



Metagenomic Sequencing and Quantitative Real-Time PCR for Fecal Pollution Assessment in an Urban Watershed

Kyle D. Brumfield^{1,2}, Joseph A. Cotruvo³, Orin C. Shanks⁴, Mano Sivaganesan⁴, Jessica Hey⁴, Nur A. Hasan², Anwar Huq¹, Rita R. Colwell^{1,2,5*} and Menu B. Leddy^{6*}

¹ Maryland Pathogen Research Institute, University of Maryland, College Park, MD, United States, ² University of Maryland Institute for Advanced Computer Studies, University of Maryland, College Park, MD, United States, ³ Joseph Cotruvo and Associates LLC, Washington, DC, United States, ⁴ U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH, United States, ⁵ CosmosID Inc., Rockville, MD, United States, ⁶ Essential Environmental and Engineering Systems, Huntington Beach, CA, United States

OPEN ACCESS

Edited by:

Edmond Sanganyado,
Shantou University, China

Reviewed by:

Peiyong Hong,
King Abdullah University of Science
and Technology, Saudi Arabia
Kishor Acharya,
Newcastle University, United Kingdom
Christopher Staley,
University of Minnesota Twin Cities,
United States

*Correspondence:

Rita R. Colwell
rcolwell@umiacs.umd.edu
Menu B. Leddy
menuleddy@gmail.com

Specialty section:

This article was submitted to
Environmental Water Quality,
a section of the journal
Frontiers in Water

Received: 06 November 2020

Accepted: 22 January 2021

Published: 15 February 2021

Citation:

Brumfield KD, Cotruvo JA,
Shanks OC, Sivaganesan M, Hey J,
Hasan NA, Huq A, Colwell RR and
Leddy MB (2021) Metagenomic
Sequencing and Quantitative
Real-Time PCR for Fecal Pollution
Assessment in an Urban Watershed.
Front. Water 3:626849.
doi: 10.3389/frwa.2021.626849

Microbial contamination of recreation waters is a major concern globally, with pollutants originating from many sources, including human and other animal wastes often introduced during storm events. Fecal contamination is traditionally monitored by employing culture methods targeting fecal indicator bacteria (FIB), namely *E. coli* and enterococci, which provides only limited information of a few microbial taxa and no information on their sources. Host-associated qPCR and metagenomic DNA sequencing are complementary methods for FIB monitoring that can provide enhanced understanding of microbial communities and sources of fecal pollution. Whole metagenome sequencing (WMS), quantitative real-time PCR (qPCR), and culture-based FIB tests were performed in an urban watershed before and after a rainfall event to determine the feasibility and application of employing a multi-assay approach for examining microbial content of ambient source waters. Cultivated *E. coli* and enterococci enumeration confirmed presence of fecal contamination in all samples exceeding local single sample recreational water quality thresholds (*E. coli*, 410 MPN/100 mL; enterococci, 107 MPN/100 mL) following a rainfall. Test results obtained with qPCR showed concentrations of *E. coli*, enterococci, and human-associated genetic markers increased after rainfall by 1.52-, 1.26-, and 1.11-fold log₁₀ copies per 100 mL, respectively. Taxonomic analysis of the surface water microbiome and detection of antibiotic resistance genes, general FIB, and human-associated microorganisms were also employed. Results showed that fecal contamination from multiple sources (human, avian, dog, and ruminant), as well as FIB, enteric microorganisms, and antibiotic resistance genes increased demonstrably after a storm event. In summary, the addition of qPCR and WMS to traditional surrogate techniques may provide enhanced characterization and improved understanding of microbial pollution sources in ambient waters.

Keywords: whole metagenome sequence, fecal indicator bacteria, microbial source tracking, quantitative real-time PCR, metagenomic analysis, culture, rainfall-runoff

INTRODUCTION

Microbiological degradation of surface water used for recreation, i.e., rivers, lakes, streams, and beaches, is a major water quality concern globally. Pathogenic microorganisms contaminating surface water can cause a range of food- and water-borne diseases. In the United States, maintenance and safety of water used for recreation are covered in the United States Environmental Protection Agency (USEPA) Clean Water Act of 1972, which regulates discharge of pollutants into waters and sets water quality standards for surface water (EPA, 2002). However, many U.S. waters fail to meet designated microbiological thresholds of water used for drinking and recreation due to high levels of fecal contamination.

Sources of surface water fecal contamination typically include wastewater treatment plants, septic systems, domestic and wild animal feces, and municipal sewer overflows after heavy rainfall events. Stormwater runoff can lead to surface water contamination by accumulation of microbiological and chemical pollutants on land during dry weather periods and subsequent transport into nearby waterways. Notable examples of waterborne outbreaks include *Escherichia coli* O157:H7 and *Campylobacter jejuni* in Walkerton, Ontario (Auld et al., 2004), *Salmonella* spp. in Georgia, USA (Lee D. et al., 2019), and *Cryptosporidium* spp. in Milwaukee, USA (Curriero et al., 2001). Thus, characterizing the collective microbial composition of water and identifying potential contaminating sources are priorities for local management groups.

Fecal microbial contamination levels are routinely determined by culturing indicator microorganisms, namely enterococci and *E. coli*, common in fecal waste of humans and other warm-blooded animals. A correlation has been established between many general fecal indicator bacteria (FIB) and gastrointestinal illnesses (Wade et al., 2010). However, culture-based monitoring has limitations, particularly an inability to detect presence of all potential pathogens in surface waters since presence of some enteric microorganisms is poorly correlated with FIB (Harwood et al., 2005; Pusch et al., 2005). FIB are not only shed by humans and other warm-blooded animals, but also cold-blooded animals such as amphibians (Gibb et al., 2017), making it difficult to identify the responsible animal pollution source(s) for targeted remediation. Furthermore, *E. coli* and enterococci of non-fecal origin capable of proliferating in the natural environment are now widely recognized and may confound FIB water quality monitoring in some areas (Byappanahalli et al., 2012). Thus, there is a benefit of combining water quality molecular methods that can discern between contaminating sources and track potentially public health relevant targets with traditional methods for FIB monitoring; for review, see (Meays et al., 2004; Hamilton et al., 2020; Mathai et al., 2020).

Molecular methods that target specific genetic regions of a microbial genome have long been suggested to complement culture methods for monitoring FIB. As a result, molecular tools, namely quantitative real-time PCR (qPCR) and digital droplet PCR, that can measure fecal contamination levels rapidly, identify sources of contamination, and track public health relevant targets, such as antimicrobial resistance (AR) genes

and specific enteric microorganisms, have been developed (EPA, 2004; Yang et al., 2017; Staley et al., 2018). These methods are useful for profiling factors contributing to non-point source contamination, e.g., humans (Shanks et al., 2009), cows (Shanks et al., 2008), dogs (Rojas et al., 2017), pigs (Mieszkin et al., 2009), horses (Dick et al., 2005), waterfowl (Ohad et al., 2016), gull (Lee et al., 2013), geese, ducks, and chicken (Green et al., 2012). However, detection and enumeration of one or more of these genetic markers by themselves does not provide a sufficiently comprehensive set of information that is needed to identify the full range of microorganisms and AR genes that may be present in surface waters contaminated with fecal waste.

