanderGheynst, J., Verbeke, T.J., Vollmers, J., Vuillemin, A., Waldo, N.B., Walsh, D. A., Weimer, B.C., Whitman, T., van der Wielen, P., Wilkins, M., Williams, T.J., Woodcroft, B., Woolet, J., Wrighton, K., Ye, J., Young, E.B., Youssef, N.H., Yu, F.B., Zemskaya, T.I., Ziels, R., Woyke, T., Mouncey, N.J., Ivanova, N.N., Kyrpides, N.C., Eloe-Fadrosh, E.A., 2020. A genomic catalog of Earth's microbiomes Nat. Biotechnol. 39. https://doi.org/10.1038/s41587-020-0718-6. https://doi.org/.
Nazarian, E.J., Bopp, D.J., Saylors, A., Limberger, R.J., Musser, K.A., 2008. Design and implementation of a protocol for the detection of Legionella in clinical and environmental samples. Diagn. Microbiol. Infect. Dis. 62, 125–132. https://doi.org/10.1016/j.diagmicrobio.2008.05.004 https://doi.org/.
NOAA NCEI, 2020. Billion-Dollar Weather and Climate Disasters [WWW Document]. https://doi.org/10.25921/stkw-7w73 https://doi.org/.
NRDC, 2017. Threats on Tap: Drinking Water Violations in Puerto Rico.
Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. MetaSPAdes: a new versatile metagenomic assembler. Genome Res. 27, 824–834. https://doi.org/10.1101/gr.213959.116 https://doi.org/.
Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2015. OK-Package 'vegan.' Community Ecol. Packag. version.
Olm, M.R., Brown, C.T., Brooks, B., Banfield, J.F., 2017. dRep : a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication 11, 2864–2868. https://doi.org/10.1038/ismej.2017.126 https://doi.org/.
Olsen, L.R., 2021. groupdata2: Creating Groups from Data.
Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., Phillippy, A.M., 2016. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol 17, 1–14. https://doi.org/10.1186/s13059-016-0997-x https://doi.org/.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412 https://doi.org/.
Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., Hugenholtz, P., 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat. Biotechnol. 36, 996. https://doi.org/10.1038/nbt.4229 https://doi.org/.
Patterson, C.L., Adams, J.Q., 2011. Emergency response planning to reduce the impact of contaminated drinking water during natural disasters 5, 341–349. https://doi.org/10.1007/s11707-011-0196-8 https://doi.org/.
Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033 https://doi.org/.
R Development Core Team, 2016. R: a language and environment for statistical computing. R Found. Stat. Comput. https://doi.org/10.1017/CBO9781107415324.004 https://doi.org/.
Radomski, N., Lucas, F.S., Moilleron, R., Cambau, E., Haenn, S., Moulin, L., 2010. Development of a real-time qpcr method for detection and enumeration of Mycobacterium spp. in surface water. Appl. Environ. Microbiol. 76, 7348–7351. https://doi.org/10.1128/AEM.00942-10 https://doi.org/.
Rice, P., Longden, L., Bleasby, A., 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. 16, 276–277. https://doi.org/10.1016/S0168-9525(00)02024-2 https://doi.org/.
Rodriguez-R, L.M., Gunturu, S., Tiedje, J.M., Cole, J.R., Konstantinidis, K.T., 2018. Nonpareil 3: fast estimation of metagenomic coverage and sequence diversity. mSystems 3, 1–9. https://doi.org/10.1128/msystems.00039-18 https://doi.org/.
Schwab, K.J., Gibson, K.E., Williams, D.L., Kulbicki, K.M., Lo, C.P., Mihalic, J.N., Breysse, P.N., Curriero, F.C., Geyh, A.S., 2007. Microbial and chemical assessment of regions within New Orleans, LA impacted by Hurricane Katrina. Environ. Sci. Technol. 41, 2401–2406. https://doi.org/10.1021/es062916x https://doi.org/.
Shachor-Meyouhas, Y., Geffen, Y., Arad-Cohen, N., Zaidman, I., Ben-Barak, A., Davidson, S., Kassis, I., 2014. Mycobacterium phocaicum bacteremia: an emerging infection in pediatric hematology-oncology patients. Pediatr. Infect. Dis. J. 33, 1299–1301. https://doi.org/10.1097/INF.0000000000000477 https://doi.org/.
Shaffer, M., Borton, M.A., McGivern, B.B., Zayed, A.A., La Rosa, S.L., Solden, L.M., Liu, P., Narrowe, A.B., Rodríguez-Ramos, J., Bolduc, B., Gazitúa, M.C., Daly, R.A., Smith, G.J., Vik, D.R., Pope, P.B., Sullivan, M.B., Roux, S., Wrighton, K.C., 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. Nucleic Acids Res. 48, 8883–8900. https://doi.org/10.1093/nar/gkaa621 https://doi.org/.
