

1 **Additional File 2: Supplemental Methods**

2

3 **Sample Collection**

4

5 *In Field Sample Collection*

6

7 Sampling to collect cells and particulate matter in source water was conducted with a tandem
8 pair of 12 cm x 3.5 cm-diameter point-of-use 0.1 micron hollow fiber membrane filters (Sawyer
9 Products, Inc.). The tandem unit was assembled by attaching two filters together with tubing; the
10 first filter in the tandem pair was labeled “Filter A” and the second filter in the tandem pair was
11 labeled “Filter B”. Field kits containing pre-labeled and pre-assembled tandem filter units,
12 detailed instructions, and accessories to perform systematic sampling were assembled at Hope
13 College, and delivered to the field collection sites in the Dominican Republic (Supplemental
14 Figure 1). Clean 18.9 L (5 gallon) plastic buckets were fitted with filters and tubing. Buckets
15 were rinsed with source water then filled with 16 L of source water which was allowed to gravity
16 drain through the filter. After a bucket was drained, the tandem filter unit was detached and 4, 50
17 mL volumes of air were pushed through the sample with a syringe to flush out residual water.
18 Each tandem filter unit was then capped at both ends, placed in a zip-sealed plastic bag and
19 shipped back to Hope College. Each field kit contained materials for the collection of three
20 biological replicates at each source water location. An image of a tandem filter setup in the field
21 is shown in Supplemental Figure 1.



22

23 Figure S1. Kit components and field setup of tandem filter sampling of drinking water sources.

24

25 **Sample Processing**

26

27 *Collection of Particulates and Bacteria from Filters*

28

29 An apparatus (Supplemental Figure 2) was designed and constructed to backflush filters returned
 30 from the field in order to collect the source water particulates and bacterial cells captured on the
 31 filter. The apparatus was sterilized using a 70% ethanol solution each day prior to use and at the
 32 filter attachment point prior to attaching each filter. The air intake was fitted with a HEPA filter
 33 to prevent airborne contamination.

34 A filter was attached to the apparatus in a reverse orientation to filtering in the field. A slug of
 35 autoclaved 18M Ω water was flushed through the filter with 103 kPa air, and the sample
 36 containing removed particulates and bacteria was collected into a sterile 250 mL centrifuge

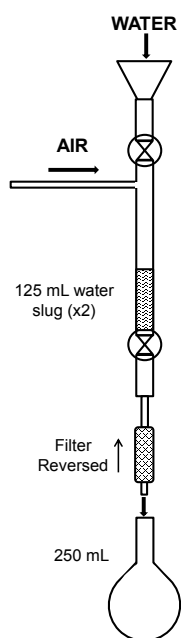
37 bottle. Numerous trials with known amounts of material loaded onto filters determined that using
38 two sequential slugs of 125 mL of water, at 103kPa air pressure, provided the highest systematic
39 yield of filtered material ($95 \pm 5\%$).

40 Backflushed samples were centrifuged at 10,000 x G for 20 minutes to pellet cells and
41 particulates. The supernatant was removed via vacuum aspiration and the pellet was resuspended
42 in 30 mL of 18 M Ω water. A 200 μ L aliquot of the homogenized pellet resuspension was
43 pipetted directly into Qiagen DNeasy PowerViral lysis tubes and stored at -80C until DNA
44 isolation was performed. The remaining pellet resuspension was used for particulate analysis.

45

46 For each day that backflushing was performed, controls were produced to assess the bacterial
47 community associated with 1) the water used for backflushing (MQ Control, 200 μ l directly into
48 a DNA extraction tube; Blank Pellet, 250 mL centrifuged as above) and 2) unused filter devices
49 (Control Pellet; 250 mL backflushed through a brand new filter).

50



51

52 Figure S2. Schematic illustrating the laboratory back flushing apparatus to retrieve the suspended load particulates
 53 and cells captured by the filter at the drinking water source.