More recently, metagenomic sequencing, coupled with bioinformatics, has gained attention as an effective water quality assessment tool (Wang et al., 2016; Burcham et al., 2019; Acharya et al., 2020; Brumfield et al., 2020a). However, many waterborne microbial surveys have relied upon targeted sequencing (Uyaguari-Diaz et al., 2016; Su et al., 2017; Jin et al., 2018), which generally employ universal PCR primers to amplify hypervariable regions of the 16S rRNA gene to infer taxonomic identification of bacteria and archaea by mapping sequencing reads to genomic databases. By employing whole metagenome sequencing (WMS), the bacterial, archaeal, viral, fungal, and protozoan microbiome community members can be profiled, and in some instances, identified to sub-species taxonomic level (Brumfield et al., 2020b). Profiling the complete microbiome by culture-independent technologies provides an assessment of overall microbial community diversity, which can be used by engineers for the development and optimization of biological systems pertaining to functional processes and nutrient cycling, such as nitrogen and phosphorus removal bacteria; for review, see (Ferrera and Sánchez, 2016). A few studies have used WMS to analyze surface water quality (Shanks et al., 2013; Brown et al., 2015; Fisher et al., 2015; Wu et al., 2018; Hamner et al., 2019; Lee et al., 2020), and microbial community composition following rainfall events remains understudied.

The objective of this pilot study was to demonstrate the feasibility and application of using WMS, host-associated and FIB qPCR, and FIB culture to characterize fecal pollution trends in ambient waters through analysis of samples collected in an urban watershed before and after a rainfall event. Relative abundance (RA) of bacteria, archaea, fungi, protists, and viruses, and carriage of AR associated genes was determined to assess microbiological pollution and provide a comparison with results of traditional water quality methods (FIB culture and host-associated targets). Results showed that added benefit can be achieved by incorporating WMS as a complement to culture and qPCR for water quality monitoring.

MATERIALS AND METHODS

Site Description and Sample Collection

A total of eight water samples were collected from sites located along a creek in an urban watershed (**Figure 1**). Sampling sites were consecutively numbered, starting upstream (site 1) and moving downstream (site 4). Grab samples were collected at each site during a dry weather period (no precipitation during



FIGURE 1 | Map of sampling locations. Map shows sites of surface water collection during dry weather, September 23, 2019, and post rainfall, October 24, 2019. Water samples were collected at all four sites during each sampling event. Blue, waterway; green, land. Scale bars are shown to indicate distance.

the previous week) on September 24, 2019 and ~12 h after the completion of a rain event (42.6 mm total precipitation during the previous 72 h) on October 23, 2019. The sliding 31-day precipitation averages for the study area on September 30 and October 31, 2019 were reported as 0.75 and 4.99 mm, respectively, with the greatest precipitation event (34 mm) occurring on October 16, 2019. Time series of the area-averaged daily precipitation rate for September and October 2019 were reported by Goddard Earth Sciences Data and Information Services Center Interactive Online Visualization and Analysis Infrastructure (Acker and Leptoukh, 2007). Surface water (2 L) was collected at each location using a sterile Nalgene carboy (Thermo Fisher Scientific, Waltham, MA, USA) treated previously with hydrochloric acid (10% v/v), ethanol (95% v/v), and autoclaved. Water samples were transported to the laboratory on ice. Temperature of the water samples was monitored to ensure that it did not reach above 8 °C during transportation by using a LogTag® single trip temperature alert indicator (LogTag Recorders, Auckland, New Zealand). Samples were processed as described below within 2 h of collection and following recommendations of EPA Clean Water Act Analytical Methods (EPA, 2019).

***Escherichia coli* and Enterococci Surrogate Testing**

Escherichia coli and enterococci concentrations in the grab samples were enumerated, following manufacturer's instructions for most probable number (MPN) method per 100 mL of sample water using IDEXX Quanti-Tray System with commercial Colilert and Enterolert media (IDEXX, Westbrook, ME, USA), respectively, providing a culturable bacteria detection range between two and 2,491.6 MPN per 100 mL of water (additional details can be found in the **Supplementary Material**). During each sampling run, 100 mL of nuclease free water was prepared at site 4 with the Colilert and Enterolert media, respectively, and transported to the lab for processing as mentioned to serve as a trip sterility blank. Cultures of *Enterococcus faecalis* (ATCC® 29212™) and *E. coli* (ATCC® 29212™) obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA) were prepared under standard growth conditions in Luria-Bertani broth at 37 °C overnight (16 h) with aeration, and 100 µL was added to 99.9 mL of nuclease free water containing IDEXX media to serve as positive controls for the Enterolert and Colilert assays, respectively.

Quantitative Real-Time PCR for Host-Associated and General FIB Genetic Markers

Sample Filtration and DNA Purification

At each sampling date and site, 100 mL (dry weather event) and 20 mL (post rainfall event) were filtered in triplicate using 0.45 µm polycarbonate filters (Fisher Scientific, Pittsburg, PA) to capture larger microorganisms. Because dry weather and post rainfall sampling events contained a variable turbidity content, the volume of water filtered was dependent on the filter and when it clogged. Filtrates in sterile 2 mL screw cap tubes

containing silica bead mill matrix (GeneRite, North Brunswick, NJ) were shipped on dry ice overnight to the USEPA research laboratory (Cincinnati, OH, USA), and stored at –80 °C until DNA purification (< 30 days). DNA purification of filtrates from 24 filters (4 sites × 2 sampling events × 3 replicates/sample) was done using the DNA-EZ RW02 kit (GeneRite LLC, North Brunswick, NJ, USA), as previously described (Li et al., 2019). Three method extraction blanks (MEB) served as controls. DNA extracts were stored in GeneMate Slick low-adhesion microcentrifuge tubes (ISC BioExpress, Kaysville, UT, USA) at 4 °C prior to qPCR amplification (< 48 h).

Reference DNA Materials

Reference DNA consisted of two plasmid constructs (Integrated DNA Technologies, Coralville, IA, USA) and salmon sperm DNA (Sigma-Aldrich, St. Louis, MO, USA). Plasmid constructs for internal amplification controls (IAC) and calibration standards (all DNA targets in a single construct) were prepared, as previously described (Li et al., 2019). The Reference DNA was stored in GeneMate Slick low-adhesion microcentrifuge tubes (ISC BioExpress, Kaysville, UT, USA) at –20 °C.

qPCR Amplification

Four host-associated, i.e., human-associated (HF183/BacR287) (USEPA, 2019), ruminant-associated (Rum2Bac) (Mieszkien et al., 2010), canine-associated (DG3) (Green et al., 2014), and avian-associated (GFD) (Green et al., 2012), and two general, i.e., enterococci (Enterol1a) (Ludwig and Schleifer, 2000; Siefiring et al., 2008) and *E. coli* (EC23S857) (Chern et al., 2011), FIB qPCR assays were used in this study along with a sample processing control (SPC) assay (Sketa22) (Shanks et al., 2016), as reported previously. Oligonucleotide sequences used in this study can be found in the supporting information (**Supplementary Table 1**). All reactions contained either 2 µL of DNA extract or between 10 and 1.0×10^5 target gene copies of reference DNA calibration standards. HF183/BacR287 multiplex reactions also contained 100 copies of IAC template. All reference DNA and water samples were analyzed in triplicate. The log fluorescence threshold was manually set to either 0.03 (HF183/BacR287, DG3, Rum2Bac, Enterol1a, EC23S857, and Sketa22) or 0.08 (GFD). Quantification cycle (Cq) values were exported to Excel (Microsoft, Redmond, WA, USA) for further analysis.

