1 **Spatial-temporal targeted and non-targeted surveys to assess microbiological**

2 **composition of drinking water in Puerto Rico following Hurricane Maria.**

3

4 Maria Sevillano[a], Solize Vosloo[a], Irmarie Cotto[a], Zihan Dai[b, c], Tao Jiang[a], Jose M. Santiago

5 Santana[d], Ingrid Y. Padilla[e], Zaira Rosario-Pabon[f], Carmen Velez Vega[f], José F. Cordero[g],

6 Akram Alshawabkeh[a], April Gu[h], and Ameet J. Pinto[a]

7 [a] Department of Civil and Environmental Engineering, Northeastern University, Boston, MA,

8 USA

9 [b] Key Laboratory of Drinking Water Science and Technology, Research Center for Eco-

10 Environmental Sciences, Chinese Academy of Sciences, Beijing, China

11 [c] University of Chinese Academy of Sciences, Beijing, China

12 [d] Department of Natural Sciences, University of Puerto Rico, Carolina, PR, USA

13 [e] Department of Civil Engineering and Surveying, University of Puerto Rico, Mayagüez, PR,

14 USA

15 [f] University of Puerto Rico—Medical Sciences Campus, San Juan, PR, USA

16 [g] Department of Epidemiology and Biostatistics, University of Georgia, Athens, Georgia, USA

17 [h] School of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA

18

19

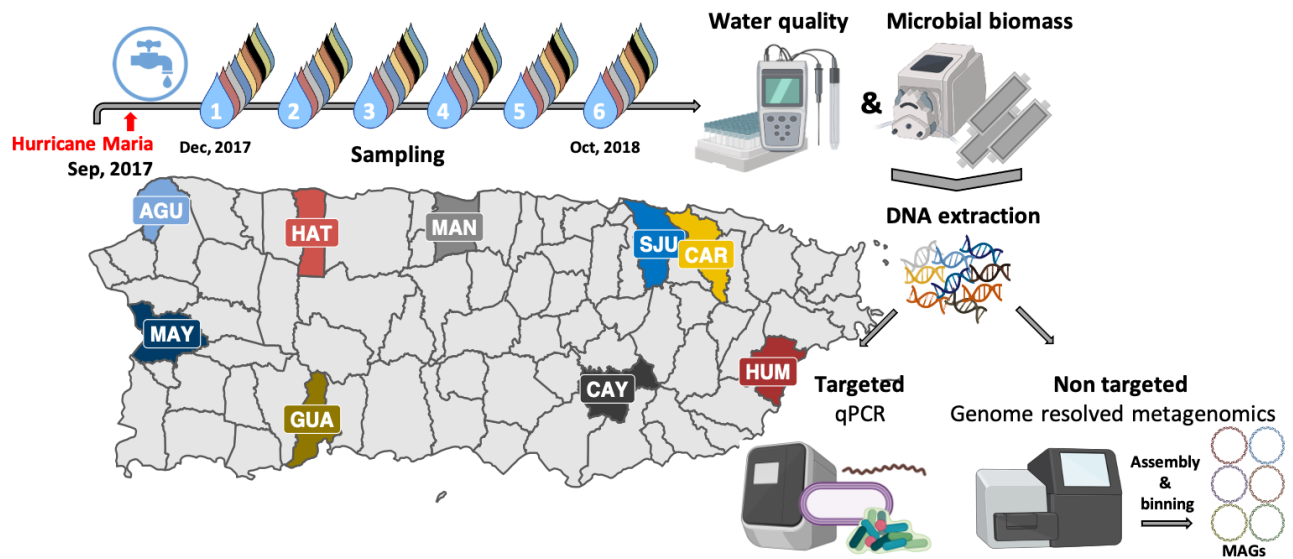20 *Corresponding author: a.pinto@northeastern.edu

21

22

23 **Keywords:** drinking water quality, hurricane Maria, metagenomics, qPCR, genome resolved

24 metagenomics

25

26 **Graphical Abstract**

27

2

28 **Abstract**

29 Loss of basic utilities, such as drinking water and electricity distribution, were sustained for

30 months in the aftermath of Hurricane Maria's (HM) landfall in Puerto Rico (PR) in September

31 2017. The goal of this study was to assess if there was deterioration in biological quality of

32 drinking water due to these disruptions. This study characterized the microbial composition of

33 drinking water following HM across nine drinking water systems (DWSs) in PR and utilized an

34 extended temporal sampling campaign to determine if changes in the drinking water microbiome

35 were indicative of HM associated disturbance followed by recovery. In addition to monitoring

36 water chemistry, the samples were subjected to culture independent targeted and non-targeted

37 microbial analysis including quantitative PCR (qPCR) and genome-resolved metagenomics. The

38 qPCR results showed that residual disinfectant was the major driver of bacterial concentrations

39 in tap water with marked decrease in concentrations from early to late sampling timepoints.

40 While *Mycobacterium avium* and *Pseudomonas aeruginosa* were not detected in any sampling

41 locations and timepoints, genetic material from *Leptospira* and *Legionella pneumophila* were

42 transiently detected in a few sampling locations. The majority of metagenome assembled

43 genomes (MAGs) recovered from these samples were not associated with pathogens and were

44 consistent with bacterial community members routinely detected in DWSs. Further, whole

45 metagenome-level comparisons between drinking water samples collected in this study with

46 samples from other full-scale DWS indicated no significant deviation from expected community

47 membership of the drinking water microbiome. Overall, our results suggest that disruptions due

48 to HM did not result in significant and sustained deterioration of biological quality of drinking

49 water at our study sites.

50

51

3

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 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 faecal 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 faecal 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

4

83   remote areas either recovered quickly or months later. However, electricity services took longer

84   to recover irrespective of remoteness category. Even when water services were restored,

85   intermittent water supply was common due to unreliable electrical supply; this could potentially

86   degrade water quality via stagnation and loss of disinfectant residual, intrusion, and

87   backflows(Bautista-de los Santos et al., 2019). Boil advisories and point of use chlorination were

88   in place after water services resumed and were reported by the Puerto Rico Aqueducts and

89   Sewers Authority (PRASA) through mid-January 2018(Exum et al., 2018). Previously, Lin et

90   al.(Lin et al., 2020) provided insights into metals, micropollutants, and molecular toxicity of pre-

91   and post-HM drinking water samples in PR and showed the impact of HM on chemical water

92   quality, suggesting that trace metals were potential drivers of cumulative risk from drinking

93   water. Additionally, Keenum et al.(Keenum et al., 2021) characterized five unregulated small

94   scale DWS and one large PRASA DWS in PR six months after HM. In this study, targeted

95   culture and molecular based analyses (i.e., quantitative PCR (qPCR), 16S rRNA amplicon

96   sequencing) demonstrated similar microbial communities and concentrations of opportunistic

97   premises plumbing pathogens (OPPPs) compared to those reported in the continental US. In our

98   study, we also aim to evaluate the microbial water quality in the aftermath of HM. However,

99   unlike Keenum et al(Keenum et al., 2021), we conducted a recurrent sampling campaign

100   beginning in December 2017 spanning nine locations across PR for a duration of a year. Despite

101   the magnitude of HM in PR, there hasn't been a large effort to characterize microbial water

102   quality. To date, there have been two reports focused on chemical contamination(Lin et al., 2020;

103   Warren, 2019) and two (including this one) on microbial composition(Keenum et al., 2021) of

104   DWDs on the island. These studies are essential to establish relationships, sampling

105   infrastructure, and methodologies needed to respond to future storms, as well as to communicate

106   risk and execute corrective actions to decrease exposure risk and unwanted health outcomes.

107   Thus, our goals were (1) to utilize an extended spatial-temporal sampling campaign to determine

108   if changes in drinking water microbiome were indicative of disturbance followed by recovery,

109   (2) if this disturbance-recovery dynamic was associated with presence of potential pathogens, (3)

110   whether potential pathogen presence was persistent or transient, and finally (4) whether

111   microbial composition of PR drinking water was consistent with or deviated significantly from

112   other drinking water systems.

113

5

2. Materials and methods

2.1 Drinking water sampling and water quality analyses. Nine sampling locations were chosen across different geographic locations in PR. Tap water 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 20L volume for each filter. 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, qPCR and shotgun sequencing. 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 seconds 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 with 310 µg lyophilized carrier RNA (QIAGEN, Cat. no.

6

145   1068337)) was mixed with 600 μL of recovered supernatant. This was then purified using the

146   automated DNA purification protocol with DNeasy PowerWater Kit on a QIACube system

147   (QIAGEN, Cat. no. 9001292). In addition to the samples, controls were processed identically and

148   consisted of unused transported Sterivex filter membranes (filter blank), no input material

149   (reagent blank), and sterilized deionized water filtered through Sterivex (water blank). This set of

150   three controls were included with each sampling campaign (n=6) and extraction run.

151

152   qPCR was performed on a QuantStudio™ 3 Real-Time PCR System (ThermoFisher Scientific

153   Cat. no. A28567). PCR reactions were carried out in a 20 μl volume containing Luna Universal

154   qPCR Master Mix (New England Biolabs, Inc., Cat. no. NC1276266), F515

155   (GTGCCAGCMGCCGCGGTAA) and R806 (GGACTACHVGGGTWTCTAAT) primer

156   pair(Caporaso et al., 2011) (IDT), DNAse/RNAse-Free water (Fisher Scientific, Cat. no.

157   10977015), and 5 μL of 10X diluted DNA template. Reactions were prepared by an epMotion

158   M5073 liquid handling system (Eppendorf, Cat. no. 5073000205D) in triplicate. The cycling

159   conditions were as follows, initial denaturing at 95 °C for 1 min followed by 40 cycles of

160   denaturing at 95 °C for 15 s, annealing at 50 °C for 15 s, and extension 72 °C for 1 min. Melting

161   curve analyses was performed by ramping from 72 °C to 95 °C for 15 s, and 60 °C for 1 min, 95

162   °C for 15 s. A negative control (NTC) and a standard curve consisting of 7 points ranging from

163   10^1-10^7 copies of 16S rRNA gene were included in every qPCR run. Genomic DNA from

164   selected samples were sent to University of Illinois Roy J. Carver Biotechnology Center (UI-

165   RJCBC) for library preparation using a low input DNA kit (NuGEN, Cat. no 0344NB). Libraries

166   were loaded into two SP lanes on a NovaSeq 6000 instrument with an output of 2x150nt reads.

167   The prepared libraries included 33 samples and 3 pooled blanks (i.e., filter blank, reagent blank,

168   and water blank) from all locations.

169

170   **2.3 qPCR analyses for waterborne pathogens.** Previously published primers targeting

171   *Legionella* spp.(Nazarian et al., 2008; Yáñez et al., 2005), *Mycobacteria* spp.(Chern et al., 2015;

172   Radomski et al., 2010), pathogenic Leptospira(Stoddard et al., 2009), and *Pseudomonas*

173   *aerugionosa*(Anuj et al., 2009) were used for qPCR assays. Reactions were set up by an

174   epMotion M5073 liquid handling system in triplicate. The assays consisted of 2X PrimeTime

175   Gene Expression Master Mix (IDTDNA, Cat no. 290479057) with low reference ROX dye,

7

176   target primers and probe (IDTDNA), 10-fold diluted template and water (UltraPure™
177   DNase/RNase-Free Distilled Water, Thermo Fisher Scientific, Cat. no. 10977015). Single target
178   reactions were conducted in a total volume of 20 µL, whereas duplex qPCR (i.e., *Pseudomonas*
179   *aeruginosa* assay) were conducted in 25 µL. Primer and probe sequences and cycling conditions
180   are described in Table S2. Standard cycling conditions for all reactions were programmed on a
181   QuantStudio™ 3 Real-Time PCR System. Target gene copy numbers were determined by
182   comparing threshold cycle with standard curve generated using gblocks gene fragments as
183   standards (Table S2).

184

185   **2.4 Metagenomic reads pre-processing.** The raw reads obtained from UI-RJCBC were
186   processed with fastp(Chen et al., 2018) v0.19.7 to remove homopolymer stretches using the
187   following flags '--trim_poly_g --trim_poly_x'. Further, trimmed reads were mapped with BWA-
188   MEM(Li, 2013) v0.7.12 against the UniVec database build 10.0 (National Center for
189   Biotechnology Information 2016) to perform vector screening and retained unmapped paired
190   reads. Subsequently, Nonpareil(Rodriguez-R et al., 2018) v3.303 was used on the quality filtered
191   reads to estimate average community coverage and metagenomic dataset diversity using kmer
192   algorithm with a kmer size of 20 and default parameters.

193

194   **2.5 Metagenome assembly and mapping.** Sample reads were co-assembled based on each
195   sampling location using metaSPAdes(Nurk et al., 2017) v 3.11.1 with the following flags '-meta
196   -t 16 --phred-offset 33 -m 500 -k 21,33,55,77,99,119' and further filtered to a minimum scaffold
197   length of 500 bp. Reads from samples and controls (i.e., extraction blank, filter blank, DI water
198   blank) were mapped to co-assemblies using BWA-MEM. An approach similar to Dai et al.(Dai
199   et al., 2020) was used to remove potential contaminant scaffolds. Briefly, BWA-MEM with flag
200   -F4 and -f2 was used to map sample and control reads against co-assemblies. The
201   BEDtools(Quinlan and Hall, 2010) genomecov using flags -g, and -d was used to calculate
202   coverage and per base coverage using generated BAM files. Relative abundances (RA) and
203   normalized coverage deviation (NCD) were calculated for each scaffold with coverage and per
204   base coverage information, respectively. Scaffolds that were not detected in the blanks or for
205   which sample RA was greater than the blank RA and the sample NCD was less than the blank
206   NCD were considered true scaffolds and were used in downstream analyses. All scaffolds that

8

207 did not meet these criteria were considered contaminant scaffolds and removed from further

208 analyses. Assembly statistics were obtained from QUAST(Gurevich et al., 2013) 5.0.2. To

209 contextualize the metagenome assemblies with respect to other distribution systems, assemblies

210 from other drinking water systems (considered unperturbed systems because samples were not

211 associated with any natural disaster or water quality issues) were compared against our co-

212 assemblies. MASH(Ondov et al., 2016) v2.1.1 was used to estimate the dissimilarity between

213 assemblies using '-r' and '-m 2' flags, and a sketch size of 100000.

214

215 **2.6 Taxonomic classification of metagenomic assemblies.** The taxonomic classification of

216 scaffolds was performed with a contig annotation tool(Von Meijenfeldt et al., 2019) (CAT

217 v5.0.4) program in which open reading frames (ORF) are predicted with Prodigal(Hyatt et al.,

218 2010) and used as alignment queries by DIAMOND(Buchfink et al., 2014) against the NCBI

219 non-redundant (nr) protein database (downloaded ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/,

220 2020-03-04). Selected genera known to contain pathogenic species, as well as non-pathogenic

221 species, that are relevant to drinking water systems were further examined. The CAT annotations

222 of true scaffolds were evaluated against annotated controls and only scaffolds with rpoB

223 normalized coverage above controls and additional annotation support were considered to avoid

224 false positives at the genus level. Specifically, additional support consisted of classification with

225 kaiju(Menzel et al., 2016) v1.7.2 using reference indexes containing NCBI BLAST nr database

226 and microbial eukaryotes with default parameters (downloaded http://kaiju.binf.ku.dk/server,

227 nr_euk 2019-06-25) and/or by kraken2(Wood et al., 2019) v2.0.9-beta against RefSeq database

228 (downloaded https://lomanlab.github.io/mockcommunity/mc_databases.html). Additionally, we

229 examined the scaffolds of eukaryotic origin with metaEuk(Levy Karin et al., 2020) v3.8dc7e0b

230 to assign taxonomy using a publicly available MMseqs2 database containing protein profiles

231 from the marine eukaryotic reference catalogue (MERC), Marine Microbial Eukaryote

232 Transcriptome Sequencing Project (MMETSP), and Uniclust50 (downloaded

233 http://wwwuser.gwdg.de/~compbiol/metaeuk/2020_TAX_DB/).