54

55 ***Metal Retrieval***

56

57 Returned field foams were processed through an acid-wash procedure to recover adsorbed
 58 metals. Foam blocks were first dried at 100 °C for 24 hours then weighed. Blocks were cut into
 59 two pieces and placed in a desiccator. Prior to rinsing, one half-block piece was reweighed and
 60 placed in an acid-washed 50 mL syringe equipped with a 0.45 µm syringe tip filter. 30 mL of 3%
 61 trace-metal grade nitric acid (pH~0.7) was added to the open syringe holding the foam, and a
 62 plunger was inserted. After 5 minutes the plunger was depressed and the acid was filter-pressed
 63 through the foam block into a 50 mL metal-free conical tube. Rinsates were analyzed
 64 immediately or refrigerated until analysis. Because the foam blocks employed for sequestering

65 metals contained some of the target analyte metals as part of the chelating formulation, rinsates
66 from background control foams were analyzed as well as controls of field, storage, and
67 processing equipment.

68

69 *Amplicon Sequencing of 16S rRNA V4 Region*

70

71 DNA was extracted from backflushed material derived from filters using DNeasy PowerViral
72 kits (MoBio Laboratories, Inc., CA, a Qiagen company) per the manufacturer's protocol. DNA
73 samples were prepared for Illumina sequencing by following the MiSeq Wet Lab SOP [1].
74 Libraries for 94 DNA samples and two controls (a negative control and a positive, mock
75 community control (ZymoBIOMICS Microbial Community Standard, Zymo Research, Irvine,
76 CA) were processed in a 96 well plate; libraries for four, 96 well plates were combined for a
77 single sequencing run. High throughput library preparation was performed with an EpMotion
78 5075 automated liquid handling device (Eppendorf North America, Enfield, CT). Amplification
79 of the V4 region of the bacterial 16S rRNA gene was performed using universal primers (515F,
80 5'-GTGCCAGCMGCCGCGGTAA-3' and 907R 5'-CCGTCAATTCMTTTRAGTTT-3';
81 modified to include unique index pairs and adapters for Illumina sequencing) and Accuprime Pfx
82 Supermix under the following conditions: 95 °C for 2 min; 30 cycles of 9 °C for 20 s, 55 °C for
83 15 s and 72 °C for 5 min; final elongation at 72 °C for 10 min. Amplicon size was checked on a
84 1% agarose gel and confirmed with High Sensitivity D1000 ScreenTape on the Agilent 2200
85 TapeStation system (Agilent Technologies Inc., CA). Amplicons were then purified using
86 calibrated Ampure XP beads (Beckman Coulter; Brea, CA) and normalized with a SequelPrep
87 kit (Invitrogen Corporation, CA). Normalized samples from a single 96 well plate were pooled

88 into a single 1.5 mL tube. The pooled DNA was quantified using a Qubit dsDNA high sensitivity
89 assay kit (Thermo Fisher Scientific Inc.; Waltham, MA). Four, pooled PCR amplicon libraries
90 (equivalent of four, 96 well plates) were combined in equimolar concentrations into a single
91 library. This combined library was sequenced using an Illumina MiSeq (500 cycle, 2 x 250bp
92 paired-end v2 chemistry, Illumina, Inc., San Diego, CA) according to the manufacturer
93 instructions.

94

95 **Analysis**

96

97 *Particulates-(Suspended Load)*

98

99 The suspended load present in source waters was estimated by resuspending backflush samples
100 and immediately analyzing using a Microlab[®] FX522 spectrophotometry system. Attenuation
101 (transmittance and absorbance) and scattering were measured at multiple wavelengths (refer to
102 [2] for more detail). Particulate concentrations were estimated by comparison to standard curves
103 of known suspended load that were developed by using individual common rock-forming
104 minerals. These standard curves were considered representative of major types of geologic
105 terrains, based on the assumption that the suspended load in any location is systematically
106 reflective of the eroding substrate [3–6]. Many of these minerals were identified in the
107 particulate material recovered upon back-flushing. Mineralogy was determined and/or estimated
108 by powder X-ray diffraction (PXRD) techniques (Rigaku[®] MiniFlex+) and SEM-EDS
109 (Hitachi[®] TM-3000) analysis. SEM-EDS was also used to identify elemental content of
110 particulates.