Quality Controls

To monitor for potential extraneous DNA contamination during qPCR amplification, six no-template controls (NTC) with purified water substituted for template DNA were performed with each instrument run. SPC protocol was used to identify suitable and consistent DNA recovery from each water sample, as previously described (Shanks et al., 2016). HF183/BacR287 multiplex IAC procedure was used to monitor for amplification inhibition. For each GFD instrument run, a melt curve analysis with a resolution of 0.3 °C was used after thermal cycling to identify spurious amplicons that could confound data interpretation (no spurious amplicons detected; data not shown).

qPCR Data Analysis

DNA calibration models were generated for each qPCR assay instrument run using the “single” Bayesian Markov Chain Monte Carlo approach (Sivaganesan et al., 2010). Amplification efficiency (E) for each reference DNA calibration model was calculated as follows: $E = 10^{(-\frac{1}{\text{slope}})} - 1$. The lower limit of quantification (LLOQ) was defined as the 95% credible interval upper-bound from repeated measurement ($n = 3$) of 10 copies per reaction reference DNA standard dilution. To investigate the influence of rainfall, water samples were organized into dry weather and post rainfall sample groups (4 sites \times 2 sampling events \times 3 filters/sampling event \times 3 replicates/filter = 36 reactions per sample group). A fecal score ratio can be used to estimate the relative level of host-associated or general fecal contamination present between dry and post-rainfall sampling events based on the weighted average source-specific gene concentration observed in each group. Weighted average fecal score ratios (average \log_{10} copies \pm 95% Bayesian credible interval) were estimated for each qPCR assay based on rainfall (post rainfall or dry weather) data group definition utilizing all measurements including non-detects (ND), detections below the LLOQ (BD), and measurements within the range of quantification (ROQ), as reported elsewhere (Cao et al., 2018; Shrestha et al., 2020). A sample group was eligible for fecal score ratio determination if each sample group (post rainfall/dry weather) had at least one BD measurement. To account for different sample volumes between post rainfall (20 mL) and dry weather (100 mL) samples, fecal scores were adjusted to indicate a 20 mL test volume (subtraction of $\log_{10}(5)$ from dry weather fecal scores prior to calculating ratios). All statistical data analyses were conducted using WinBUGS v.1.4.3 (University of Cambridge, 2020), Statistical Analysis Software (SAS Institute, Cary, NC, USA), and Excel (Microsoft, Redmond, WA, USA).

Whole Metagenome DNA Sequencing

A total of 600 mL of water from each sampling event was concentrated by using a combination of vacuum and syringe filtration. Water samples were first passed through six 0.6 μm pore size 25 mm polycarbonate Whatman Nuclepore Track-Etch Membranes (Millipore Sigma, St. Louis, MO, USA) by vacuum filtration, to remove trace minerals and other particulates and expedite subsequent filtration. Resulting filtrate was collected aseptically and passed consecutively through a single 0.22 μm pore size Sterivex™ Filter Unit (Millipore Sigma, St. Louis, MO, USA) by using syringe filtration and six 0.1 μm pore size 47 mm polycarbonate Whatman Nuclepore Track-Etch Membranes (Millipore Sigma, St. Louis, MO, USA) by vacuum filtration to capture smaller microorganisms, including some viruses and phages. All filter membranes were stored at -80°C until DNA was prepared (< 30 days).

For each sample, total DNA was isolated from the microbial biomass collected on the Sterivex™ Filter Unit using the Qiagen DNeasy Power Water Sterivex Kit (Qiagen, Germantown, MD, USA), following manufacturer's instructions, to obtain a final elution volume of 60 μL . DNA was prepared from all 12 filter membranes using the Qiagen DNeasy Power Soil Kit (Qiagen,

Germantown, MD, USA), with the following modifications for DNA extraction from the filter membranes. The 12 filter membranes for each sample were cut into ribbons ~ 2 mm by 10 mm and distributed evenly amongst six Power Bead Tubes included in the Qiagen Power Soil Kit (Qiagen, Germantown, MD, USA), and final elution volume for each of the six membrane preparations was 20 μL . Eluted DNA was pooled for each Sterivex™ Filter Unit and the six membrane preparations to achieve 180 μL . DNA was purified using DNA Clean and Concentrator™-25 Kit (Zymo Research, Irvine, CA, USA), following manufacturer's instructions, providing a final elution volume of 50 μL .

Concentrations of genomic dsDNA were measured using Qubit® dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) on an Invitrogen Qubit® 4.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). DNA libraries were prepared using Nextera XT DNA Library Prep Kit (Illumina Inc., San Diego, CA, USA). All DNA libraries were quantified, as previously mentioned, and sequenced using an Illumina HiSeq4000 Instrument (Illumina Inc., San Diego, CA, USA) with a 2×150 bp run. A negative sequencing control, consisting of nuclease-free water, and a sequencing standard, i.e., ZymoBIOMICS™ Microbial Community Standard (Zymo Research, Irvine, CA, USA), were included for quality assurance of high-throughput sequencing. Metagenomic samples were sequenced with an average of 50.9 (min = 44.2; max = 56.8) million sequence read depth across samples.

Whole Metagenome Sequence Analysis

General sequencing statistics for all samples and mean sequence quality distribution, measured by FastQC (v.0.11.6) (Andrews, 2019), are detailed in **Supplementary Table 2**. Base-calling error probabilities (P) were evaluated using Phred Quality Score (Q), defined by: $Q = -10 \log_{10}(P)$. Using a previously defined read quality threshold (Roy et al., 2018; Brumfield et al., 2020a), read libraries were above a Phred Quality Score of 17 for at least 80% of the read lengths, i.e., probability of correct base call was at least 98%, so the reads were not subjected to quality trimming. The average Illumina sequencing read lengths across all libraries was 151 bp.

Unassembled metagenomic sequencing reads were analyzed, as described previously (Lax et al., 2012; Ponnusamy et al., 2016; Roy et al., 2018; Connelly et al., 2019; Brumfield et al., 2020a), using CosmosID Metagenomics Cloud Application v.1.0 (Cosmos ID, 2019) to achieve multi-kingdom microbiome analysis and profiling of AR associated genes and quantification of the organism RA, defined as the proportion of unique organism-specific k-mers annotated by each database relative to the total number of unique sequencing reads generated for that sample. Additional information on the bioinformatics pipeline employed for taxonomic classification of sequencing reads can be found in the supporting information.

Principal coordinate analysis (PCoA) employing Bray-Curtis distance measure and alpha diversity via CHOA1 index were performed based on RA of bacterial taxa in each sample. Analysis of community resistome was achieved by identifying AR associated genes based on percent coverage as a function

of gene-specific k-mer frequency in each sample. Sunburst visualizations of taxonomic composition for each sampling event, were generated using Krona (Ondov et al., 2011).

Data Availability

Illumina paired metagenomic sequencing data generated for all samples in this study are deposited in the NCBI Sequence Read Archive database (<https://www.ncbi.nlm.nih.gov/sra>) under BioProject PRJNA655751. Accession numbers for individual sample sequencing read libraries are provided in the supplementary information.

RESULTS

Enumeration of *Escherichia coli* and Enterococci

Concentrations of cultural FIB were determined using the IDXX/Quanti-Tray FIB assay, and results for each sample are shown in **Figure 2A**. Within dry weather sampling events, the concentrations of *E. coli* ranged between 117.8 and 248.1 MPN/100 mL, and enterococci ranged from 36.8 to 313 MPN/100 mL. During the post rainfall sampling event, variation in the concentration of *E. coli* and enterococci could only be assessed qualitatively because dilutions tested yielded results exceeding the upper limit of quantification; that is, concentrations of *E. coli* (sites 1 and 2) and enterococci (sites 1–3) in these samples were >2,491.6 MPN/100 mL. Thus, post rainfall, all samples exceeded the local single sample recreational water quality thresholds of 410 MPN/100 mL and 107 MPN/100 mL for *E. coli* and enterococci, respectively.