234

235 **2.7 Metagenomic assembled genomes.** Co-assemblies were binned with CONCOCT(Alneberg

236 et al., 2014) within anvi'o(Eren et al., 2015) v5.1 by clustering scaffolds 2500 bp or longer into

237 metagenome assembled genomes (MAGs) and manually refining them within the anvi'o

9

238   platform. Furthermore, dRep(Olm et al., 2017) v2.3.2 was used to dereplicate MAGs and obtain

239   representative genomes with flags '-comp 50 -con 10' and default values. GTDB-tk(Parks et al.,

240   2018) v0.1.3 was used to assign taxonomy to MAGs with the flag 'classify_wf'. Sample reads

241   were mapped to corresponding single pseudo contig MAGs. Pseudo contigs were generated with

242   the union command in EMBOSS(Rice et al., 2000) utility. Mapping was performed with

243   BBMap(Bushnell, 2015) v38.24 using a 90% identity threshold and setting flags

244   'ambiguous=best', 'mappedonly=t', and 'pairedonly=t'. Detection of a MAG in a sample was

245   established when ≥ 25% of its bases were covered by at least one read from the corresponding

246   sample. Coverage was determined with samtools(Li et al., 2009) v1.10 'coverage' . The

247   abundance of a MAG in a sample was calculated as sample reads mapped per million reads per

248   genome length in kbp (RPKM). Further information about MAGs, such as number of 5S rRNA,

249   16S rRNA, 23S rRNA, and tRNA counts was obtained by annotating the MAGs using

250   DRAM(Shaffer et al., 2020) v1.0.6. The databases used with DRAM were downloaded with the

251   following flags 'DRAM-setup.py prepare_databases --output_dir DRAM_data --skip_uniref'.

252   MAGs from this study were compared to 52,515 MAGs recovered from environmentally diverse

253   metagenomes by Nayfach et al.(Nayfach et al., 2020) (downloaded

254   https://portal.nersc.gov/GEM/genomes/fna.tar, 2020-11-10) using FastANI(Jain et al., 2018)

255   v2.3.2 with default parameters. Metadata linked with this genomic catalogue of earth

256   microbiomes, hereafter referred to as JGI MAGs (downloaded

257   https://portal.nersc.gov/GEM/genomes/genome_metadata.tsv, 2020-11-30) was used to address

258   niche association. Further, SEARCH-SRA(Stewart et al., 2015; Torres et al., 2017; Towns et al.,

259   2014) online portal was used to interrogate the SRA database (246,329 records) by aligning

260   metagenomic datasets to our MAGs. Only records that mapped 10 or more reads from the SRA

261   collection were further inspected. The metadata associated with SRA accession numbers

262   (downloaded https://s3.amazonaws.com/starbuck1/sradb/SRAmetadb.sqlite.gz, 2021-04-08) was

263   incorporated through custom scripts in R software(R Development Core Team, 2016) that rely

264   on the dbplyr(Wickham et al., 2021) package. Records that were classified as retrieved from

265   metagenomic library sources and with a whole genome sequence strategy were retained for

266   analysis. Considering that SRA metadata is user provided, manual curation to ensure consistency

267   and retrieve the same ecosystem categories as in JGI MAGs metadata was performed. Data with

268   missing context (lacking information in title or description) were considered as "others" and

269   removed from analyses. The association between a MAG and an ecosystem category was
270   determined by multiplying the total number of reads from the ecosystem category mapping to the
271   MAG by the ratio of the number of unique SRA records associated with a MAG and ecosystem
272   category and the number of unique SRA records within the entire dataset assigned to the
273   ecosystem category.

274

275   **2.8 Data analyses and statistics.** Statistical analyses were conducted in R and visualizations
276   generated with ggplot2(Wickham, 2011) package. PCA analyses of water quality parameters
277   were performed with centered and scaled data in R base prcomp(). Linear regression models of
278   log10 (volume normalized 16S rRNA gene copies) against water quality parameters were fit
279   using base R lm(). Pearson correlation between log10(volume normalized 16S rRNA gene
280   copies) and chlorine concentrations (mg/L) was obtained with base R cor() and exponential
281   decay curve with nls(). Euclidian distances between pairwise Mash distance of DW
282   metagenomes was calculated with vegan(Oksanen et al., 2015) function vegdist() and clustered
283   with complete linkage method with base R hclust(). PCoA ordination of Mash distances was
284   performed with ape(Paradis et al., 2004) function pcoa(). Group wise non parametric testing was
285   performed with R base statistic packages using function kruskal.test() or wilcox.test().
286   Permutational hypothesis testing (n iterations=10000) of differences in group means between
287   two groups was performed after up-sampling group data to balance observations using
288   upsample() from the groupdata2(Olsen, 2021) package. Analyses of variance (ANOVA) was
289   performed with aov() and followed up with post hoc Tukey-Kramer testing using TukeyHSD() in
290   base R.

291

292   **3. Results and discussion**

293   **3.1 Bacterial concentrations were associated with water quality parameters, particularly**
294   **total chlorine concentrations.** Water quality parameters were recorded for each sampling
295   location and timepoint (Figure1A, Table S1). PCA analyses was conducted to assess whether
296   water chemistry varied spatially and/or temporally. Nitrogen species were excluded from PCA
297   analyses as their concentrations were below detection limit at several locations/timepoints and
298   nitrate concentrations also strongly correlated with conductivity (Pearson's R = 0.88, p<0.001)
299   and phosphate concentrations (Pearson's R = 0.47, p<0.001). No clear clustering of samples by

11

300 location or timepoint was observed despite drinking water samples being obtained from variable

301 source waters (Figure 1B). For instance, CAY, GUA, and MAN had higher mean and variable

302 conductivities compared to other locations, which is consistent with source water type of these

303 locations being groundwater. Phosphate, on the other hand, was relatively narrowly distributed

304 across samples, and in lower concentration compared to other DW systems(Gooddy et al., 2015),

305 with CAY location consistently higher than other PR locations.

306
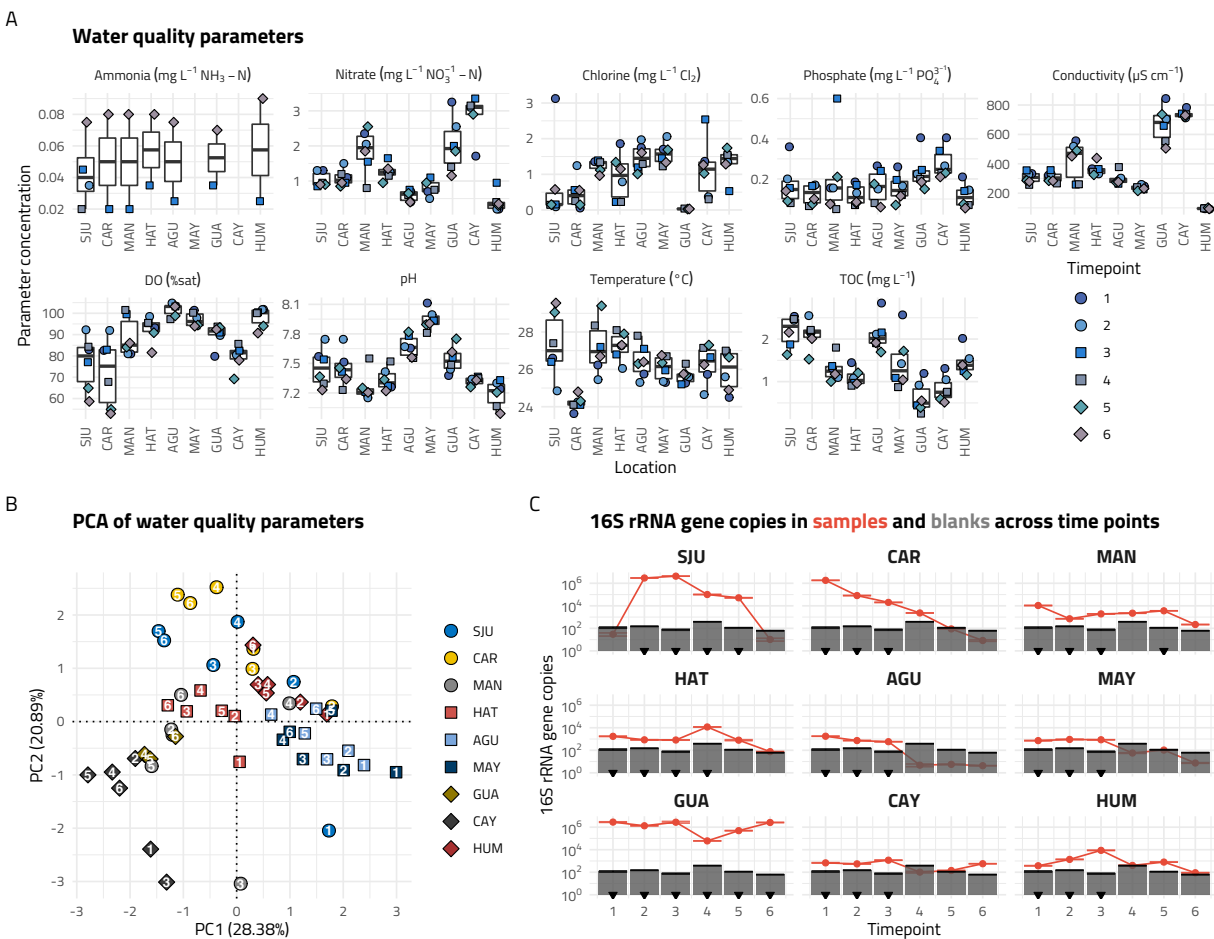
308 **Figure 1:** (A) Concentrations of water quality parameters for nine locations from December 2017 to October 2018. Shapes and
309 colors correspond to timepoints. All nitrite measurements were below detection limit (BDL, not shown), for ammonia 36
310 measurements were BDL, and for chlorine 3 samples were BDL. (B) Principal component analyses of water quality parameters.
311 Colors and shapes correspond to sampling location. Direct labels of timepoints within sampling location, in white. The PCA does
312 not show a clear clustering between samples or timepoints. (C) Gene copies of 16S rRNA gene for samples and blanks. Red
313 lines correspond to samples and grey bars correspond to blanks. Comparison between samples and blanks guided sample
314 selection for sequencing. Inverted black triangles denote samples that were sequenced.

315

12

316   We quantified the abundance of 16S rRNA genes in all samples as a measure of bacterial
317   concentrations (Figure 1C). The average PCR efficiencies for these qPCR assays was 90±2.9%.
318   Volume normalized 16S rRNA gene copies (16S rRNA gene copies mL^{-1}) ranged from 6.4×10^{-2}-
319   9.1×10^4 copies mL^{-1}. Independently regressing variables in the PCA as descriptors of log10(16S
320   rRNA gene copies mL^{-1}) for each location resulted in six significant (p<0.05) linear models. The
321   goodness of fit for all models was relatively high with an average adjusted R^2 of 0.704 ± 0.062,
322   and significant associations between log10(16S rRNA mL^{-1}) and DO in CAR, pH in HUM,
323   temperature in CAR, MAY and GUA, and TOC in AGU. However, regressing parameters
324   against log10(16S rRNA mL^{-1}) for all locations only resulted in statistically significant
325   associations with temperature (adj R^2 = 0.067, p < 0.05) and total chlorine (adj R^2 = 0.227, p <
326   0.001). A multiple linear regression model with all water quality parameters as descriptors of
327   log10(16S rRNA mL^{-1}) (n=51), indicated that total chlorine was the major driver associated with
328   decreasing 16S rRNA gene concentrations (Adj R^2 = 0.28, p<0.001, Figure 2A, Table S3).
329   Chlorine concentrations measured in samples were comparable to those reported in other US
330   studies(Stanish et al., 2016), except for GUA where it was either below detection limit (BDL) or
331   very low for all timepoints. All samples, except timepoint 1 in SJU, had total chlorine
332   concentrations below 3 mg L^{-1} (Figure 2B, Table S1). The negative relationship between
333   bacterial load and chlorine concentration is particularly evident for SJU and CAR, where
334   decreasing total chlorine concentration were associated with increased bacterial concentrations,
335   relatively stable chlorine concentrations correspond to stable bacterial concentrations, and
336   absence of chlorine shows high bacterial concentration, respectively (Figure 2B). Not
337   surprisingly, these results suggest that maintaining chlorine residual is critically important for
338   ensuring low bacterial concentrations which could be particularly challenging due to
339   infrastructure damage from natural disasters. Some locations did exhibit significant variation in
340   chlorine concentrations between December 2017 and February 2018 (e.g., SJU, CAR, HAT),
341   with some of these variations associated with water main breaks (e.g., CAR in December 2017).
342

13
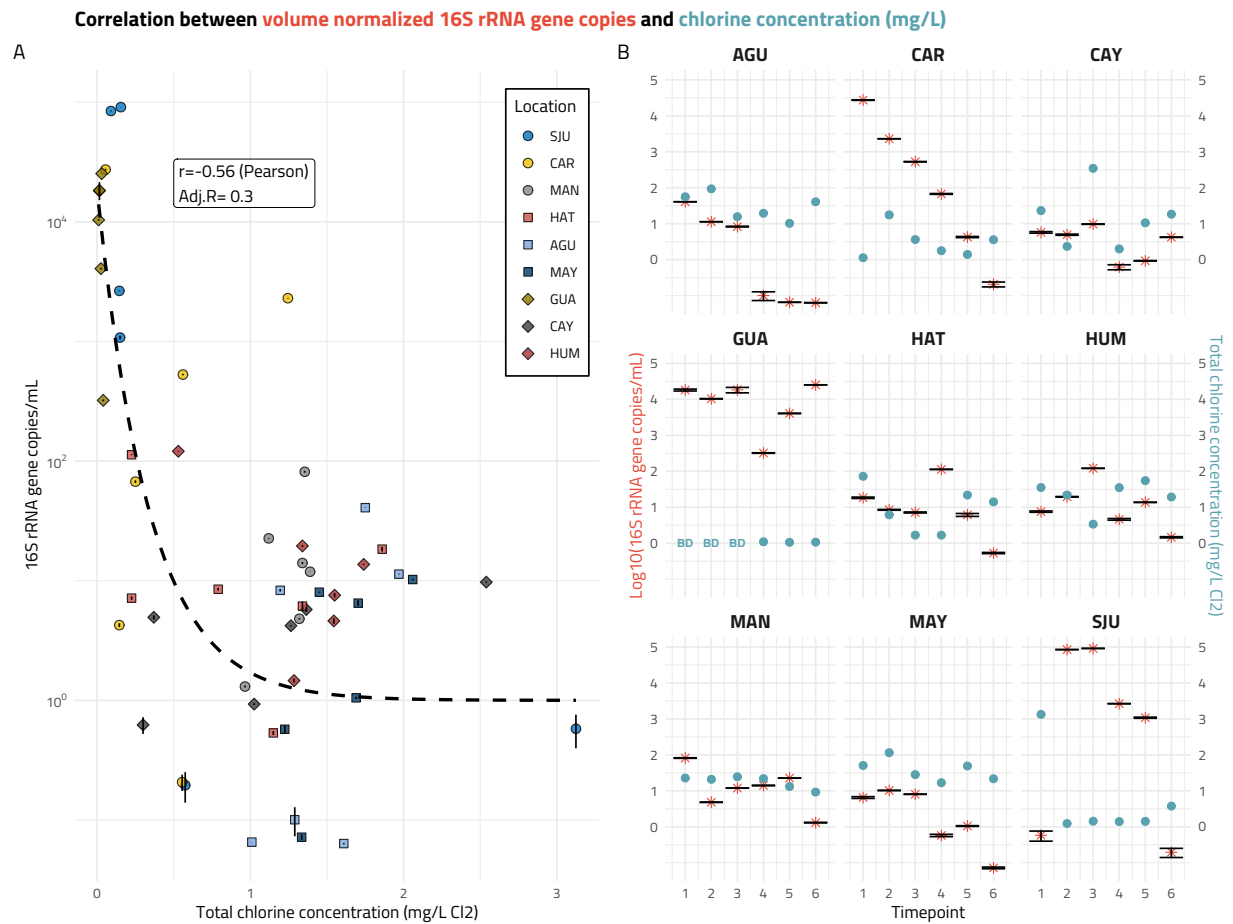
Figure 2: (A) 16S rRNA gene copies normalized by volume correlated with total chlorine concentration. Colors and shapes correspond to sampling location. Total chlorine concentration was the major driver among all recorded water quality parameters when all samples were aggregated in a multiple linear 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.