Shukla, M.A., Woc-Colburn, L., Weatherhead, J.E., 2018. Infectious diseases in the aftermath of hurricanes in the United States. Curr. Trop. Med. Reports 5, 217–223. https://doi.org/10.1007/s40475-018-0162-6 https://doi.org/.
Stanish, L.F., Hull, N.M., Robertson, C.E., Kirk Harris, J., 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, 1–21. https://doi.org/10.1371/journal.pone.0157966 https://doi.org/.
Stewart, C.A., Hancock, D., Stanzioneb, D., Turnerd, G., Cockerill, T.M., Merchant, N., Taylor, J., Vaughn, M., Foster, I., Skidmore, E., Tuecke, S., Gaffney, N.I., 2015. Jetstream: a self-provisioned, scalable science and engineering cloud environment. ACM Int. Conf. Proceeding Ser. 2015-July. https://doi.org/10.1145/2792745.2792774 https://doi.org/.
Stoddard, R.A., Gee, J.E., Wilkins, P.P., McCaustland, K., Hoffmaster, A.R., 2009. Detection of pathogenic Leptospira spp. through TaqMan polymerase chain reaction targeting the LipL32 gene. Diagn. Microbiol. Infect. Dis. 64, 247–255. https://doi.org/10.1016/j.diagmicrobio.2009.03.014 https://doi.org/.
Sutter, J.D., Sosa Pascual, O., 2018. Deaths from bacterial disease in Puerto Rico spiked after Maria.
Torres, P.J., Edwards, R.A., McNair, K.A., 2017. PARTIE: a partition engine to separate metagenomic and amplicon projects in the Sequence Read Archive. Bioinformatics 33, 2389–2391. https://doi.org/10.1093/bioinformatics/btx184 https://doi.org/.
Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G.D., Roskies, R., Scott, J.R., Wilkins-Diehr, N., 2014. XSEDE: accelerating scientific discovery. Comput. Sci. Eng. 120, 62–74.
Truitt, Z.G., Poon-Kwong, B., Bachoon, D.S., Otero, E., 2020. Seasonal shifts in the presence of pathogenic leptospires, Escherichia coli, and physicochemical properties in coastal rivers and streams of Puerto Rico. J. Environ. Qual. 49, 1264–1272. https://doi.org/10.1002/jeq2.20091 https://doi.org/.
Von Meijenfeldt, F.A.B., Arkhipova, K., Cambuy, D.D., Coutinho, F.H., Dutilh, B.E., 2019. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. Genome Biol. 20, 1–14. https://doi.org/10.1186/s13059-019-1817-x https://doi.org/.
Vosloo, S., Sevillano-Rivera, M., Pinto, A.J., 2019. Modified DNeasy PowerWater Kit® protocol for DNA extractions from drinking water samples. https://doi.org/10.17504/protocols.io.66khhcw.
Walker, J.T., 2018. The influence of climate change on waterborne disease and Legionella: a review. Perspect. Public Health 138, 282–286. https://doi.org/10.1177/1757913918791198 https://doi.org/.
Wang, H., Edwards, M., Falkinham, J.O., Pruden, A., 2012. Molecular survey of the occurrence of legionella spp., mycobacterium spp., pseudomonas aeruginosa, and amoeba hosts in two chloraminated drinking water distribution systems. Appl. Environ. Microbiol. 78, 6285–6294. https://doi.org/10.1128/AEM.01492-12 https://doi.org/.

Warren, M.R., 2019. Challenges With Sampling and Understanding Lead Corrosion Issues in Puerto Rico After Hurricane Maria. University of Colorado.
Wickham, H., 2011. Ggplot2. Wiley Interdiscip. Rev. Comput. Stat. 3, 180–185. https://doi.org/10.1002/wics.147 https://doi.org/.
Wickham, H., Girlich, M., Ruiz, E., 2021. dbplyr: A "dplyr" Back End for Databases.
Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20, 1–13. https://doi.org/10.1186/s13059-019-1891-0 https://doi.org/.
Wynwood, S.J., Graham, G.C., Weier, S.L., Collet, T.A., McKay, D.B., Craig, S.B., 2014. Leptospirosis from water sources. Pathog. Glob. Health 108, 334–338. https://doi.org/10.1179/2047773214Y.0000000156 https://doi.org/.
Yáñez, M.A., Carrasco-Serrano, C., Barberá, V.M., Catalán, V., 2005. Quantitative detection of Legionella pneumophila in water samples by immunomagnetic purification and real-time PCR amplification of the dotA gene. Appl. Environ. Microbiol. 71, 3433–3441. https://doi.org/10.1128/AEM.71.7.3433-3441.2005 https://doi.org/.