Host-Associated and General FIB qPCR qPCR Quality Controls

Calibration model performance parameters (slope and y-intercept parameters, LLOQ, linearity (R^2), and E) are provided in **Supplementary Table 3**. Calibration model R^2 -values were ≥ 0.997 and E -values ranged from 0.93 (Rum2Bac) to 1.01 (DG3 and GFD). Extraneous DNA control reactions indicated 94.4% were DNA-free. False positives Cq values ($n = 10$) were higher than respective LLOQ in all cases (lowest false positive = 38.3 Cq). Amplification inhibition was absent in all samples based on multiplex IAC HF183/BacR287 testing. IAC acceptance thresholds and competition thresholds ranged from 37.2 to 37.8 Cq and 27.8 to 31.2 Cq, respectively. All 24 water filter DNA extracts passed SPC testing, exhibiting negligible matrix interference. The SPC acceptance threshold was 33.0 Cq. IAC proficiency testing (Run 1 = 0.62 Cq and Run 2 = 0.45 Cq) and SPC proficiency screening (Batch = 0.57 Cq) indicated acceptable consistent implementation of these control experiments.

Host-Associated and General FIB qPCR Results

Water quality genetic markers for host-associated and general FIB were measured from 24 filters representing temporal sampling before and after a rain event. **Table 1** provides the number of qPCR measurements for each assay organized into ND, BD, and ROQ categories. It is important to note the difference in water sample volumes between dry weather

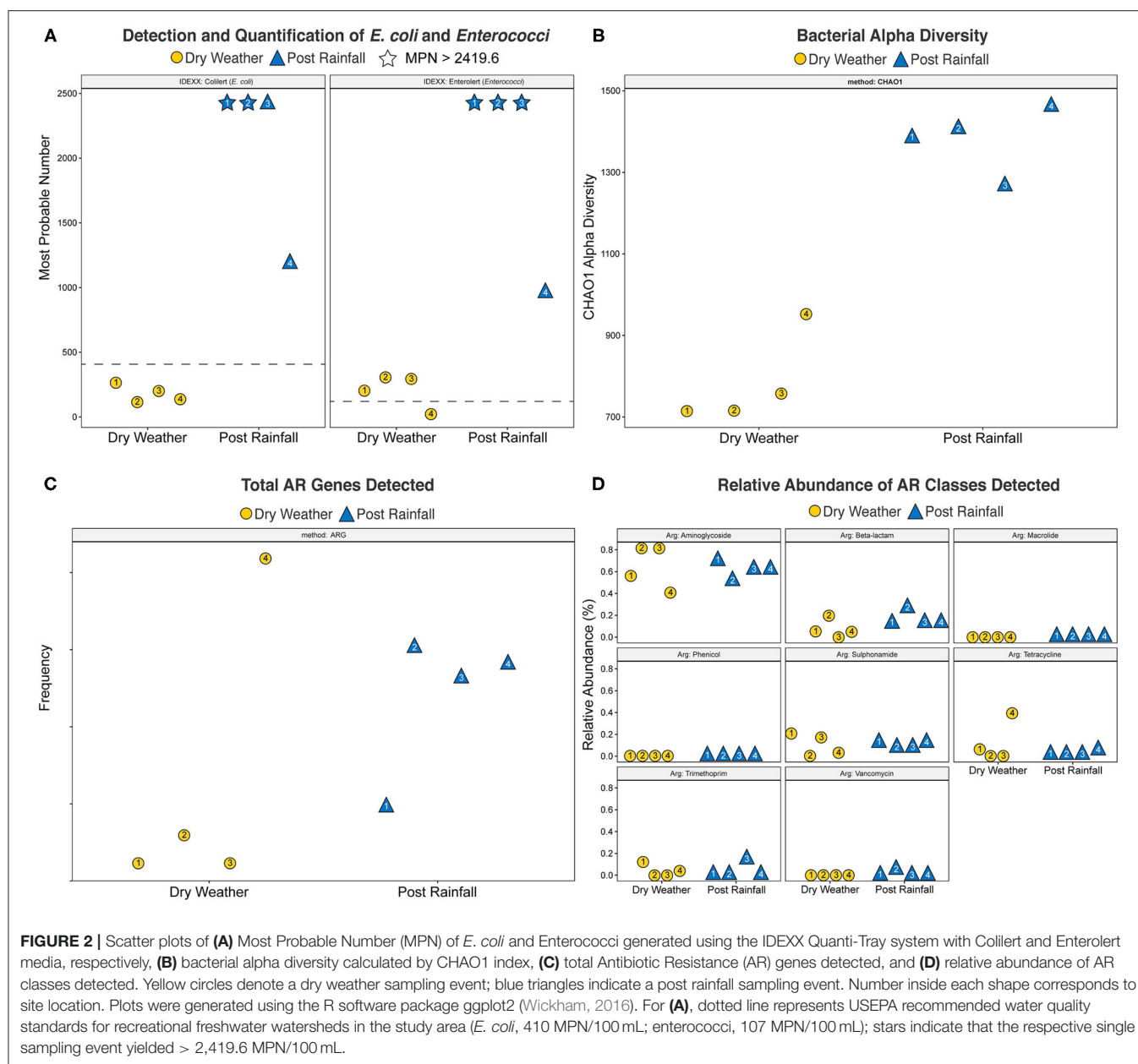
(100 mL) and post rainfall (20 mL) samples. Of the total 432 measurements, 49.8% were ND, 24.8% BD, and 25.5% ROQ. General FIB (Enterol1a and EC23S857) accounted for 91.8% of all ROQ ($n = 101$) followed by HF183/BacR287 ($n = 9$). The frequency of ROQ was higher post rainfall (29.2%; $n = 63$) compared to dry weather (21.8%; $n = 47$), despite the five-fold difference in sample volume. Enterococci and *E. coli* as well as avian and human host-associated genetic markers were detected in post rainfall and dry weather samples. Ruminant-associated genetic markers were detected in dry weather samples, while dog fecal waste was detected in post rainfall samples. DG3 (dog) was not detected in dry weather large volume samples (100 mL) but were detected in 41.7% of samples post rainfall (20 mL). In addition, Rum2Bac (ruminant) was not detected in dry weather samples; however, Rum2Bac was detected in post rainfall samples.

Fecal Score Ratio

Due to the large number of ND and BD results (74.5% of total), a censored data fecal score ratio (Cao et al., 2018; Shrestha et al., 2020) was used to calculate weighted average \log_{10} copies per reaction 95% Bayesian Credible Interval (BCI) ratio (post rainfall/dry weather) for each eligible qPCR assay data set, which included Enterol1a, EC23S857, HF183/BacR287, and GFD qPCR assays (**Table 1**). \log_{10} fecal score ratios for Enterol1a [−1.26 (−1.07 to −1.46 95% BCI)], EC23S857 [−1.52 (−1.22 to −1.82 95% BCI)], and HF183/BacR287 [−1.11 (−0.81 to −1.45 95% BCI)] indicate significantly higher concentrations of fecal sources for post rainfall when compared to dry weather sample groups (**Figure 3**). However, the avian marker, GFD, exhibited a different trend, with significant difference between post rainfall and dry weather sample fecal scores (95% BCI intersects at 0; −0.73 to 0.01).