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, Figure 1C, 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 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

14

361   were generated after quality filtering and 1.16 Gb reads were not mapped against UniVec,

362   resulting in less than 2.6% of the reads being discarded for the majority of samples (n=31), only

363   2 samples (i.e. HUM_1, HUM_2) retained less than 92.5% of the raw reads (Table S4). Within

364   sample diversity (Nd), as assessed by Nonpareil curves, ranged from 15.33-19.16 (Figure 3A,

365   Table S5). These indices rely on redundancy of reads in a metagenome to estimate diversity of

366   metagenome, with higher Nd corresponding to more diverse communities. The Nd values

367   observed here are consistent with those seen in other chlorinated drinking water systems(Dai et

368   al., 2020). The spread of observed Nd values within location was larger for GUA, HUM, and

369   HAT, indicating higher temporal variation in diversity, while low variability in Nd values at

370   CAR indicative of low temporal differences. Kruskal-Wallis test of Nd by location reveal

371   significant differences in the median of at least one of the groups ($p<0.05$), however multiple

372   hypothesis correction with Dunn test did not identify any significant pairwise differences.

373   Significant and positive correlations were observed between Nd and pH at AGU ($p<0.05$), and

374   Nd and DO at HUM ($p<0.01$), and significant negative correlations between HAT diversity and

375     $\log10(16S \ rRNA \ mL^{-1})$ and CAY diversity and nitrate ($p<0.05$ for both locations).

376

377   The reads were subsequently assembled and scaffolds identified as potential contamination were

378   removed as outlined in the materials and methods section. A summary of statistics for the 9 co-

379   assemblies that were generated can be found in Table S6. CAT was used to annotate true

380   scaffolds (i.e., scaffolds retained post contamination analysis) and coverage information allowed

381   us to obtain per sample profiles (Figure 3B, Table S7). Bacteria generally constituted the largest

382   portion in the samples from SJU, CAR, GUA, and CAY, with a mean relative abundance (RA)

383   of 96.6±1.72%. In contrast, scaffolds of eukaryotic origin (18.5±19.5%) and unclassified

384   scaffolds (5.62±2.85%) constituted a significant proportion of the community in MAN, HAT,

385   AGU, MAY, and HUM. The RA of eukaryotic contigs is not consistent within locations,

386   suggesting high temporal variation of the eukaryotic fraction in the systems. Despite highest

387   eukaryotic contigs RA at timepoint 1 for AGU and MAY and timepoint 2 at HUM, water quality

388   parameters at these sites had relatively low temporal variation. On the other hand, MAN had its

389   highest RA at timepoint three when DO and phosphate concentrations were higher than mean

390   values. In contrast with other sites with high contribution of eukaryotic contigs, MAN water

391   source is groundwater. Phosphate concentration is associated with source water type and

392   treatment processes (i.e., corrosion control). Douterelo et al.(Douterelo et al., 2018) compared

393   metagenomic samples from sites with different source waters and saw dominance of bacteria and

394   no significant differences in the RA of eukaryotes. Moreover Inkinen et al.(Inkinen et al., 2019)

395   correlated the presence of phosphate concentrations with active eukaryotes in a chlorinated

396   groundwater DWDS in contrast with surface waters. At the HAT site, the largest eukaryotic RA

397   was seen at timepoint 1 consistent with highest chlorine concentration. Disinfection is a driver of

398   microbial composition in DWS and it likely reduced the contribution of bacterial community in

399   these samples, resulting in an observed increase in the RA of eukaryotes. We further investigate

400   the eukaryotic component of the microbial communities identified through CAT with MetaEuk

401   (Figure S1, Table S8 and S9). MetaEuk based classification indicated that free living amoeba

402   (FLA) capable of supporting intracellular growth of opportunistic pathogens (e.g., *Vermamoeba*,

403   *Acanthamoeba*, etc.) were transiently detected at low RAs at CAR, MAN, HAT, AGU, MAY,

404   CAY, and HUM (Figure S2, Table S9). Waterborne parasites like Giardia and Cryptosporidium

405   were not detected.

406

16
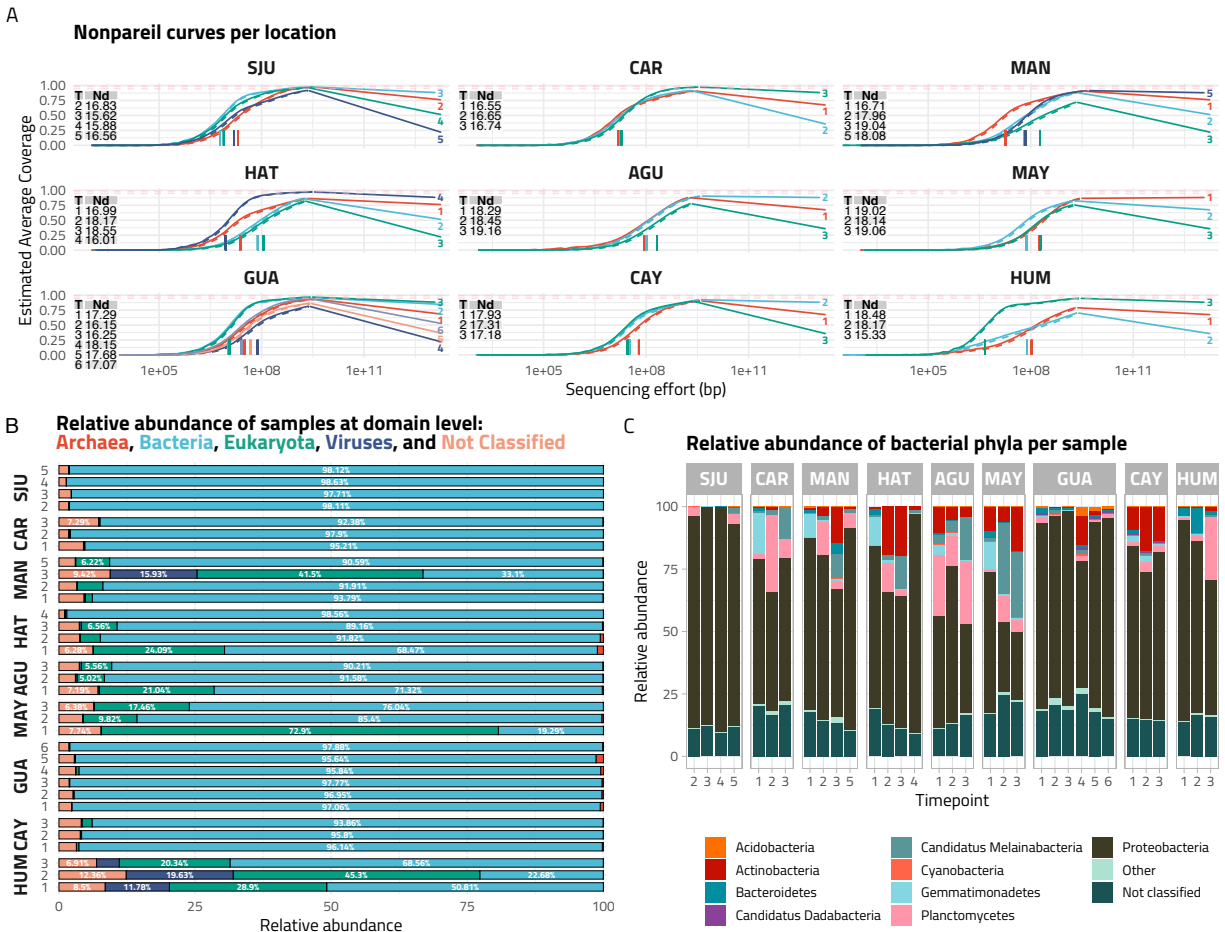
Figure 3: Diversity of microbial communities in metagenomes. (A) Nonpareil curves faceted by location. Each curve within facet denotes a sample directly labelled by timepoint. Respective Nonpareil diversity (Nd) is shown in table on the top left corner of each facet. The bars correspond to the sample 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.

The coverage of scaffolds classified as bacteria was normalized by rpoB gene coverage in the respective samples to assess the bacterial community (Figure 3C, Table S10). The proportion of bacterial scaffolds not classified beyond the domain level ranged from 9.20-24.80% and patterns were consistent within location. Similar to previous studies, Proteobacteria was the dominant phylum in the majority of samples, ranging from 27.31-90.07%, 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,

17

424   with an average RA of 9.10±6.95% in these samples, a mean 1.21±2.77% in other samples, and a

425   global 5.27±6.61% RA. Planctomycetes is present at a RA greater than 1% in 72.72% of

426   samples, but is predominant in CAR and AGU, with a mean RA of 13.41±15.08% and

427   20.58±7.24%, respectively, compared to a global average of 6.25±8.4%. On average,

428   81.2±9.61% of sample cumulative RA was not classified up to genus level. The dominant

429   classified genera were Bradyrhizobium, Gemmata, Gemmatimonas, Hyphomicrobium,

430   Methylobacterium, Mycobacterium, Novosphingobium, Pseudorhodoplanes, and Sphingomonas.

431   We further compared the metagenomic assemblies recovered from the nine PR samples to other

432   DWS (not impacted by natural disasters, i.e., undisturbed, Table S11), to assess if there were

433   indications of significant deviation that could be attributed to HM. There was no clear clustering

434   of metagenomes as shown by PCoA ordination of pairwise Mash distances including the nine co-

435   assemblies from PR and 52 co-assemblies from other DWS (Figure 4A). Furthermore, there was

436   no statistical difference between pairwise Mash distances grouped as PR vs other DWS and other

437   DWS vs other DWS using permutational t tests (p>0.3, Figure 4B). This suggests that the

438   differences in metagenomes between HM impacted and other DWSs are similar to those

439   observed between other DWSs. Additionally, complete linkage clustering indicated that SJU,

440   MAN, HAT, and AGU, and GUA and CAY clustered closely, both within and between each

441   other; CAR did not cluster directly with another location and MAY and HUM were similar, but

442   separate from the rest of the PR locations (Figure S3A). The respective Mash distances of early

443   and late samples clustered identically as the previous analyses when leveraging coverage data

444   and CAT classification of scaffolds to subset scaffolds pertinent to these categories (Figure S3B-

445   C). This indicates that the metagenomes from the samples collected in PR were largely

446   consistent with what would be expected from drinking water samples, irrespective of time of

447   collection (i.e., December 2017 or October 2018).

448

18
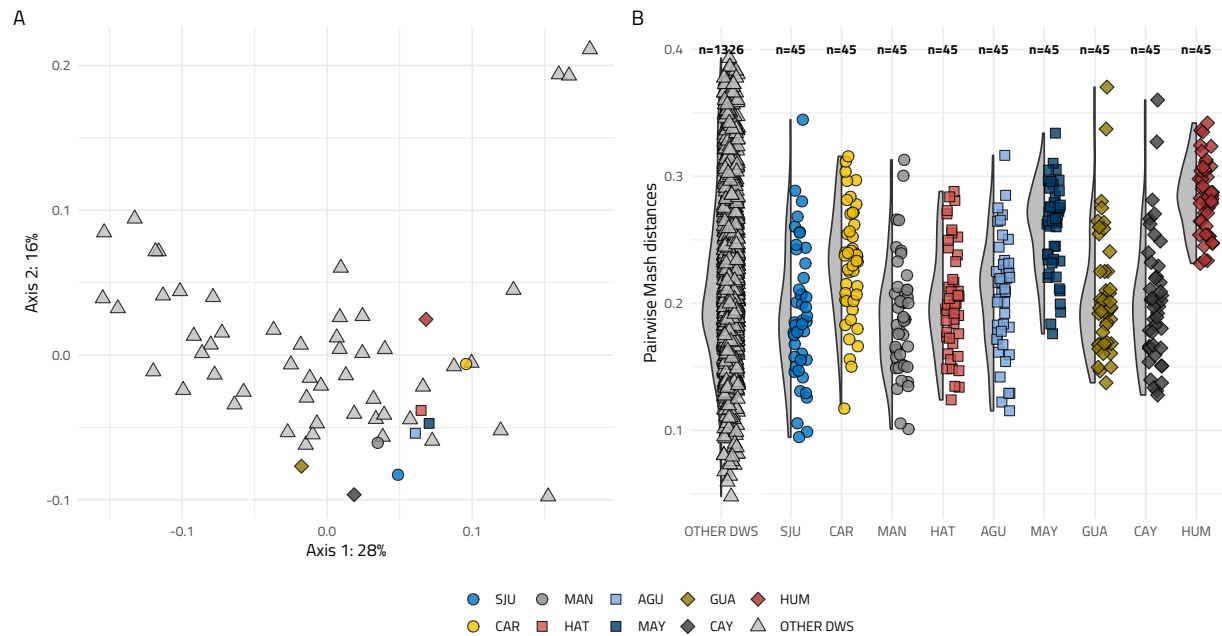
Figure 4: (A) PCoA ordination of Mash distances between DWS, including PR co-assemblies and reference DWSs. (B) Distribution of Mash distances prior to up sampling used for permutational t-test with group 1: other vs other and group 2: PR vs others. No significant differences were observed (p>0.3) between 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 (Figure 5A) 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(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

19

471   detected in MAN, HAT, and HUM. Pathogenic Leptospira was only detected in GUA. There

472   were statistically significant (Wilcoxon test, $p<0.05$) temporal differences in relative abundances

473   of OPPPs for Legionella in MAN, Mycobacterium in HAT, MAY, GUA, and HUM, and

474   Pseudomonas in HAT.

475

476   We used qPCR to quantify the abundance of *Legionella spp*. and *Mycobacterium spp*., while also

477   conducting more targeted assays to detect and quantify the abundance of *Legionella*

478   *pneumophila*, *Mycobacterium avium*, and *Pseudomonas aeruginosa*, and pathogenic species of

479   the Leptospira genus (Figure 5B). The average qPCR efficiencies for these assays was ~92.9%.

480   Negative controls for each sampling campaign were included in every assay. *Mycobacterium*

481   *avium* and *Pseudomonas aeruginosa* were not detected in any of the samples. The mean

482   concentration for *Legionella pneumophila*, *Legionella spp.*, pathogenic Leptospira, and

483   *Mycobacterium spp.* in the samples was 0.3, 7.8, 0.01, and 1 copies mL^{-1}, respectively.

484   *Mycobacterium spp.* were observed in all locations with a general frequency of detection of 74%.

485   *Mycobacterium spp.* were detected at every sampling timepoint in MAN, GUA, and CAY at very

486   low concentrations (0.81 ± 1.09 copies mL^{-1}), while their concentrations were as high as 10 copies

487   mL^{-1} at CAR and only detected in the first three timepoints. Consistent with metagenomic

488   results, *Legionella spp.* was widely observed across all sampling locations at an 81% frequency

489   of detection, with highest concentrations observed in SJU, CAR, and GUA. *Legionella spp.*

490   concentrations decreased from 50 and 127 copies mL^{-1} at SJU and CAR, respectively to below 1

491   copy mL^{-1} from December 2017 to October 2018. *Legionella spp.* thrive in warmer

492   temperatures(Lesnik et al., 2016), such as those in PR. Interestingly, concentrations *Legionella*

493   *spp.* and *Mycobacterium spp.* are several orders of magnitude lower that what has been published

494   in literature(Huang et al., 2021; Isaac and Sherchan, 2020; Ley et al., 2020; Liu et al., 2019) (i.e.,

495    $1\text{-}10^4$ copies mL^{-1}). A potential reason for this could be over-chlorination in the systems, which

496   had been reported in the aftermath of HM(Brown et al., 2018), including in early phase of

497   sampling as this study showed.

498

499   The dotA gene assay to target *Legionella pneumophila* revealed low concentrations (i.e.,

500   0.29 ± 0.44 copies mL^{-1} for SJU and GUA locations. In SJU, *L. pneumophila* was detected in

501   timepoints 4 and 5 only, while being detected at GUA at all timepoints. This is consistent with

502   observations from other DWS where *Legionella pneumophila* was detected at low frequency and

503   low concentrations(Lu et al., 2016; Wang et al., 2012).The LipL32 gene of pathogenic

504   Leptospira were observed only in GUA and at timepoints 1, 4, and 6 (i.e., 0.01 copies mL^{-1},

505   5.6% frequency). The detection of Leptospira at this location was consistent with the detection of

506   the genus Leptospira using metagenomics. Leptospira is not routinely reported in DWS, apart

507   from the recent study by Keenum et al.(Keenum et al., 2021), possibly due to the efficacy of

508   routine disinfection practices in the elimination of this pathogen(Wynwood et al., 2014).