111 *Metals in Rinsates-(Dissolved Metals)*

112

113 Rinsates were analyzed by ICP-OES techniques with a PerkinElmer® Avio 200 instrument.

114 Quality assurance (QA) and quality control (QC) checks were consistent with a modified EPA

115 Method 200.7 protocol, as summarized in methods provided by Perkin Elmer [7]. Analytes in the

116 study include As, Se, Zn, Pb, Cd, Ni, Fe, Mn, Cr, Mg, Cu, Ce, Sb, and Ba. Additional

117 information on detection wavelengths, ICP parameters and QA/QC for this study can be accessed

118 in Peterson et al, 2020. Raw data were processed through a statistical comparison routine and

119 reverse protocol algorithm to estimate metal concentrations in field drinking water sources.

120 Several working assumptions were made for results obtained from metal retrieval and analysis

121 procedures and element retention and recovery testing [2].

122

123 Single-element foam retention/recovery testing was performed to evaluate the relative validity of

124 the working assumptions employed to estimate the heavy metal content in field source waters.

125 Results from retention tests indicated a trend from higher recovery at low input solution

126 concentrations to lower recoveries at high input solution concentrations. They also showed

127 different recoveries for different elements. For example, within the 20-200 ppb range of input

128 concentrations, the average recovery from all foams for Cu was: Cu = 241% ($\pm 210\%$). Cu is a

129 high concentration background element within the polyurethane foams. These results

130 substantiate the understanding that, except for Cu, metal concentrations determined in this study

131 should be considered minimum levels. However, there are some cautionary considerations

132 regarding the foam retention data. These are: 1) The tests were for single-metal solutions, the

133 behavior of which may not reflect the processes active when multiple elements are present in

134 solution with subsequent competition for adsorption sites on the foam; and, 2)The aqueous
135 matrix of the test solutions (RO water) is not representative of the complex multi-constituent
136 matrix of natural field waters.

137
138 As another check on the working assumptions, field tests on local lake water (Holland, MI, USA)
139 were performed in which direct analysis of water was compared to analysis of rinsates from lake
140 samples subjected to the field foam collection and rinsing protocol. Rinsates of foam-sequestered
141 lake samples were significantly distinguishable from average background foam concentrations
142 for As, Cd, Ni, Cr, Mg, Cu, Ce, Ba, and Se. However, only the concentrations of Mg, Ce, and Cu
143 were significantly distinguishable when comparing direct lake samples with lake water rinsates
144 after average background foam concentrations were subtracted. Mg concentrations were the only
145 values above the LOQ. This experiment indicated that foam retention from actual field samples
146 could only be calculated from sources with high enough concentration levels to distinguish from
147 background, and further supports that field concentrations of most metals estimated in the current
148 study should be considered minimum values.

149

150 *16S rRNA Communities*

151 Sequencing data from Illumina MiSeq runs were demultiplexed using the MiSeq Reporter
152 Generate FASTQ workflow, producing forward and reverse read files in FASTQ format for each
153 filter in a kit used to sample a drinking water source. Reads files from multiple sequencing runs
154 were processed together according to the MiSeq SOP [8, 9] using mothur v.1.44.0 [10] via a
155 batch script of mothur commands to produce a set of high quality operational taxonomic units
156 (OTUs) at a similarity cutoff of 97%. Reads were aligned and taxonomy assigned using the Silva

157 Release 132 alignment and database [11]. Chimeras were removed from the OTUs using vsearch
158 v2.13.3 [12]. After processing, the data set included 7,312,060 sequences, representing 88,319
159 unique sequences (OTUs).

160 The data set was analyzed using the *R* packages, Phyloseq [13] and vegan[14]. A phyloseq data
161 object was generated from the mothur “shared” and “taxonomy” files and a metadata file
162 describing characteristics of each sample. Each sequencing reaction is represented by a row in
163 the metadata spreadsheet, with each row representing backflushed material from a single filter.
164 Beta diversity analyses were performed by calculating Bray-Curtis pairwise distances followed
165 by NMDS ordination (k ranging from 2 to 15, 200 iterations). The quality of sequencing
166 reactions from each filter in each kit was initially evaluated by looking at the degree of clustering
167 of filters from a kit in NMDS ordinations alongside control sequencing reactions associated with
168 the date a kit was processed (backflush controls) and the library preparation plate the filters were
169 processed on (plate controls). It was expected that the three Filter A replicates from a kit (water
170 source) should cluster in the ordination and be separated from Filter B replicates of the same
171 (and other) kit(s). The number of reads from a sequencing reaction were also used to evaluate the
172 quality of the sequencing reactions. All sequencing reactions with fewer than 5000 reads were
173 eliminated from further analysis. Five kits were eliminated from further analysis of bacterial
174 communities (kits 221, 234, 241, 245, 247) after quality control evaluations were completed.
175 PERMANOVA [15] and betadisper were performed on pairwise sample groupings based on
176 these metadata categories: tandem_filter, filter, kit, water_type, and all subcategories of
177 water_type.