Metagenomic Data Analysis Community Microbiome

WMS, using DNA prepared from the water samples, generated ~815 million reads across the raw sequence libraries equating to roughly 620 million unique reads (supplementary information). Total bacterial alpha diversity was calculated using CHAO1 index (**Figure 2B**) and ranged from 712 to 952 in dry weather samples and from 1,273 to 1,468 in samples post rainfall. Bacterial communities in water samples collected during dry weather and post rainfall were analyzed by three-dimensional PCoA using Bray-Curtis dissimilarity index (**Supplementary Figure 1**), where distance between points indicates degree of difference in bacterial DNA sequence composition. Each sample contained a relatively distinct bacterial composition; however, like samples clustered more closely with like samples, e.g., post rainfall samples clustered more closely compared to those collected during dry weather. Bacteria, archaea, fungi, protozoa, and viruses (including bacteriophages) identified by WMS characterization are shown in Krona plots, representing RA of microbial species gamma-diversity, i.e., total microbial species diversity, detected in dry weather (**Figure 4**) and post rainfall (**Figure 5**).



Community Resistome

The total number of AR and classes of AR associated genes are shown in **Figures 2C,D**, respectively. With the exception of site 4, the number of AR genes detected across each of the locations was lower in samples collected during dry weather compared to those collected after rainfall. Furthermore, site 4 was dominated by tetracycline resistance genes that were not detected in the dry weather samples obtained from the other locations but were detected in all samples post rainfall. Overall, the relative abundance of various antibiotic classes detected by WMS did not vary demonstrably across sampling events, and AR genes associated with the aminoglycoside class of antibiotics were dominant in all samples.

Wastewater-Associated Enteric Microorganisms

WMS allowed detection of multiple wastewater-associated enteric microorganisms (**Table 2**). RA of *Aeromonas hydrophila* was roughly one log greater at all sites following rainfall. *E. coli* was detected at all sites at roughly equal RA; however, at site 1 during dry weather sampling, *E. coli* comprised 0.36% RA. *Vibrio cholerae* was detected at all locations, except site 2. *Enterococcus* spp. (*E. casseliflavus*, *E. faecalis*, and *E. faecium*), were detected at sites 1 and 3 during dry weather sampling, while enterococci were detected at all sites except site 3 following rainfall. *Legionella pneumonophila* was detected at site 4 only following rainfall at extremely low RA. Fast growing *Mycobacteria* (*M. chelonae*, *M. abscessus*, and *M. phlei*) were detected at all locations following

TABLE 1 | Enumeration using qPCR.

Assay	Determinant	Weather	ND ^a	BD ^b	ROQ ^c	Fecal score ratio eligible
EC23S857	<i>E. coli</i>	Dry weather	6	18	12	Yes
		Post rainfall ^d	0	6	30	
Entero1a	Enterococci	Dry weather	1	9	26	Yes
		Post rainfall ^d	0	3	33	
GFD	Avian	Dry weather	22	14	0	Yes
		Post rainfall ^d	30	6	0	
HF183/BacR287	Human	Dry weather	22	5	9	Yes
		Post rainfall ^d	7	29	0	
Rum2Bac	Ruminant	Dry weather	34	2	0	No
		Post rainfall ^d	36	0	0	
DG3	Dog	Dry weather	36	0	0	No
		Post rainfall ^d	21	15	0	

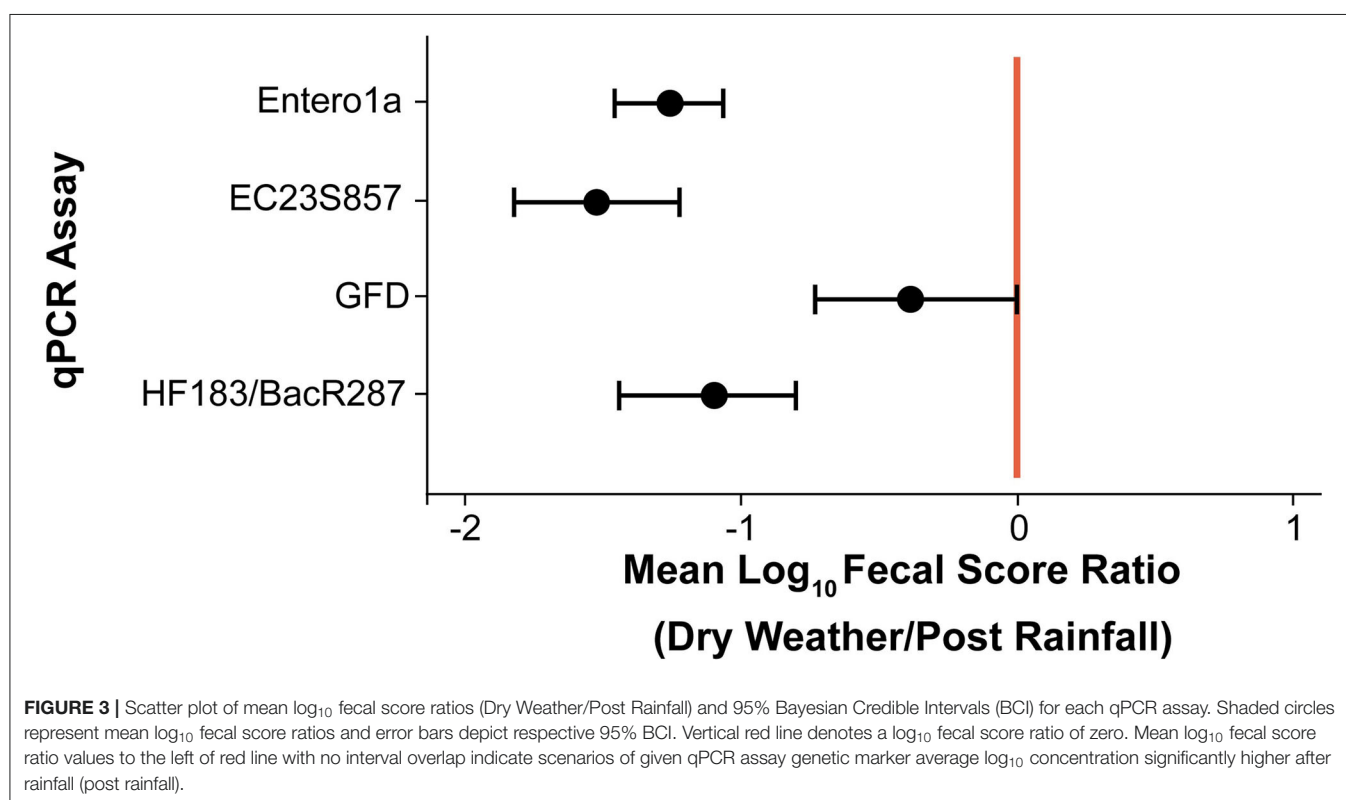
qPCR counts when no-detection (ND) was obtained, detections below LLOQ (BD), and within the range of quantification (ROQ) organized by sampling event type (post rainfall/dry weather) are given.

^ano-detection.

^bbelow the respective lower limit of quantification.

^cmeasurement within the respective range of qualification.