509   However, its importance has been highlighted in rivers and creeks when used as drinking water

510   without proper treatment, particularly in situations of water scarcity, such as hurricanes(Keenum

511   et al., 2021; Truitt et al., 2020). The presence of Leptospira in GUA is possibly exacerbated by

512   the absence of residual chlorine at this location.

513
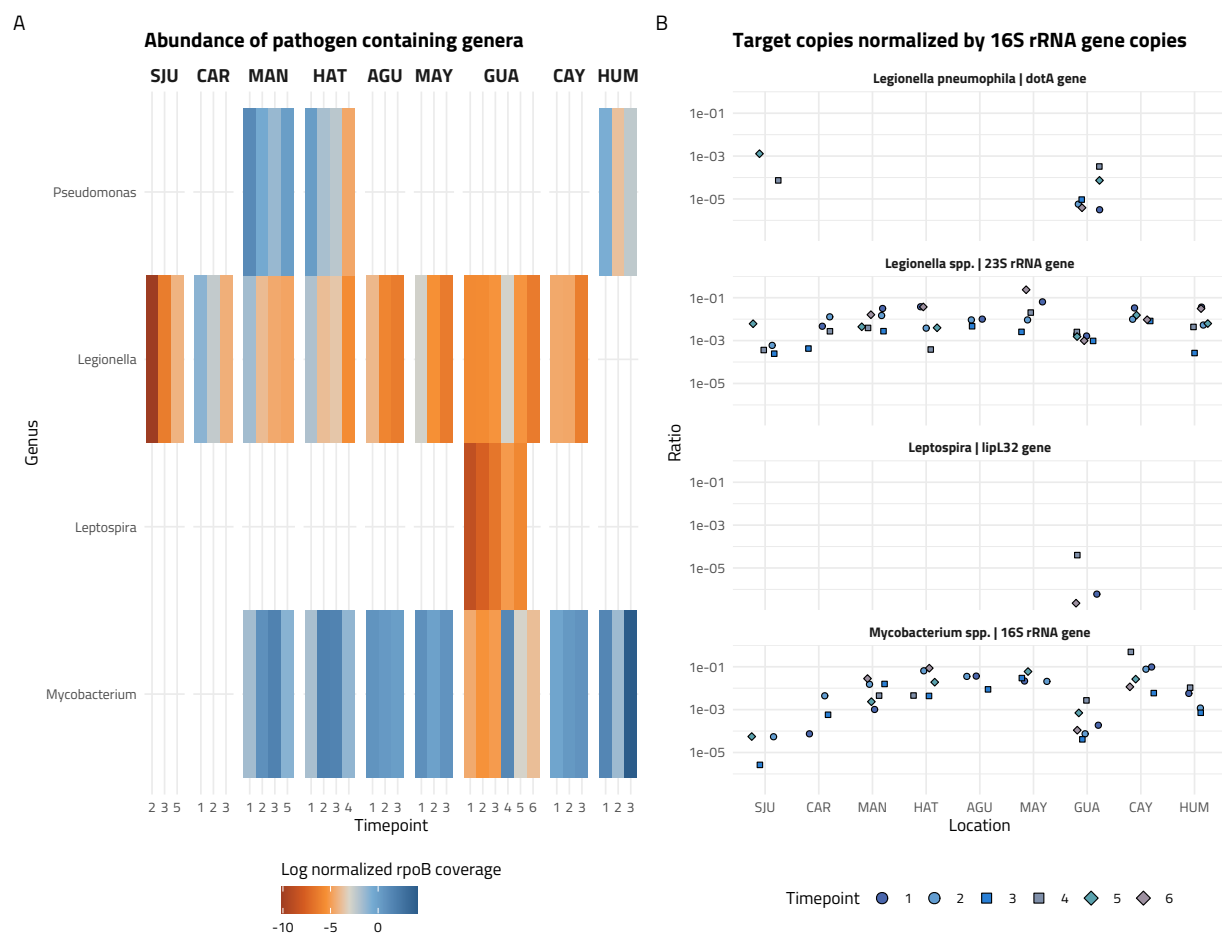
514

515   **Figure 5**: (A) rpoB normalized coverage of selected genera relevant to drinking water systems. Annotations are based on CAT
516   classification and support from Kaiju or/and Kraken. (B) Ratio of copy numbers of targets (i.e. *Legionella pneumophila*,

517 *Legionella spp.*, pathogenic Leptospira, and *Mycobacterium spp.*) to 16S rRNA gene copies. Colors and shapes correspond to
518 timepoints. Legionella spp. and Mycobacterium spp. are ubiquitous and more abundant than other targets. *Pseudomonas*
519 *aeruginosa* or *Mycobacterium avium* were not detected in any samples.
520

521 **3.4 A small fraction of recovered metagenome assembled genomes were associated with**

522 **pathogens.** Metagenomes were co-assembled by location, binned, and manually refined with

523 anvi'o. 105 bacterial MAGs were recovered after dereplication with dRep and quality filtering

524 for completeness greater than 50% and percent redundancy lower than 10%. We identified one

525 or more 16S rRNA in 39% of the MAGs. Further, we compared the differences in abundances

526 between samples by accounting for MAGs read recruitment. Of these, 37% of the MAGs were

527 detected in a quarter or more of the samples (Figure 6, Table S12). The PR MAGs were shared

528 homology with 4.55% of a recently published JGI MAG collection based on ANI values ranging

529 from 74.65-99.49%. JGI MAGs ecosystem categories represent aquatic (36.75%), human

530 (31.31%), terrestrial (6.5%), built environment (5.03%), and wastewater (5%) environments. In

531 contrast, the ecosystem distribution of the PR MAGs pairwise comparisons with JGI MAGs was

532 comprised of aquatic (32.49%), terrestrial (20.57%), built environment (13.64%), plants (12.5%),

533 and lab enrichment (4.95%) habitats. The aquatic ecosystem had the largest number of same

534 species representation (20% of PR MAGs, Figure 6) with species boundaries level set at 83%

535 cutoff threshold(Jain et al., 2018), and included 2 JGI MAGs also recovered from DWS.

536

537 Despite this observation, the JGI MAG dataset suffers from lack of representation of MAGs

538 assembled from DWS habitats (n=7). Therefore, a complimentary approach was used by

539 mapping metagenomic reads from diverse ecosystems against the MAGs assembled in this study

540 using the SearchSRA tool. This analysis indicated that the aquatic ecosystem was found to be the

541 top environmental association for 63.8% of our recovered MAGs, the other top ecosystems were

542 terrestrial (21.9%), human (13.3%), and mammal (1.0%) associated environments (Figure 6).

543 However, if we consider the top four environments, aquatic ecosystem category is represented in

544 all of our MAGs. There were statistically significant differences (ANOVA, $p<0.001$) between

545 the proportion of reads mapping from each ecosystem category mapping to the PR MAGs, the

546 only pairwise comparisons that were not statistically significant were terrestrial vs aquatic

547 (Tukey's, $p=0.96$) and mammals vs human environments (Tukey's, $p=0.06$). Additionally, more

548 than 13% of the aquatic ecosystem metagenomes were associated with DWS. Altogether, these

22

549   analyses show that our MAGs are widely distributed in the environment, but are largely
550   associated with aquatic and DWS associated environments.

551

552   The classification of resulting representative MAGs consisted of 64.8% Proteobacteria, followed
553   by 14.3% Cyanobacteria, 12.4% Planctomycetota, and 3.81% Actinobacteriota. All of the
554   Actinobacteria MAGs were classified as Mycobacterium. More than 50% of the MAGs were not
555   classified to genus level. The most abundant genera among the MAGs included
556   Hyphomicrobium (n=6), Bradyrhizobium (n=4), Gemmata (n=4), Mycobacterium (n=4), and
557   Porphyrobacter (n=4). There was no relationship between environmental parameters and MAG
558   abundance as assessed by Mantel statistic (r =0.133, p > 0.05) or constrained redundancy
559   analyses. Three of the four Mycobacterium MAGs were classified up to species level and
560   correspond to *Mycobacterium gordonae*, *Mycobacterium paragordonae*, and *Mycobacterium*
561   *phocaicum,* all of which have been recovered from drinking water systems previously and are
562   associated with infections in immunocompromised individuals(Shachor-Meyouhas et al., 2014).
563   M. *gordonae* was more abundant (RPKM=2.27±2.77) and frequently detected (58%) than M.
564   *paragordonae* (RPKM=0.77±0.41, 42% detection). Nevertheless, their abundance and frequency
565   of detection was higher than for M. *phocaicum* (RPKM=1.8±1.9, 15% detection).
566   Mycobacterium MAGs were not detected from SJU, and infrequently detected at CAR, GUA,
567   and HUM. A single Pseudomonas MAG was recovered and classified as *Pseudomonas*
568   *alcaligenes*. This MAG was only detected once and at the first timepoint at CAR, HAT, AGU,
569   and MAY. At the MAN location, it was detected in the initial and final timepoint and at higher
570   abundance at the final timepoint. *Pseudomonas alcaligenes* carries multiple antibiotic resistance
571   genes, are considered opportunistic human pathogens and have been identified in previous
572   literature characterizing drinking water systems, particularly in chlorinated systems(Jia et al.,
573   2019; Ma et al., 2019).

574
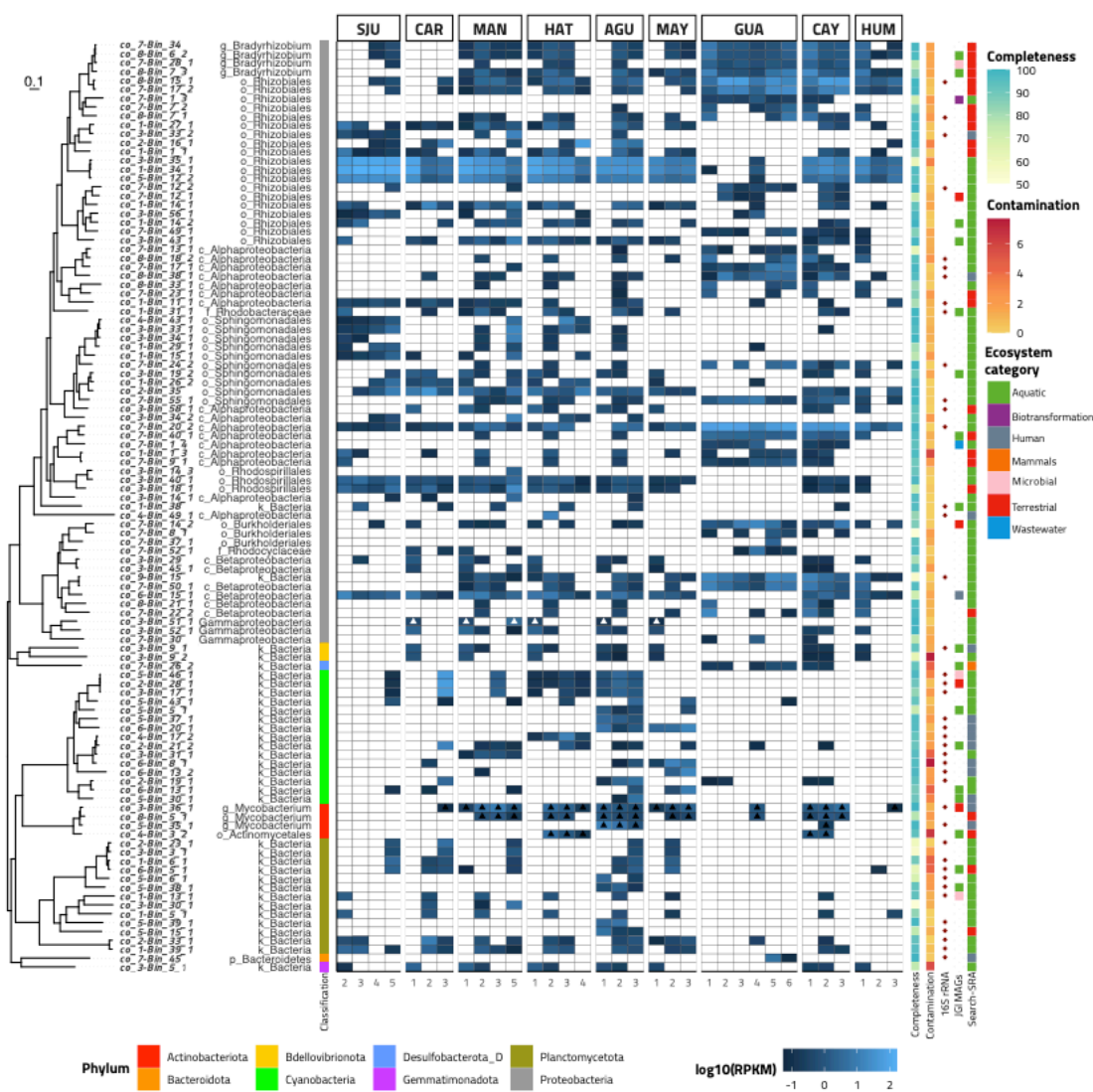
Figure 6: Metagenomic assembled genomes (MAG) information. From left to right: Phylogenomic tree of 105 recovered MAGs. Each branch is labelled by MAG code (italics) and marker lineage of MAG classified by GTDB-tk and annotated by color of phylum. Heatmap rows correspond to respective MAG and color gradient denotes log10(RPKM) values of MAGs detected in samples. A MAG was considered detected if ≥ 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 50-100% with color gradient from pale green to blue increasing value and contamination ranges from 0-10% 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.

4. Conclusions

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

24

592 manage microbial concentration at the taps, yet sampling locations showed significant variation
593 in the earlier timepoints. The estimated bacterial concentrations based on 16S rRNA gene
594 abundance at the sampling locations were consistent with literature established values
595 characterizing DWSs and generally decreased over time. Additionally, members of the microbial
596 community were comparable to those found in other DWSs which were not impacted by natural
597 disasters. Regardless of the ubiquity of some targeted OPPPs, such as *Legionella spp.* and
598 *Mycobacterium spp.*, they were present at low concentrations. Interestingly, pathogenic
599 Leptospira was only detected at a single location and its presence could be associated with a lack
600 of disinfectant residual at that site. A small fraction of metagenome assembled genomes were
601 associated with potential pathogens, and other recovered MAGs represent previously reported
602 taxa routinely found in drinking water systems. Altogether, the water disruptions (i.e., no water
603 or intermittent supply) that were sustained after HM did not have a significant impact on the
604 microbiological quality of drinking water in our study sites.

605

Data availability

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

610

618

References (MLA)

620 Allaire, M., Wu, H., Lall, U., 2018. National trends in drinking water quality violations 115, 1–6.
621   https://doi.org/10.1073/pnas.1719805115
622 Alneberg, J., Bjarnason, B.S., Bruijn, I. De, Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson,
623   A.F., Quince, C., 2014. Binning metagenomic contigs by coverage and composition 11.

624   https://doi.org/10.1038/nmeth.3103

625   Anuj, S.N., Whiley, D.M., Kidd, T.J., Bell, S.C., Wainwright, C.E., Nissen, M.D., Sloots, T.P., 2009. Identification

626   of Pseudomonas aeruginosa by a duplex real-time polymerase chain reaction assay targeting the ecfX and the

627   gyrB genes. Diagn. Microbiol. Infect. Dis. 63, 127–131. https://doi.org/10.1016/j.diagmicrobio.2008.09.018

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

629   Ashbolt, N.J., 2015. Microbial Contamination of Drinking Water and Human Health from Community Water

630   Systems 95–106. https://doi.org/10.1007/s40572-014-0037-5

631   Bautista-de los Santos, Q.M., Chavarria, K.A., Nelson, K.L., 2019. Understanding the impacts of intermittent supply

632   on the drinking water microbiome. Curr. Opin. Biotechnol. 57, 167–174.