178 In order to determine the presence/absence of 18 bacterial genera in backflush material from
179 each filter, relative abundances of each OTU in a sample were calculated and OTUs were filtered

180 to remove those below 0.1%. Each genus was considered present in a drinking water sample if
181 the following criteria were met: 1) present in at least one Filter A replicate, 2) absent in all
182 associated Filter B replicates, and 3) absent in all associated backflush controls and the negative
183 plate control. Results for each water source (kit) are shown in Additional File 1.

184 **References**

- 185 1. SchlossLab/MiSeq_WetLab_SOP. SchlossLab; 2020.
186 https://github.com/SchlossLab/MiSeq_WetLab_SOP. Accessed 21 Oct 2020.
- 187 2. Peterson JW, Fry B, Wade D, Fishman F, Stid J, Tarp C, et al. Application of point-of-use
188 water filters and polyurethane foams as a global reconnaissance method to characterize
189 particulates and heavy metals in untreated drinking water sources. In Review.
- 190 3. Meybeck M, Laroche L, Dürr HH, Syvitski JPM. Global variability of daily total suspended
191 solids and their fluxes in rivers. *Glob Planet Change*. 2003;39:65–93.
- 192 4. Garzanti E, Andó S, France-Lanord C, Censi P, Vignola P, Galy V, et al. Mineralogical and
193 chemical variability of fluvial sediments 2. Suspended-load silt (Ganga–Brahmaputra,
194 Bangladesh). *Earth Planet Sci Lett*. 2011;302:107–20.
- 195 5. Blake JM, Peters SC. The occurrence and dominant controls on arsenic in the Newark and
196 Gettysburg Basins. *Sci Total Environ*. 2015;505:1340–9.
- 197 6. Nasrabadi T, Ruegner H, Schwientek M, Bennett J, Valipour SF, Grathwohl P. Bulk metal
198 concentrations versus total suspended solids in rivers: Time-invariant & catchment-specific
199 relationships. *PLOS ONE*. 2018;13:e0191314.

- 200 7. Sarojam P. Application Note: ICP-Optical Emission Spectroscopy. 2010.
201 https://www.perkinelmer.com/lab-solutions/resources/docs/APP_MetalsinWastewater.pdf.
- 202 8. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a Dual-
203 Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the
204 MiSeq Illumina Sequencing Platform. *Appl Environ Microbiol.* 2013;79:5112–5120.
- 205 9. Patrick D. Schloss. MiSeq SOP. MiSeq SOP. https://mothur.org/wiki/miseq_sop/. Accessed
206 13 Oct 2020.
- 207 10. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing
208 mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing
209 and Comparing Microbial Communities. *Appl Environ Microbiol.* 2009;75:7537–41.
- 210 11. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal
211 RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.*
212 2013;41:D590–6.
- 213 12. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool
214 for metagenomics. *PeerJ.* 2016;4:e2584.
- 215 13. McMurdie PJ, Holmes S. Phyloseq: An R Package for Reproducible Interactive Analysis and
216 Graphics of Microbiome Census Data. *PLoS ONE.* 2013;8.
- 217 14. Jari Oksanen FGB, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, et al. vegan:
218 Community Ecology Package version 2.5-4. 2019. <https://CRAN.R-project.org/package=vegan>.

219 15. Anderson MJ. Permutational Multivariate Analysis of Variance (PERMANOVA). In: Wiley
220 StatsRef: Statistics Reference Online. American Cancer Society; 2017. p. 1–15.
221 doi:10.1002/9781118445112.stat07841.

222