^dsample volumes of 20 mL per filter were used.



rainfall and at site 4 in the dry weather samples. MAC (*M. avium* and *M. intracellulare*) was detected at 0.03% RA at site 1 and at site 2 (and at lower RA) following rainfall but not detected in any other samples. *M. tuberculosis* was detected at 0.02% RA (low abundance) at sites 1 and 3 following rainfall and site 2 before rainfall. Following rainfall, *Yersinia enterocolitica* was detected at low RA at all sites. *Campylobacter coli* was detected only at site 4 in the dry weather samples. *Burkholderia*

pseudomallei was detected at all sites following rainfall and detected at sites 3 and 4 in the dry weather samples. Similarly, *Salmonella enterica* was detected at all sites except site 3 following a rainfall and detected at sites 3 and 4 in the dry weather samples. Human *mastadenovirus* C was detected only at site 4 after rainfall. *Cryptosporidium muris* was detected at all post rainfall sites except site 3 and detected at sites 1 and 2 in the dry weather samples. *Acanthamoeba spp.* were detected at

at 0.4% RA at site 4. In contrast, following rainfall, both *Bacteroides spp.* and *Bifidobacterium spp.* were detected at all sites. *Clostridium spp.* were detected at low RA at site 4 after rainfall and at all locations before rainfall. *Citrobacter spp.* were detected at roughly equal RA in all samples except site 3 in dry weather samples. *Escherichia spp.* were most abundant at site 1 in dry weather samples (0.36%) and detected at roughly equal RA at all other locations during dry weather and post rainfall sampling. *Enterobacter spp.* were detected in all samples (except site 2

TABLE 2 | Relative sequencing read abundance of wastewater enteric microorganisms detected in urban watershed microbiomes.

Genus	Species	Illness	Dry weather				After rainfall			
			Site 1	Site 2	Site 3	Site 4	Site 1	Site 2	Site 3	Site 4
Bacteria										
<i>Aeromonas</i>	<i>A. hydrophilia</i>	Gastroenteritis; Septicemia	<0.01	0.01	0.01	0.01	0.07	0.10	0.10	0.08
<i>Burkholderia</i>	<i>B. pseudomallei</i>	Melioidosis	-	-	0.03	0.01	<0.01	<0.01	<0.01	<0.01
<i>Campylobacter</i>	<i>Campylobacter spp.</i>	Gastroenteritis	0.03	0.02	0.02	<0.01	0.02	-	0.03	<0.01
<i>Yersinia</i>	<i>Y. enterocolitica</i>	Diarrhea, reactive arthritis	-	-	-	-	<0.01	<0.01	<0.01	0.04
<i>Enterococcus</i>	<i>E. casseliflavus</i> ; <i>E. faecalis</i> ; <i>E. faecium</i>	Gastroenteritis	<0.01	-	<0.01	-	<0.01	<0.01	-	0.01
<i>Escherichia</i>	<i>E. coli</i>	Gastroenteritis	0.36	0.02	0.04	0.02	0.03	0.04	0.02	0.03
<i>Legionella</i>	<i>L. pneumophila</i>	Respiratory illness (pneumonia, Pontiac fever)	-	-	-	-	-	-	-	<0.01
<i>Mycobacteria</i>	<i>M. tuberculosis</i>	Tuberculosis	<0.01	<0.01	-	-	<0.01	-	<0.01	0.02
	<i>M. avium</i> ; <i>M. intracellulare</i>	MAC	-	-	-	-	0.03	<0.01	-	-
	<i>M. chelonae</i> ; <i>M. abscessus</i> ; <i>M. phlei</i>	Extrapulmonary infection	-	-	-	0.01	<0.01	0.02	0.01	0.02
<i>Salmonella</i>	<i>S. enterica</i>	Gastroenteritis; reactive arthritis	-	-	<0.01	<0.01	<0.01	<0.01	-	<0.01
<i>Vibrio</i>	<i>V. cholerae</i> (non-O1/O139)	Cholera	0.03	<0.01	-	<0.01	<0.01	<0.01	<0.01	<0.01
Protozoa										
<i>Cryptosporidium</i>	<i>C. muris</i>	Gastroenteritis	1.35	0.57	-	-	0.10	0.10	-	0.23
<i>Acanthamoeba</i>	<i>Acanthamoeba spp.</i>	Amoebic meningoencephalitis, keratitis, encephalitis	23.66	38.25	1.26	13.89	8.24	14.42	1.94	6.68
Viruses										
<i>Adenoviridae</i>	Human mastadenovirus C	Gastroenteritis, respiratory illness, eye infection	-	-	-	-	-	-	-	<0.01

Relative abundance (percentage) of unique sequencing reads for each sample, annotated as respective microbial taxa, is shown. Sample RA values are shown as proportion within each kingdom, i.e., bacteria, protozoa, and viruses. Characterization of wastewater enteric microorganisms is defined elsewhere (Poffé and de Beeck, 1991; Shannon et al., 2007; Ramírez-Castillo et al., 2015; WHO, 2017; Richardson and Rautema-Richardson, 2019). “-”, not detected; MAC, *Mycobacterium avium* complex.

times higher, respectively. These findings are in agreement with previous studies where precipitation was found to contribute to high concentrations of FIB in surface water (Lee et al., 2020). High FIB concentrations are commonly reported in many urban watersheds throughout the U.S. For example, mean *E. coli* concentration of 1,156 MPN/100 mL was reported for a tributary of the Des Moines River in Iowa, USA (Schilling et al., 2009).

Generally, when FIB were detected in post rainfall samples by culture and qPCR, WMS supported the increased RA of FIB in samples collected after rainfall (Table 3). Unlike culture and qPCR methods, taxonomic profiling of metagenomic sequencing reads associated with *E. coli* and enterococci indicated comparatively low numbers. The results suggest that WMS may not be suitable for FIB monitoring alone since results can be susceptible to shifts in occurrence associated with changes in RA of other community members. As a result, WMS is best applied as a complement to culture and qPCR by providing useful information toward understanding microbial diversity and occurrence of AR genes, presence of enteric microorganisms, as well as additional information on sources of contamination.

Host-Associated Fecal Pollution

Identifying sources of fecal pollution in surface waters is a public health challenge, and qPCR is currently employed for surface water quality testing, namely to quantify specific genes in environmental samples (Dick et al., 2005; Shanks et al., 2008, 2009; Mieszkin et al., 2009; Green et al., 2012; Lee et al., 2013; Ohad et al., 2016; Rojas et al., 2017). In the study reported here, the weighted average fecal score ratio (\log_{10} copies per 100 mL) of the HF183/BacR287 genetic marker was -1.11 or ~ 12.8 -fold greater after rainfall (Table 1, Figure 3). This indicates sewage and/or other sources of human waste, such as storm drain overflow that can contribute to water quality degradation in the urban watershed. In contrast, the avian-associated GFD genetic marker showed no significant difference between post rainfall and dry weather samples suggesting birds do impact water quality but perhaps independent of rainfall patterns. The ruminant-associated genetic marker (Rum2Bac) was detected in all samples but at very low incidence in this watershed. The canine-associated DG3 genetic marker was detected only in samples of water collected after a rainfall. Most probably surface runoff introduced canine excretory waste into the watershed.

TABLE 3 | Relative sequencing read abundance of select general fecal indicators and human-associated microorganisms detected in urban watershed microbiomes.