633   https://doi.org/10.1016/j.copbio.2019.04.003

634   Brown, P., Ve, C.M., Murphy, C.B., Welton, M., Torres, H., Rosario, Z., Alshawabkeh, A., Cordero, F., Padilla,

635   I.Y., Meeker, J.D., 2018. Hurricanes and the Environmental Justice Island : Irma and Maria in Puerto Rico 11,

636   148–153. https://doi.org/10.1089/env.2018.0003

637   Buchfink, B., Xie, C., Huson, D.H., 2014. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12,

638   59–60. https://doi.org/10.1038/nmeth.3176

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

640   Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight,

641   R., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl.

642   Acad. Sci. U. S. A. 108, 4516–4522. https://doi.org/10.1073/pnas.1000080107

643   Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. Fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34,

644   i884–i890. https://doi.org/10.1093/bioinformatics/bty560

645   Chern, E.C., King, D., Haugland, R., Pfaller, S., 2015. Evaluation of quantitative polymerase chain reaction assays

646   targeting Mycobacterium avium, M. intracellulare, and M. avium subspecies paratuberculosis in drinking

647   water biofilms. J. Water Health 13, 131–139. https://doi.org/10.2166/wh.2014.060

648   Collier, S.A., Deng, L., Adam, E.A., Benedict, K.M., Beshearse, E.M., Blackstock, A.J., Bruce, B.B., Derado, G.,

649   Edens, C., Fullerton, K.E., Gargano, J.W., Geissler, A.L., Hall, A.J., Havelaar, A.H., Hill, V.R., Hoekstra,

650   R.M., Reddy, S.C., Scallan, E., Stokes, E.K., Yoder, J.S., Beach, M.J., 2021. Estimate of burden and direct

651   healthcare cost of infectious waterborne disease in the United States. Emerg. Infect. Dis. 27, 140–149.

652   https://doi.org/10.3201/eid2701.190676

653   Dai, Z., Sevillano-Rivera, M., Calus, S., Bautista-de los Santos, Q.M., Eren, A.M., van der Wielen, P.W.J.J., Ijaz,

654   U.Z., Pinto, A., 2020. Disinfection exhibits systematic impacts on the drinking water microbiome.

655   Microbiome.

656   Douterelo, I., Calero-Preciado, C., Soria-Carrasco, V., Boxall, J.B., 2018. Whole metagenome sequencing of

657   chlorinated drinking water distribution systems. Environ. Sci. Water Res. Technol. 4, 2080–2091.

658   https://doi.org/10.1039/c8ew00395e

659   Eren, A.M., Esen, C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., Delmont, T.O., 2015. Anvi ' o : an

660   advanced analysis and visualization platform for ' omics data 1–29. https://doi.org/10.7717/peerj.1319

661 Estrada, F., Botzen, W.J.W., Tol, R.S.J., 2015. with an influence from climate change 8, 6–11.

662   https://doi.org/10.1038/NGEO2560

663 Exum, N.G., Betanzo, E., Schwab, K.J., Chen, T.Y.J., Guikema, S., Harvey, D.E., Harvey, D.E., 2018. Extreme

664   Precipitation , Public Health Emergencies , and Safe Drinking Water in the USA 305–315.

665 Gooddy, D.C., Lapworth, D.J., Ascott, M.J., Bennett, S.A., Heaton, T.H.E., Surridge, B.W.J., 2015. Isotopic

666   Fingerprint for Phosphorus in Drinking Water Supplies. Environ. Sci. Technol. 49, 9020–9028.

667   https://doi.org/10.1021/acs.est.5b01137

668 Goodess, C.M., 2012. How is the frequency , location and severity of extreme events likely to change up to 2060 ?

669   Environ. Sci. Policy 27, S4–S14. https://doi.org/10.1016/j.envsci.2012.04.001

670 Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: Quality assessment tool for genome assemblies.

671   Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

672 Huang, J., Chen, S., Ma, X., Yu, P., Zuo, P., Shi, B., Wang, H., Alvarez, P.J.J., 2021. Opportunistic pathogens and

673   their health risk in four full-scale drinking water treatment and distribution systems. Ecol. Eng. 160.

674   https://doi.org/10.1016/j.ecoleng.2020.106134

675 Hull, N.M., Holinger, E.P., Ross, K.A., Robertson, C.E., Harris, J.K., Stevens, M.J., Pace, N.R., 2017. Longitudinal

676   and Source-to-Tap New Orleans, LA, U.S.A. Drinking Water Microbiology. Environ. Sci. Technol. 51, 4220–

677   4229. https://doi.org/10.1021/acs.est.6b06064

678 Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: Prokaryotic gene

679   recognition and translation initiation site identification. BMC Bioinformatics 11. https://doi.org/10.1186/1471-

680   2105-11-119

681 Inkinen, J., Jayaprakash, B., Siponen, S., Hokajärvi, A.M., Pursiainen, A., Ikonen, J., Ryzhikov, I., Täubel, M.,

682   Kauppinen, A., Paananen, J., Miettinen, I.T., Torvinen, E., Kolehmainen, M., Pitkänen, T., 2019. Active

683   eukaryotes in drinking water distribution systems of ground and surface waterworks. Microbiome 7, 1–17.

684   https://doi.org/10.1186/s40168-019-0715-5

685 Isaac, T.S., Sherchan, S.P., 2020. Molecular detection of opportunistic premise plumbing pathogens in rural

686   Louisiana's drinking water distribution system. Environ. Res. 181, 108847.

687   https://doi.org/10.1016/j.envres.2019.108847

688 Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., Aluru, S., 2018. High throughput ANI analysis

689   of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9, 1–8.

690   https://doi.org/10.1038/s41467-018-07641-9

691 Jia, S., Wu, J., Ye, L., Zhao, F., Li, T., Zhang, X.X., 2019. Metagenomic assembly provides a deep insight into the

692   antibiotic resistome alteration induced by drinking water chlorination and its correlations with bacterial host

693   changes. J. Hazard. Mater. 379, 120841. https://doi.org/10.1016/j.jhazmat.2019.120841

694 Jiang, S.C., Han, M., Chandrasekaran, S., Fang, Y., Kellogg, C.A., 2020. Assessing the water quality impacts of two

695   Category-5 hurricanes on St. Thomas, Virgin Islands. Water Res. 171, 115440.

696   https://doi.org/10.1016/j.watres.2019.115440

697 Keenum, I., Medina, M.C., Garner, E., Pieper, K.J., Blair, M.F., Milligan, E., Pruden, A., Ramirez-Toro, G.,

698   Rhoads, W.J., 2021. Source-to-Tap Assessment of Microbiological Water Quality in Small Rural Drinking
699   Water Systems in Puerto Rico Six Months After Hurricane Maria. Environ. Sci. Technol.
700   https://doi.org/10.1021/acs.est.0c08814

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

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

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

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

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

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

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

719   Lin, Y., Sevillano-Rivera, M., Jiang, T., Li, G., Cotto, I., Vosloo, S., Carpenter, C.M.G., Larese-Casanova, P., Giese,
720   R.W., Helbling, D.E., Padilla, I.Y., Rosario-Pabón, Z., Vélez Vega, C., Cordero, J.F., Alshawabkeh, A.N.,
721   Pinto, A., Gu, A.Z., 2020. Impact of Hurricane Maria on Drinking Water Quality in Puerto Rico. Environ. Sci.
722   Technol. 54, 9495–9509. https://doi.org/10.1021/acs.est.0c01655

723   Liu, L., Xing, X., Hu, C., Wang, H., 2019. One-year survey of opportunistic premise plumbing pathogens and free-
724   living amoebae in the tap-water of one northern city of China. J. Environ. Sci. (China) 77, 20–31.
725   https://doi.org/10.1016/j.jes.2018.04.020

726   Lu, J., Struewing, I., Vereen, E., Kirby, A.E., Levy, K., Moe, C., Ashbolt, N., 2016. Molecular Detection of
727   Legionella spp. and their associations with Mycobacterium spp., Pseudomonas aeruginosa and amoeba hosts
728   in a drinking water distribution system. J. Appl. Microbiol. 120, 509–521. https://doi.org/10.1111/jam.12996

729   Ma, L., Li, B., Zhang, T., 2019. New insights into antibiotic resistome in drinking water and management
730   perspectives: A metagenomic based study of small-sized microbes. Water Res. 152, 191–201.
731   https://doi.org/10.1016/j.watres.2018.12.069

732   Maness, L.R., 2019. The Effect of Hurricanes on Pathogenic Diseases. J. Environ. Health 81, 16–20.

733   Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju.
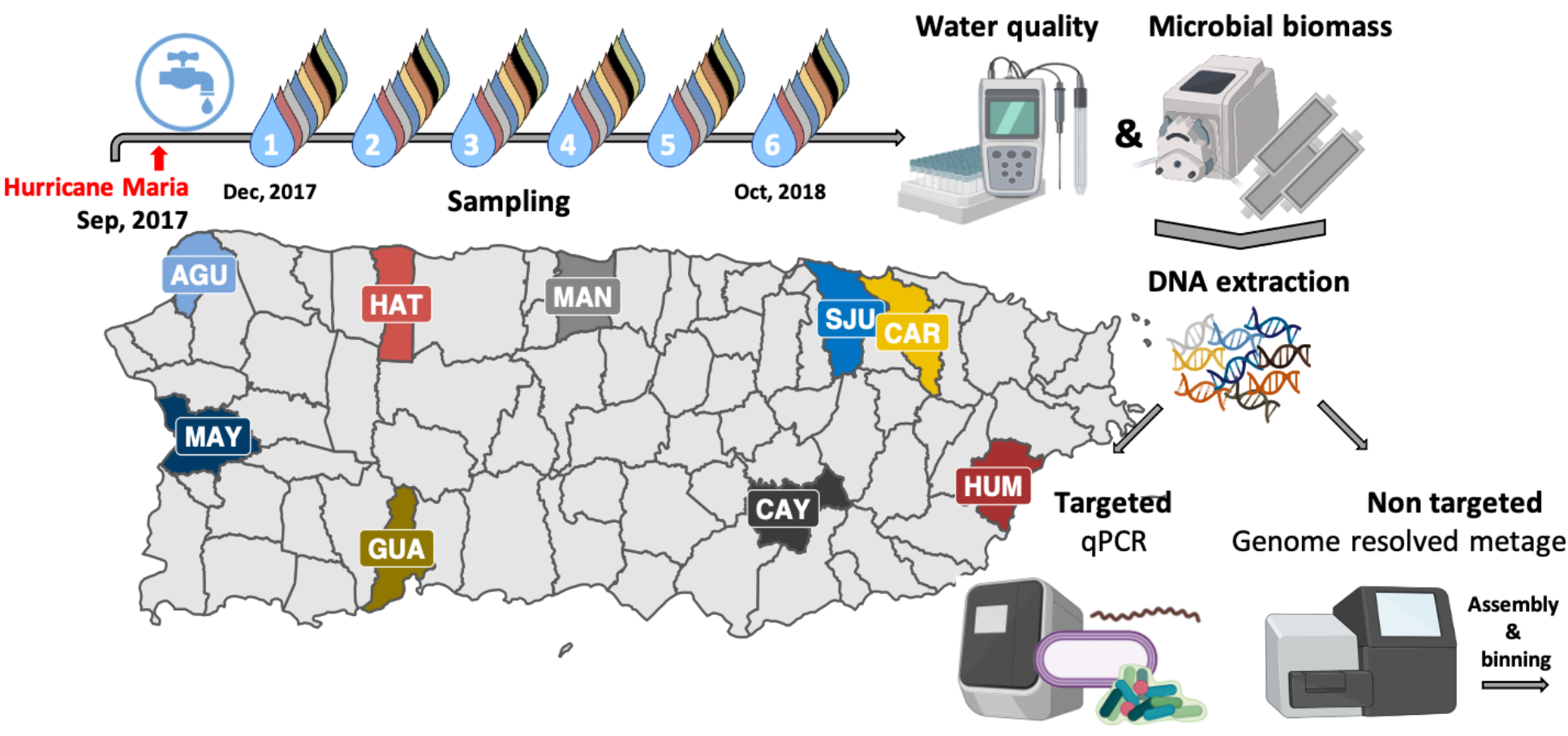734   Nat. Commun. 7, 1–9. https://doi.org/10.1038/ncomms11257

735 Nayfach, S., Roux, S., Seshadri, R., Udwary, D., Varghese, N., Schulz, F., Wu, D., Paez-Espino, D., Chen, I.M.,

736   Huntemann, M., Palaniappan, K., Ladau, J., Mukherjee, S., Reddy, T.B.K., Nielsen, T., Kirton, E., Faria, J.P.,

737   Edirisinghe, J.N., Henry, C.S., Jungbluth, S.P., Chivian, D., Dehal, P., Wood-Charlson, E.M., Arkin, A.P.,

738   Tringe, S.G., Visel, A., Abreu, H., Acinas, S.G., Allen, E., Allen, M.A., Andersen, G., Anesio, A.M.,

739   Attwood, G., Avila-Magaña, V., Badis, Y., Bailey, J., Baker, B., Baldrian, P., Barton, H.A., Beck, D.A.C.,

740   Becraft, E.D., Beller, H.R., Beman, J.M., Bernier-Latmani, R., Berry, T.D., Bertagnolli, A., Bertilsson, S.,

741   Bhatnagar, J.M., Bird, J.T., Blumer-Schuette, S.E., Bohannan, B., Borton, M.A., Brady, A., Brawley, S.H.,

742   Brodie, J., Brown, S., Brum, J.R., Brune, A., Bryant, D.A., Buchan, A., Buckley, D.H., Buongiorno, J.,

743   Cadillo-Quiroz, H., Caffrey, S.M., Campbell, A.N., Campbell, B., Carr, S., Carroll, J.L., Cary, S.C., Cates,

744   A.M., Cattolico, R.A., Cavicchioli, R., Chistoserdova, L., Coleman, M.L., Constant, P., Conway, J.M., Mac

745   Cormack, W.P., Crowe, S., Crump, B., Currie, C., Daly, R., Denef, V., Denman, S.E., Desta, A., Dionisi, H.,

746   Dodsworth, J., Dombrowski, N., Donohue, T., Dopson, M., Driscoll, T., Dunfield, P., Dupont, C.L., Dynarski,

747   K.A., Edgcomb, V., Edwards, E.A., Elshahed, M.S., Figueroa, I., Flood, B., Fortney, N., Fortunato, C.S.,

748   Francis, C., Gachon, C.M.M., Garcia, S.L., Gazitua, M.C., Gentry, T., Gerwick, L., Gharechahi, J., Girguis,

749   P., Gladden, J., Gradoville, M., Grasby, S.E., Gravuer, K., Grettenberger, C.L., Gruninger, R.J., Guo, J.,

750   Habteselassie, M.Y., Hallam, S.J., Hatzenpichler, R., Hausmann, B., Hazen, T.C., Hedlund, B., Henny, C.,

751   Herfort, L., Hernandez, M., Hershey, O.S., Hess, M., Hollister, E.B., Hug, L.A., Hunt, D., Jansson, J., Jarett,

752   J., Kadnikov, V. V., Kelly, C., Kelly, R., Kelly, W., Kerfeld, C.A., Kimbrel, J., Klassen, J.L., Konstantinidis,

753   K.T., Lee, L.L., Li, W.J., Loder, A.J., Loy, A., Lozada, M., MacGregor, B., Magnabosco, C., Maria da Silva,

754   A., McKay, R.M., McMahon, K., McSweeney, C.S., Medina, M., Meredith, L., Mizzi, J., Mock, T., Momper,

755   L., Moran, M.A., Morgan-Lang, C., Moser, D., Muyzer, G., Myrold, D., Nash, M., Nesbø, C.L., Neumann,

756   A.P., Neumann, R.B., Noguera, D., Northen, T., Norton, J., Nowinski, B., Nüsslein, K., O'Malley, M.A.,

757   Oliveira, R.S., Maia de Oliveira, V., Onstott, T., Osvatic, J., Ouyang, Y., Pachiadaki, M., Parnell, J., Partida-

758   Martinez, L.P., Peay, K.G., Pelletier, D., Peng, X., Pester, M., Pett-Ridge, J., Peura, S., Pjevac, P., Plominsky,

759   A.M., Poehlein, A., Pope, P.B., Ravin, N., Redmond, M.C., Reiss, R., Rich, V., Rinke, C., Rodrigues, J.L.M.,

760   Rossmassler, K., Sackett, J., Salekdeh, G.H., Saleska, S., Scarborough, M., Schachtman, D., Schadt, C.W.,

761   Schrenk, M., Sczyrba, A., Sengupta, A., Setubal, J.C., Shade, A., Sharp, C., Sherman, D.H., Shubenkova, O.

762   V., Sierra-Garcia, I.N., Simister, R., Simon, H., Sjöling, S., Slonczewski, J., Correa de Souza, R.S., Spear,

763   J.R., Stegen, J.C., Stepanauskas, R., Stewart, F., Suen, G., Sullivan, M., Sumner, D., Swan, B.K., Swingley,

764   W., Tarn, J., Taylor, G.T., Teeling, H., Tekere, M., Teske, A., Thomas, T., Thrash, C., Tiedje, J., Ting, C.S.,

765   Tully, B., Tyson, G., Ulloa, O., Valentine, D.L., Van Goethem, M.W., VanderGheynst, J., Verbeke, T.J.,

766   Vollmers, J., Vuillemin, A., Waldo, N.B., Walsh, D.A., Weimer, B.C., Whitman, T., van der Wielen, P.,

767   Wilkins, M., Williams, T.J., Woodcroft, B., Woolet, J., Wrighton, K., Ye, J., Young, E.B., Youssef, N.H., Yu,

768   F.B., Zemskaya, T.I., Ziels, R., Woyke, T., Mouncey, N.J., Ivanova, N.N., Kyrpides, N.C., Eloe-Fadrosh,

769   E.A., 2020. A genomic catalog of Earth's microbiomes. Nat. Biotechnol. 39. https://doi.org/10.1038/s41587-

770   020-0718-6

771 Nazarian, E.J., Bopp, D.J., Saylors, A., Limberger, R.J., Musser, K.A., 2008. Design and implementation of a

772   protocol for the detection of Legionella in clinical and environmental samples. Diagn. Microbiol. Infect. Dis.