Associated biomarker	Taxa/Species	References	Dry weather				After rainfall			
			Site 1	Site 2	Site 3	Site 4	Site 1	Site 2	Site 3	Site 4
Bacteria										
General FIB	<i>Bacteroides spp.</i>	Harwood et al., 2017	-	-	<0.01	0.40	0.03	0.03	0.02	0.05
General FIB	<i>Bifidobacterium spp.</i>	Harwood et al., 2017	-	-	-	<0.01	0.01	<0.01	<0.01	0.02
General FIB	<i>Clostridium spp.</i>	Harwood et al., 2017	<0.01	0.03	<0.01	0.02	-	-	-	0.02
General FIB	<i>Citrobacter spp.</i>	Harwood et al., 2017	<0.01	0.01	-	0.01	0.01	0.01	0.01	<0.01
General FIB	<i>Escherichia spp.</i>	Harwood et al., 2017	0.36	0.03	0.04	0.03	0.03	0.05	0.02	0.03
General FIB	<i>Enterobacter spp.</i>	Harwood et al., 2017	0.01	-	0.03	0.05	0.06	0.05	0.06	0.08
General FIB	<i>Klebsiella spp.</i>	Harwood et al., 2017	0.02	0.04	<0.01	0.02	0.09	0.09	0.10	0.07
Human-associated	<i>Bacteroides dorei</i>	Harwood et al., 2017	-	-	-	<0.01	-	-	-	-
Human-associated	<i>Bacteroides thetiotamicron</i>	Harwood et al., 2017	-	-	-	<0.01	-	-	-	0.01
Human-associated	<i>Methanobrevibacter smithii</i>	Harwood et al., 2017	-	-	-	-	-	-	-	<0.01
Human-associated	<i>Bifidobacterium adolescentis</i>	Harwood et al., 2017	-	-	-	<0.01	<0.01	<0.01	<0.01	<0.01
Wastewater-associated	<i>Acinetobacter spp.</i>	Newton et al., 2013	0.21	0.09	0.09	0.56	0.60	0.50	0.71	0.57
Wastewater-associated	<i>Arcobacter spp.</i>	Newton et al., 2013	0.10	0.03	0.10	0.06	0.05	0.27	0.02	0.18
Phages										
General FIB	Somatic Coliphage- <i>Inoviridae</i>	Jebri et al., 2017	-	-	-	-	0.10	0.17	-	0.06
General FIB	Somatic Coliphage- <i>Myoviridae</i>	Jebri et al., 2017	21.10	10.24	12.46	17.88	12.78	15.30	12.06	10.04
General FIB	Somatic Coliphage- <i>Podoviridae</i>	Jebri et al., 2017	13.70	13.17	5.77	45.95	15.01	15.89	11.16	15.96
General FIB	Somatic Coliphage- <i>Siphoviridae</i>	Jebri et al., 2017	64.54	71.87	78.37	35.14	70.92	67.39	74.29	72.80
Human-associated	crAssphage	Kongprajug et al., 2019	-	-	-	0.05	0.43	0.16	0.41	0.11
Human-associated	<i>Bacteroides</i> φ B124-14	Ogilvie et al., 2012	-	-	-	-	-	0.21	-	-

Relative abundance (percentage) of unique sequencing reads for each sample, annotated as respective microbial taxa. Sample RA values represent the proportion within each kingdom, i.e., bacteria and bacteriophages. “-”, not detected.

The WMS and bioinformatic annotation strategy employed in this study identified fecal-associated microorganisms with finer taxonomic resolution; this has been observed in an earlier study (Brumfield et al., 2020b). Microbial species reported to be closely associated with the gut microbiota of specific animal groups (Puig et al., 1999; Buchan et al., 2001; Newton et al., 2013; Shanks et al., 2013; Harwood et al., 2017; Jebri et al., 2017) suggest WMS can be used to shed light on potential sources of fecal pollution. Viruses were detected in samples collected after a rainfall, with crAssphage exhibiting a strikingly similar pattern to the HF183/BacR287 qPCR genetic marker, showing a change between dry weather and post rainfall samples. CrAssphage was recently reported as a dsDNA *Bacteroides* bacteriophage tightly associated with human waste (Dutilh et al., 2014) and highly abundant in sewage (Korajkic et al., 2020). Future experiments comparing qPCR (Stachler et al., 2017) and

WMS measurements of crAssphage in surface waters are planned to confirm this observation.

Microbial Diversity of an Urban Watershed

The autochthonous microbiome of surface water is important in sustainability of natural ecosystems, and potential microbial shifts during periods of dry weather and post rainfall can provide useful information to microbial ecology and sources impacting their occurrence. Previous investigation has associated high taxonomic diversity of lotic ecosystems in urban areas with low fecal contamination in those waterways (Paruch et al., 2019). In the current study, increased fecal pollution appeared to be paralleled with higher alpha diversity following rainfall (Figure 2B), and work is underway to confirm the impact of stormwater runoff on fecal pollution and microbial diversity.

Dominant bacterial phyla detected in the urban watershed (Figures 4, 5) were similar to those detected in surface water of other freshwater aquatic systems (Newton et al., 2011; Staley et al., 2013, 2014; Brown et al., 2015; Hamner et al., 2019), i.e., *Proteobacteria* comprised more than 60% RA in both dry weather and post rainfall samples. *Actinobacteria*, a common soil microbe often present in pristine waterbodies (Jenkins et al., 2009; Ghai et al., 2011), was also prominent in the microbiome of all samples examined in this study. The majority of viruses detected in the urban watershed were bacteriophages belonging to three major bacteriophage families, i.e., *Myoviridae*, *Siphoviridae*, and *Podoviridae* (Figures 4, 5), a finding in agreement with other reported freshwater viromes (Mohiuddin and Schellhorn, 2015). *Thaumarchaeota spp.*, a group of ammonia-oxidizing archaea detected in urban areas demonstrating high rates of nitrification (Reisinger et al., 2016; Epp Schmidt et al., 2019), were present in most samples in this study. *Nitrospirales* were detected predominantly in samples collected following rainfall. These archaea are considered to be important players in recovery of microbial species composing soil microbiomes following disruptive flooding events (Wang et al., 2019). *Clavaria fumosa*, a fungus, and the protozoan *Pseudoperonospora cubensis*, reported by Lee and colleagues (Lee et al., 2020) who used WMS to define the microbial species composition of residential urban stormwater runoff, were detected in this study.

Most watersheds research currently employs FIB identified by culture methods, and a few investigations have adopted qPCR, primarily for enterococci (USEPA, 2012), to link traditional water quality criteria standard definitions to host-associated qPCR. Rapid qPCR for enterococci and *E. coli* offers a shorter sample processing time (< 3 h), compared to traditional FIB culturing (> 18 h). Host-associated genetic markers reliably track common fecal pollution sources, e.g., human, avian, dog, and ruminant, as well as provide quantitative information useful for water quality management. However, these methods were all designed and optimized to detect and quantify a known genetic marker, usually requiring an individual assay to detect each specific gene, without being able to identify uncharacterized microorganisms, i.e., any potential pathogens not yet recognized. Here, use of metagenomic analysis includes thousands of targets including those not yet identified by culture or qPCR expanding the breadth of information regarding the water quality and our knowledge as it relates to the changes in the microbial population.

Community Resistome and Wastewater-Associated Enteric Microorganisms

WMS has been used to explore AR trends in anthropogenically impacted environments (Karkman et al., 2019) and to detect wastewater-associated enteric pathogens (Stamps et al., 2018), including viruses (Lee S. et al., 2019). Stormwater is considered to promote transmission of AR among bacteria (Di Cesare et al., 2017), as well as introduce enteric pathogens (Ahmed et al., 2018), suggesting occurrence of AR and enteric microorganisms can vary between dry weather and after rainfall in our watershed of interest. In this study, we observed an increase in AR genes

detected after a rainfall (Figure 2C). At site 4, which is a location near where the creek debouches into a larger river, there was an abundance of tetracycline resistance genes that were not detected in the dry weather samples. It is possible that the additional AR genes detected at this location were impacted by microorganisms introduced to the creek from the larger river during periods of dry weather. However, the relative abundance of various AR classes detected at each site varied only slightly between sampling events (Figure 2D). These observations are the result of a single grab sample at each location, and additional observations are needed to establish a resistome baseline prior to determining the overall impact that rainfall may have on AR gene composition.