773   62, 125–132. https://doi.org/10.1016/j.diagmicrobio.2008.05.004

774   NOAA NCEI, 2020. Billion-Dollar Weather and Climate Disasters [WWW Document].

775   https://doi.org/10.25921/stkw-7w73

776   NRDC, 2017. Threats on Tap: Drinking Water Violations in Puerto Rico.

777   Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. MetaSPAdes: A new versatile metagenomic

778   assembler. Genome Res. 27, 824–834. https://doi.org/10.1101/gr.213959.116

779   Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P.,

780   Stevens, M.H.H., Wagner, H., 2015. OK-Package 'vegan.' Community Ecol. Packag. version.

781   Olm, M.R., Brown, C.T., Brooks, B., Banfield, J.F., 2017. dRep : a tool for fast and accurate genomic comparisons

782   that enables improved genome recovery from metagenomes through de-replication 11, 2864–2868.

783   https://doi.org/10.1038/ismej.2017.126

784   Olsen, L.R., 2021. groupdata2: Creating Groups from Data.

785   Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., Phillippy, A.M., 2016. Mash:

786   Fast genome and metagenome distance estimation using MinHash. Genome Biol. 17, 1–14.

787   https://doi.org/10.1186/s13059-016-0997-x

788   Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of phylogenetics and evolution in R language.

789   Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412

790   Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., Hugenholtz, P., 2018. A

791   standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat.

792   Biotechnol. 36, 996. https://doi.org/10.1038/nbt.4229

793   Patterson, C.L., Adams, J.Q., 2011. Emergency response planning to reduce the impact of contaminated drinking

794   water during natural disasters 5, 341–349. https://doi.org/10.1007/s11707-011-0196-8

795   Quinlan, A.R., Hall, I.M., 2010. BEDTools: A flexible suite of utilities for comparing genomic features.

796   Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033

797   R Development Core Team, 2016. R: A language and environment for statistical computing. R Found. Stat. Comput.

798   https://doi.org/10.1017/CBO9781107415324.004

799   Radomski, N., Lucas, F.S., Moilleron, R., Cambau, E., Haenn, S., Moulin, L., 2010. Development of a real-time

800   qpcr method for detection and enumeration of Mycobacterium spp. in surface water. Appl. Environ.

801   Microbiol. 76, 7348–7351. https://doi.org/10.1128/AEM.00942-10

802   Rice, P., Longden, L., Bleasby, A., 2000. EMBOSS: The European Molecular Biology Open Software Suite. Trends

803   Genet. 16, 276–277. https://doi.org/10.1016/S0168-9525(00)02024-2

804   Rodriguez-R, L.M., Gunturu, S., Tiedje, J.M., Cole, J.R., Konstantinidis, K.T., 2018. Nonpareil 3: Fast Estimation

805   of Metagenomic Coverage and Sequence Diversity. mSystems 3, 1–9.

806   https://doi.org/10.1128/msystems.00039-18

807   Sanders, E.J., Rigau-Pérez, J.G., Smits, H.L., Deseda, C.C., Vorndam, V.A., Aye, T., Spiegel, R.A., Weyant, R.S.,

808   Bragg, S.L., 1999. Increase of leptospirosis in dengue-negative patients after a Hurricane in Puerto Rico in

809   1966. Am. J. Trop. Med. Hyg. 61, 399–404. https://doi.org/10.4269/ajtmh.1999.61.399

810   Schwab, K.J., Gibson, K.E., Williams, D.L., Kulbicki, K.M., Lo, C.P., Mihalic, J.N., Breysse, P.N., Curriero, F.C.,

811   Geyh, A.S., 2007. Microbial and chemical assessment of regions within New Orleans, LA impacted by

812   Hurricane Katrina. Environ. Sci. Technol. 41, 2401–2406. https://doi.org/10.1021/es062916x

813   Shachor-Meyouhas, Y., Geffen, Y., Arad-Cohen, N., Zaidman, I., Ben-Barak, A., Davidson, S., Kassis, I., 2014.

814   Mycobacterium phocaicum bacteremia: An emerging infection in pediatric hematology-oncology patients.

815   Pediatr. Infect. Dis. J. 33, 1299–1301. https://doi.org/10.1097/INF.0000000000000477

816   Shaffer, M., Borton, M.A., McGivern, B.B., Zayed, A.A., La Rosa, S.L., Solden, L.M., Liu, P., Narrowe, A.B.,

817   Rodríguez-Ramos, J., Bolduc, B., Gazitúa, M.C., Daly, R.A., Smith, G.J., Vik, D.R., Pope, P.B., Sullivan,

818   M.B., Roux, S., Wrighton, K.C., 2020. DRAM for distilling microbial metabolism to automate the curation of

819   microbiome function. Nucleic Acids Res. 48, 8883–8900. https://doi.org/10.1093/nar/gkaa621

820   Shukla, M.A., Woc-Colburn, L., Weatherhead, J.E., 2018. Infectious Diseases in the Aftermath of Hurricanes in the

821   United States. Curr. Trop. Med. Reports 5, 217–223. https://doi.org/10.1007/s40475-018-0162-6

822   Stanish, L.F., Hull, N.M., Robertson, C.E., Kirk Harris, J., Stevens, M.J., Spear, J.R., Pace, N.R., 2016. Factors

823   influencing bacterial diversity and community composition in municipal drinking waters in the Ohio River

824   basin, USA. PLoS One 11, 1–21. https://doi.org/10.1371/journal.pone.0157966

825   Stewart, C.A., Hancock, D., Stanzioneb, D., Turnerd, G., Cockerill, T.M., Merchant, N., Taylor, J., Vaughn, M.,

826   Foster, I., Skidmore, E., Tuecke, S., Gaffney, N.I., 2015. Jetstream: A self-provisioned, scalable science and

827   engineering cloud environment. ACM Int. Conf. Proceeding Ser. 2015-July.

828   https://doi.org/10.1145/2792745.2792774

829   Stoddard, R.A., Gee, J.E., Wilkins, P.P., McCaustland, K., Hoffmaster, A.R., 2009. Detection of pathogenic

830   Leptospira spp. through TaqMan polymerase chain reaction targeting the LipL32 gene. Diagn. Microbiol.

831   Infect. Dis. 64, 247–255. https://doi.org/10.1016/j.diagmicrobio.2009.03.014

832   Sutter, J.D., Sosa Pascual, O., 2018. Deaths from bacterial disease in Puerto Rico spiked after Maria.

833   Torres, P.J., Edwards, R.A., McNair, K.A., 2017. PARTIE: A partition engine to separate metagenomic and

834   amplicon projects in the Sequence Read Archive. Bioinformatics 33, 2389–2391.

835   https://doi.org/10.1093/bioinformatics/btx184

836   Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D.,

837   Peterson, G.D., Roskies, R., Scott, J.R., Wilkins-Diehr, N., 2014. XSEDE: Accelerating Scientific Discovery.

838   Comput. Sci. Eng. 120, 62–74.

839   Truitt, Z.G., Poon-Kwong, B., Bachoon, D.S., Otero, E., 2020. Seasonal shifts in the presence of pathogenic

840   leptospires, Escherichia coli, and physicochemical properties in coastal rivers and streams of Puerto Rico. J.

841   Environ. Qual. 49, 1264–1272. https://doi.org/10.1002/jeq2.20091

842   Von Meijenfeldt, F.A.B., Arkhipova, K., Cambuy, D.D., Coutinho, F.H., Dutilh, B.E., 2019. Robust taxonomic

843   classification of uncharted microbial sequences and bins with CAT and BAT. Genome Biol. 20, 1–14.

844   https://doi.org/10.1186/s13059-019-1817-x

845   Vosloo, S., Sevillano-Rivera, M., Pinto, A.J., 2019. Modified DNeasy PowerWater Kit® protocol for DNA

846   extractions from drinking water samples. https://doi.org/10.17504/protocols.io.66khhcw

847   Walker, J.T., 2018. The influence of climate change on waterborne disease and Legionella: a review. Perspect.

848   Public Health 138, 282–286. https://doi.org/10.1177/1757913918791198

849   Wang, H., Edwards, M., Falkinham, J.O., Pruden, A., 2012. Molecular survey of the occurrence of legionella spp.,

850   mycobacterium spp., pseudomonas aeruginosa, and amoeba hosts in two chloraminated drinking water

851   distribution systems. Appl. Environ. Microbiol. 78, 6285–6294. https://doi.org/10.1128/AEM.01492-12

852   Warren, M.R., 2019. Challenges with sampling and understanding lead corrosion issues in Puerto Rico after

853   hurricane Maria. University of Colorado.

854   Wickham, H., 2011. Ggplot2. Wiley Interdiscip. Rev. Comput. Stat. 3, 180–185. https://doi.org/10.1002/wics.147

855   Wickham, H., Girlich, M., Ruiz, E., 2021. dbplyr: A "dplyr" Back End for Databases.

856   Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20, 1–13.

857   https://doi.org/10.1186/s13059-019-1891-0

858   Wynwood, S.J., Graham, G.C., Weier, S.L., Collet, T.A., McKay, D.B., Craig, S.B., 2014. Leptospirosis from water

859   sources. Pathog. Glob. Health 108, 334–338. https://doi.org/10.1179/2047773214Y.0000000156

860   Yáñez, M.A., Carrasco-Serrano, C., Barberá, V.M., Catalán, V., 2005. Quantitative detection of Legionella

861   pneumophila in water samples by immunomagnetic purification and real-time PCR amplification of the dotA

862   gene. Appl. Environ. Microbiol. 71, 3433–3441. https://doi.org/10.1128/AEM.71.7.3433-3441.2005
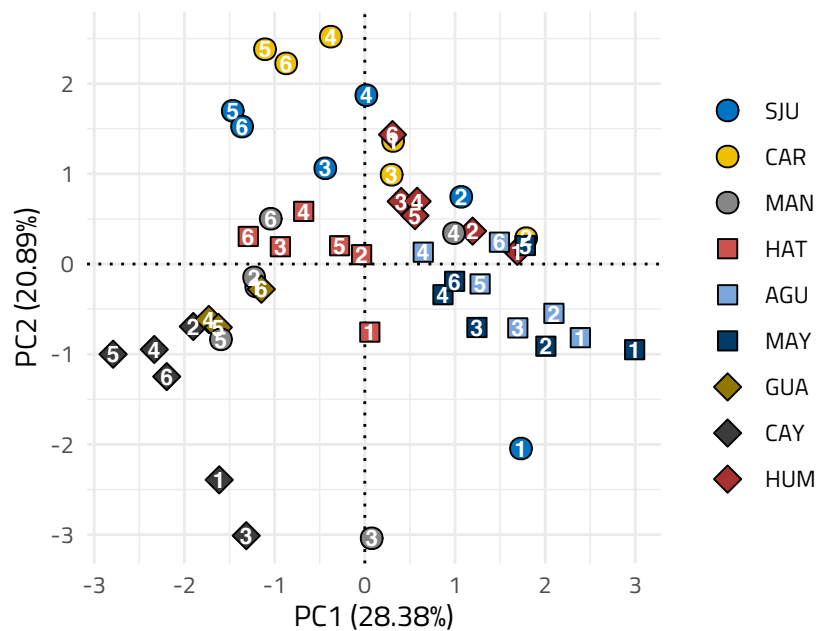
863

864

A **Water quality parameters**

Ammonia (mg L^{-1} NH$_3$ – N)

Nitrate (mg L^{-1} NO$_3^{-1}$ – N)

Chlorine (mg L^{-1} Cl$_2$)

Phosphate (mg L^{-1} PO$_4^{3-1}$)

Conductivity (µS cm^{-1})

DO (%sat)

pH

Temperature (°C)

TOC (mg L^{-1})

Timepoint

1 2 3 4 5 6

Parameter concentration

Location

SJU CAR MAN HAT AGU MAY GUA CAY HUM

B **PCA of water quality parameters**

PC2 (20.89%)

PC1 (28.38%)

SJU
CAR
MAN
HAT
AGU
MAY
GUA
CAY
HUM
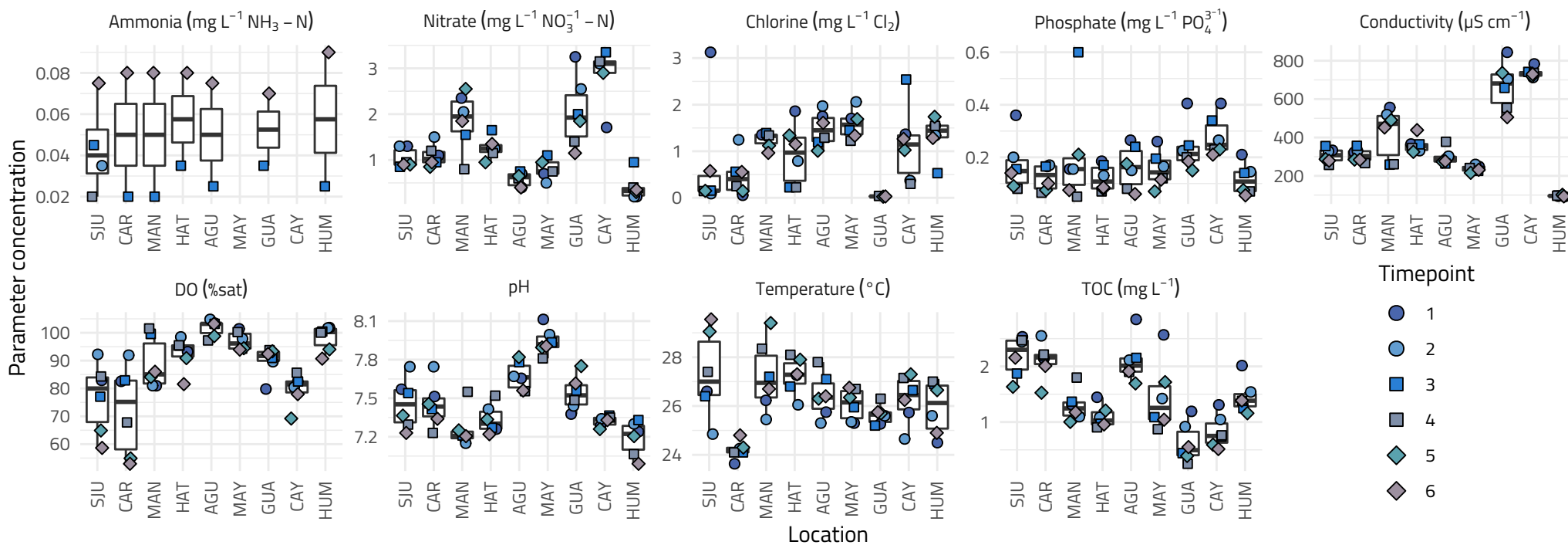
C **16S rRNA gene copies in samples and blanks across time points**

SJU   CAR   MAN

HAT   AGU   MAY

GUA   CAY   HUM

16S rRNA gene copies

Timepoint
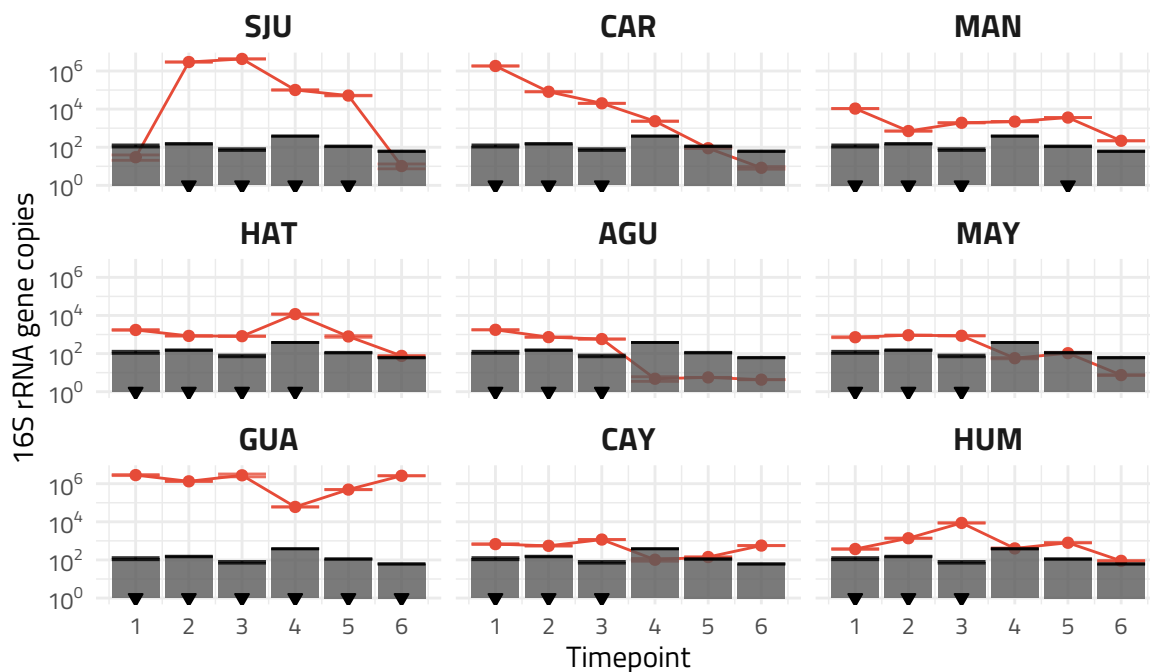
Correlation between volume normalized 16S rRNA gene copies and chlorine concentration (mg/L)
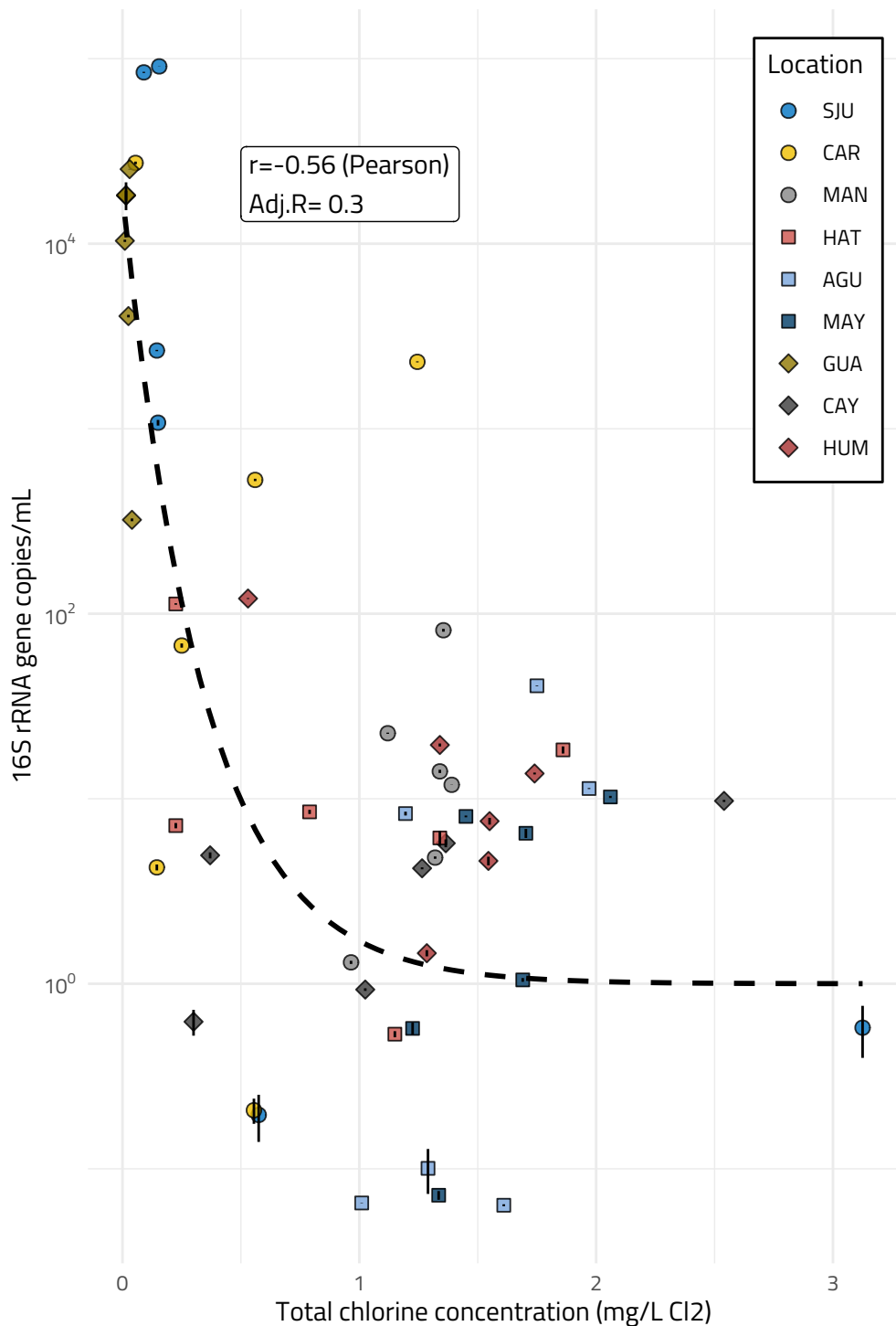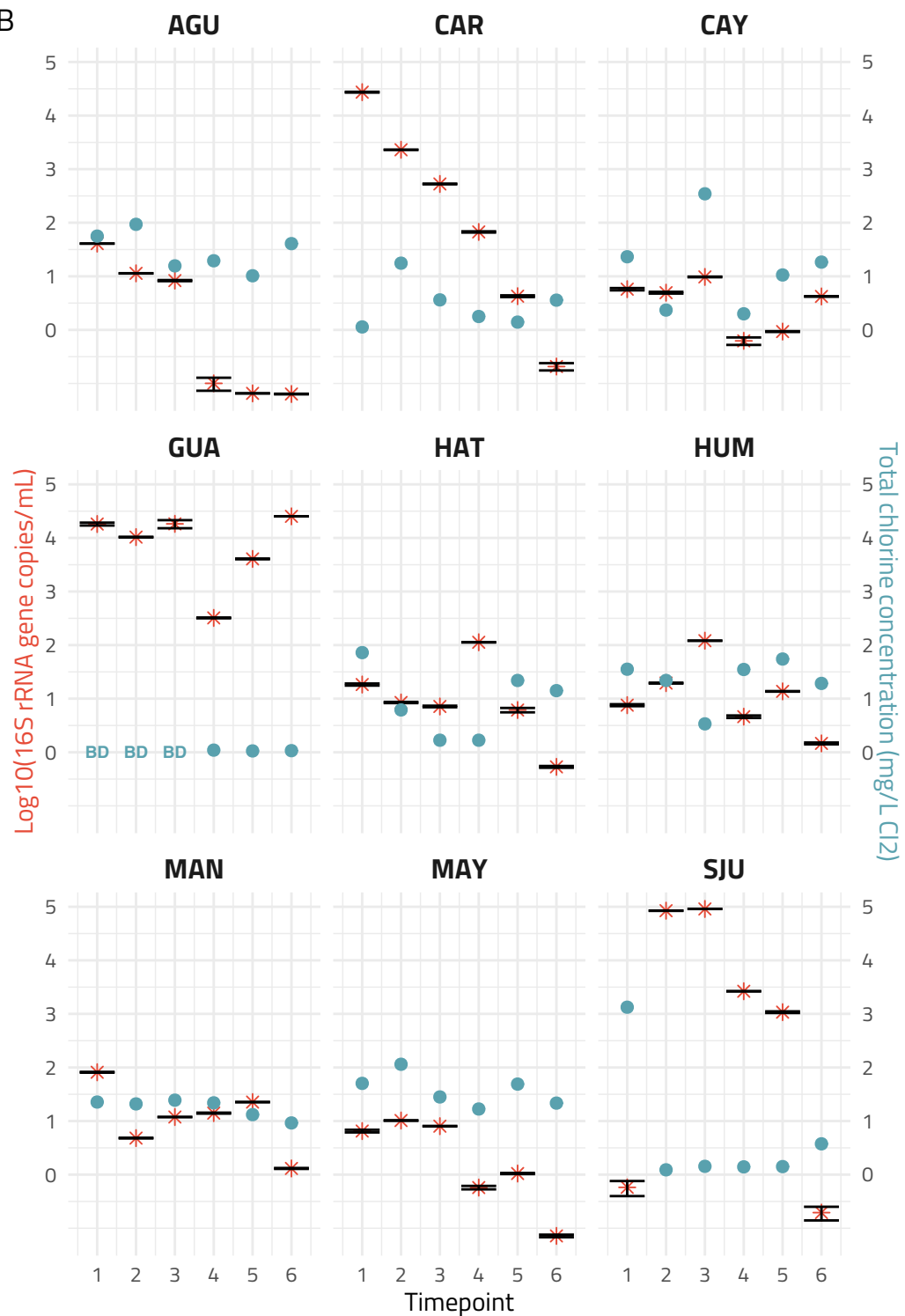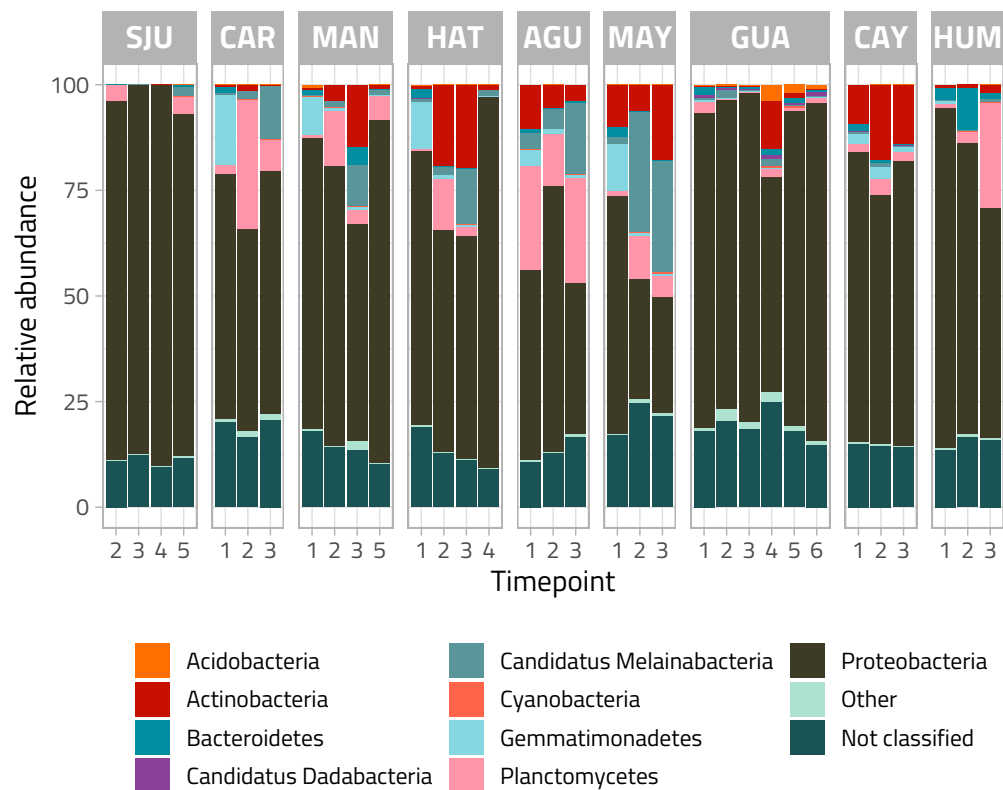
A Nonpareil curves per location
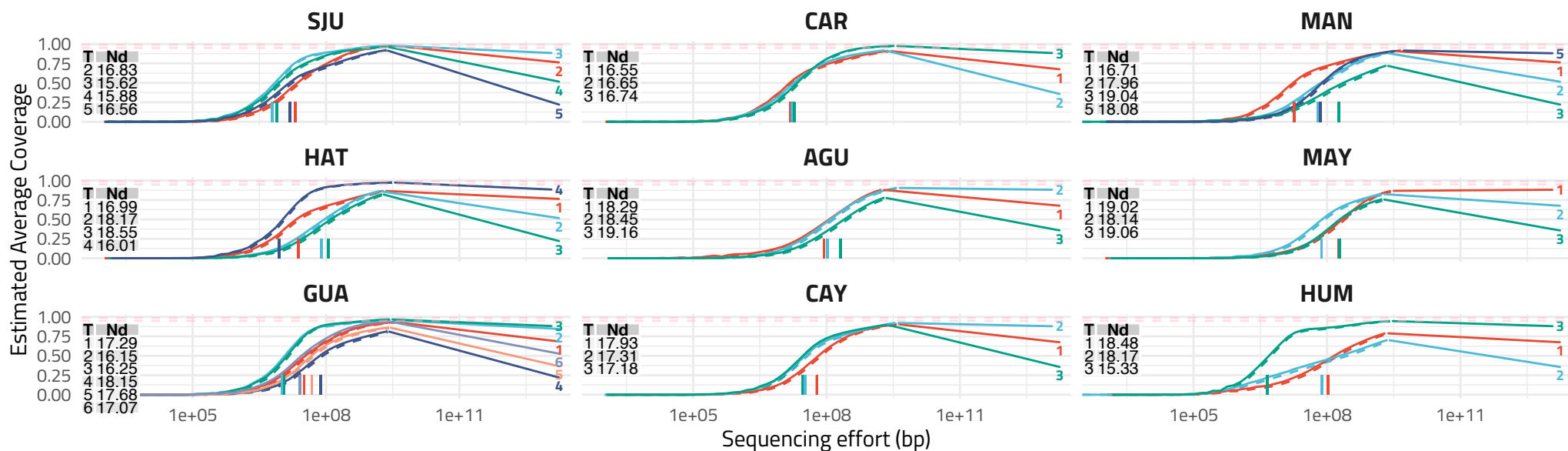
B Relative abundance of samples at domain level:
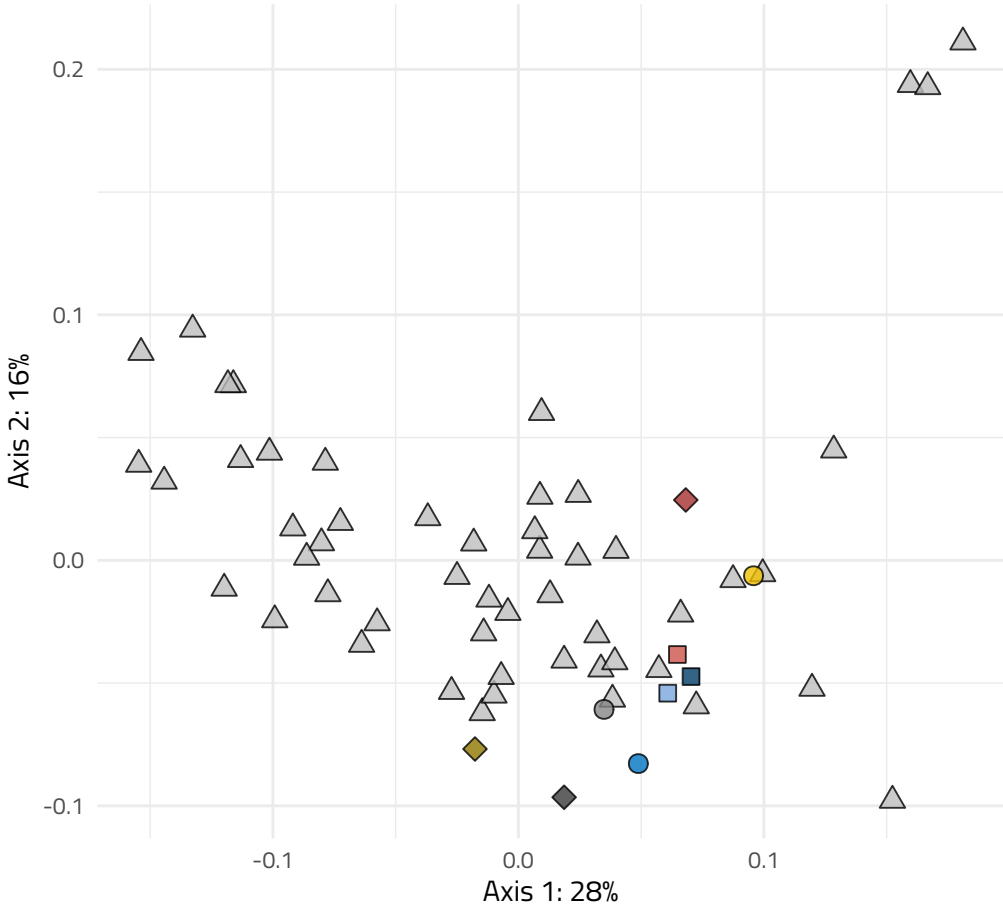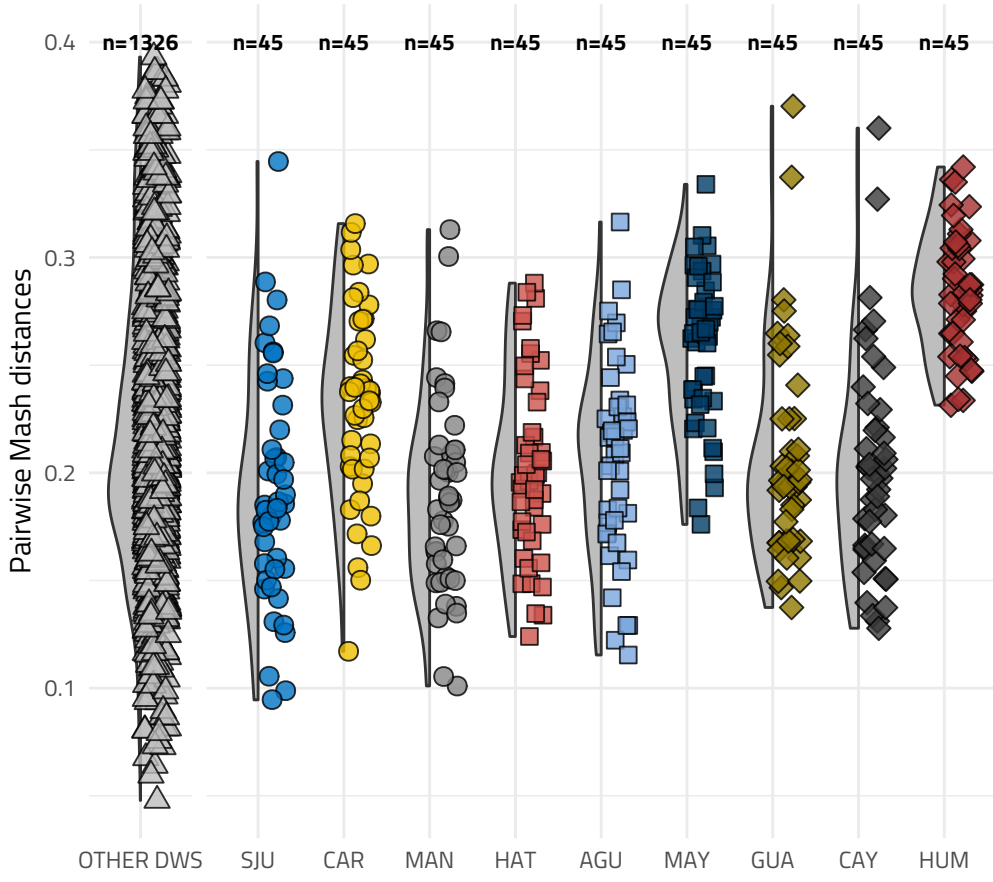Archaea, Bacteria, Eukaryota, Viruses, and Not Classified
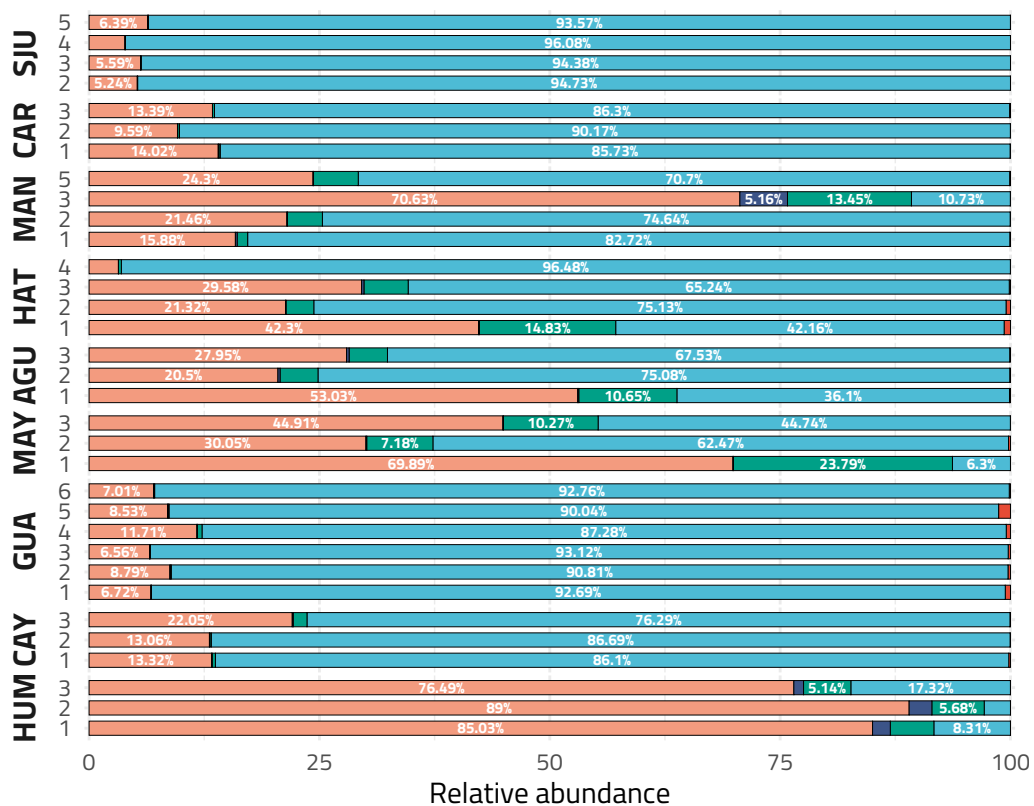
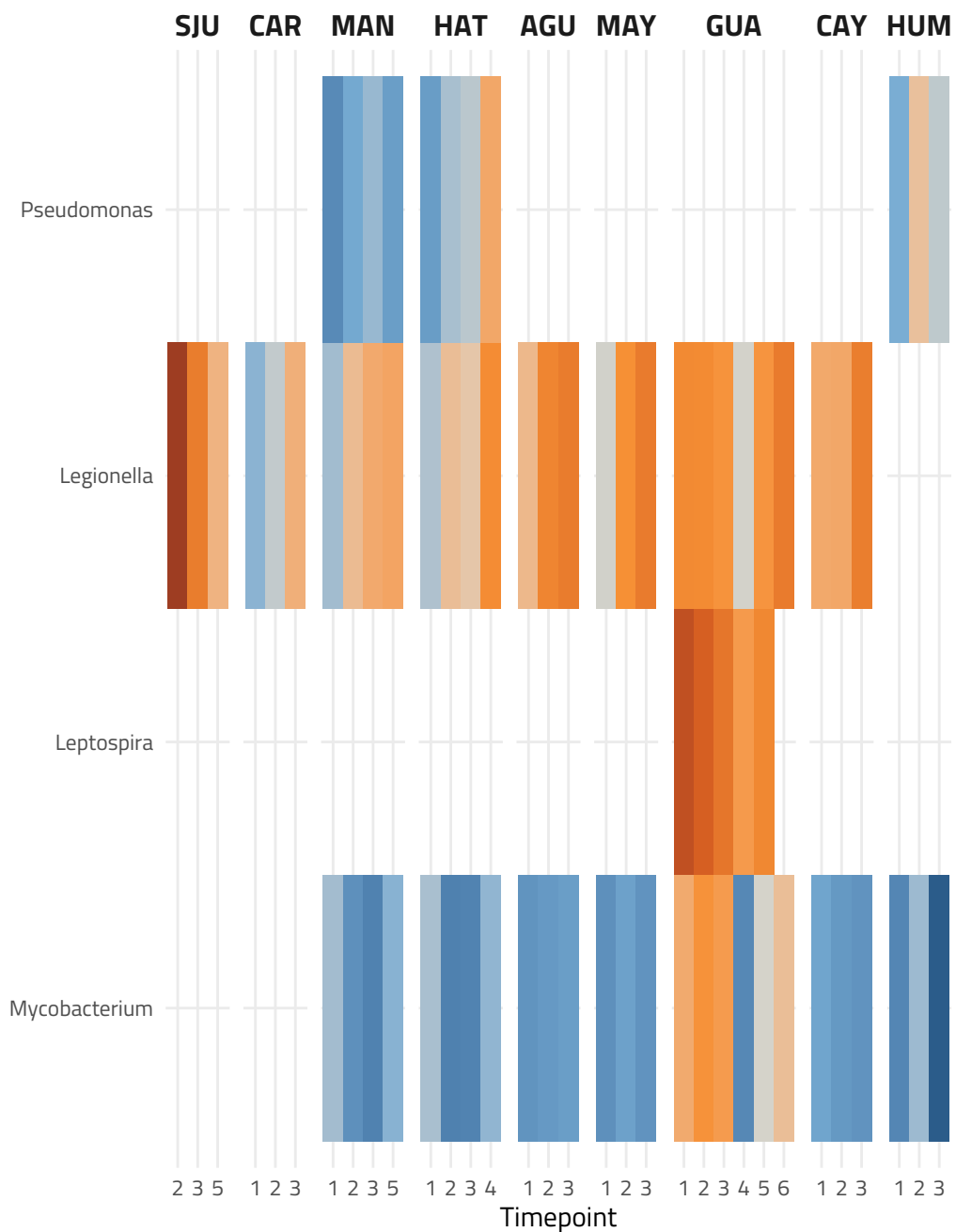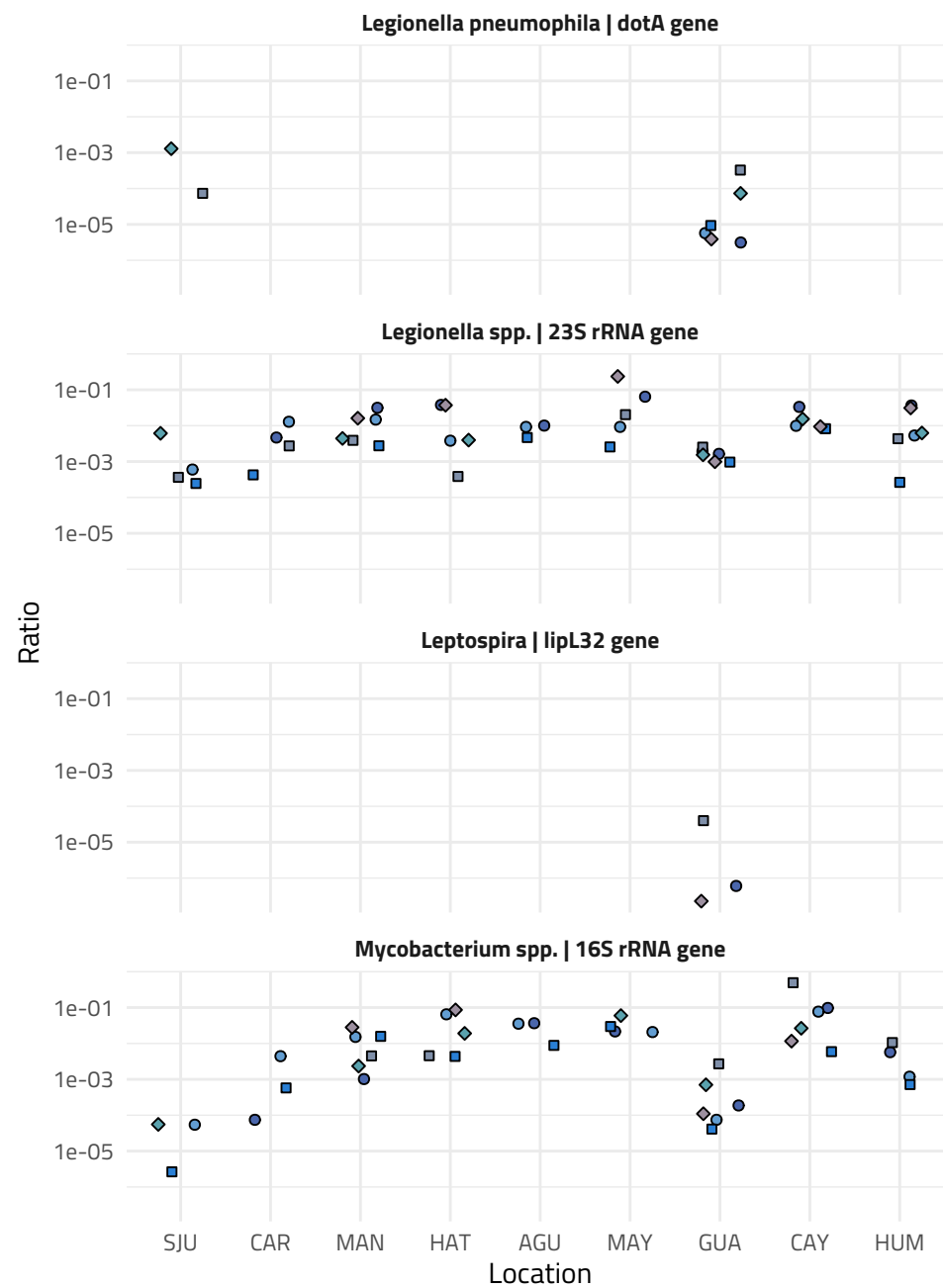**C** Relative abundance of bacterial phyla per sample

A

Axis 2: 16%

Axis 1: 28%

B

Pairwise Mash distances

n=1326 n=45 n=45 n=45 n=45 n=45 n=45 n=45 n=45 n=45

OTHER DWS SJU CAR MAN HAT AGU MAY GUA CAY HUM

SJU MAN AGU GUA HUM
CAR HAT MAY CAY OTHER DWS

A **Abundance of pathogen containing genera**

B **Target copies normalized by 16S rRNA gene copies**

Log normalized rpoB coverage

Timepoint: 1, 2, 3, 4, 5, 6