In addition to AR, WMS makes possible detection of multiple wastewater-borne enteric microorganisms without *a priori* knowledge. For example, *Cryptosporidium spp.*, detected at three of the four sites after rainfall (Table 2), currently is the leading cause of reported cases of diarrhea linked to human parasites in contaminated water in the U.S. (Gharpure et al., 2019) and are also known to infect animals. Similarly, *Aeromonas hydrophilia*, an opportunistic pathogen causing gastroenteritis and blood infections was detected in all samples and at slightly increased in RA after rainfall (Table 2). These findings may add support to other reports concluding that *Aeromonas hydrophilia* is omnipresent at low abundance in surface water (Poffé and de Beeck, 1991), but at high concentrations in raw wastewater (Shannon et al., 2007).

The source of microbial contamination of urban watersheds is often varied and can originate from farms, animal feed lots, septic tanks, combined sewer overflow, among others, and modes of delivery can introduce multiple fecal sources to the waterway which can make mitigation difficult. A microbial signature approach, to identify fecal pollution in waters off an urban coast of Lake Michigan, USA, used 16S rRNA sequencing to classify sequences to three wastewater-associated bacterial genera (*Acinetobacter*, *Arcobacter*, and *Trichococcus*) and five fecal-associated bacterial families (*Bacteroidaceae*, *Porphyromonadaceae*, *Clostridiaceae*, *Lachnospiraceae*, and *Ruminococcaceae*) (Newton et al., 2013). Newton and colleagues determined the RA of sewer and fecal signatures increased to >2% of the measured surface water bacterial communities following sewer overflow. Similarly, during a metagenomic survey of wastewater in the United Kingdom, the genera *Arcobacter* and *Aeromonas* were identified as predominant fecal pollution indicators (Acharya et al., 2020). In this study, with the exception of *Trichococcus*, these sewer and fecal signatures were detected at varying RA between dry and post-rainfall sampling events (Table 3, Figures 4, 5).

Metagenomic sequencing can be used to detect and identify a wide range of microorganisms, including bacteria, viruses, fungi, and protists, in some cases to sub-species level. However, viability or infectious potential of detected microorganisms requires additional analyses of metabolic activity. Metagenomic analysis via short read lengths limits the information available within a single read, and characterization of sophisticated genomic structures requires assembled genomes for analysis (Ayling et al., 2020). However, the coverage of microorganisms present at lower RA is often not sufficient to obtain *de novo* metagenomic

assembled genomes. Until novel metagenomic assembly tools are developed, the taxonomic annotation of unassembled sequencing reads remains a viable option for metagenomic data analysis to conserve as much of the less abundant species sequence as possible (Lax et al., 2012; Ponnusamy et al., 2016; Roy et al., 2018; Connelly et al., 2019; Brumfield et al., 2020a). Direct testing of sediments and potential fecal pollution sources harboring pathogens near the study area could also be helpful to index and further describe the microbial communities associated with stormwater runoff and to differentiate the transient stormwater population from the indigenous microbial population.

Conclusions

This pilot study demonstrates the feasibility and application of combining traditional and non-conventional techniques to characterize microbial communities in ambient waters and identify potential sources of pollution. FIB by culture, qPCR amplification of FIB and host-associated genetic markers, and WMS to detect, identify, and enumerate bacteria, archaea, fungi, protists, and viruses were employed. This comprehensive strategy provided useful insights of the microbial constituents present during dry weather and post rainfall in the surface water of an urban watershed and could potentially improve water quality management in the U.S. and globally (United Nations, 2018). Trends were characterized with a limited number of grab samples and are illustrative of a single event representing temporal shifts in microbial communities before and after rainfall. Further assessments are warranted to identify the naturally occurring community ecology and establish a comparative baseline that can be used to verify microbial shifts between dry and rain event conditions, explore the influence of antecedent dry period and stream flow, as well as characterize potential public health significance of metagenomic DNA sequence findings. Future stormwater investigations could also shed light on potential shifts in WMS and qPCR results in response to storm size and duration, soil and particulate content, season, groundwater height, elevation, and land use. It could also be useful to compare sample processing, DNA sequencing, and bioinformatic approaches used here with other strategies. It is worth noting that WMS data reporting is typically limited to shifts in the relative abundance of sequence reads. While this practice can provide a wealth of new information, it does not conclude absolute quantification of specific genomic targets of interest, viability, or infectious potential of the detected microorganisms. Furthermore, the use of WMS for monitoring microbial communities can be more expensive and time consuming than traditional techniques and require advanced bioinformatics software for analyses. Future research directions could also

include the exploration of inter-method correlation between different microbial water quality indicators (Acharya et al., 2019) and compare different WMS strategies for routine integration of these three methods for water quality monitoring. Until the above future research directions are fully characterized and considered in the context of water management, WMS is best applied as a complement to established culture and qPCR practices.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

KB, OS, NH, JC, RC, AH, and ML: conceptualization. KB, OS, MS, JH, JC, and ML: data curation. KB, OS, JC, and ML: formal analysis and writing-original draft preparation. JC, RC, AH, and ML: funding acquisition. KB, OS, MS, JH, JC, and ML: investigation. KB, OS, MS, JH, NH, JC, and ML: methodology. KB, OS, MS, JH, NH, JC, RC, AH, and ML: project administration, writing-review, and editing. KB, OS, and ML: software and visualization. OS, NH, JC, RC, AH, and ML: supervision. KB, OS, JC, NH, and ML: validation. All authors contributed to the article and approved the submitted version.

FUNDING

Research reported in this publication was funded in part by a contract with Joseph Cotruvo and Associates LLC, Washington, DC [KB, JC, AH, and ML]. Support in the form of partial salaries for authors was provided by the National Institute of Environmental Health Sciences of the National Institutes of Health (NIH, <https://www.nih.gov>) under award number R01ES030317A and the National Science Foundation (NSF; <https://www.nsf.gov>) under award number OCE1839171 to Dr. Anwar Huq at the University of Maryland, College Park, MD [KB and AH]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frwa.2021.626849/full#supplementary-material>

REFERENCES

- Acharya, K., Blackburn, A., Mohammed, J., Haile, A. T., Hiruy, A. M., and Werner, D. (2020). Metagenomic water quality monitoring with a portable laboratory. *Water Res.* 184:116112. doi: 10.1016/j.watres.2020.116112
- Acharya, K., Khanal, S., Pantha, K., Amatya, N., Davenport, R. J., and Werner, D. (2019). A comparative assessment of conventional and molecular methods, including MinION nanopore sequencing, for surveying water quality. *Sci. Rep.* 9:15726. doi: 10.1038/s41598-019-51997-x

- Acker, J. G., and Leptoukh, G. (2007). Online analysis enhances use of NASA Earth science data. *Eos Trans. Am. Geophys. Union* 88, 14–17. doi: 10.1029/2007EO020003
- Ahmed, W., Zhang, Q., Lobos, A., Senkbeil, J., Sadowsky, M. J., Harwood, V. J., et al. (2018). Precipitation influences pathogenic bacteria and antibiotic resistance gene abundance in storm drain outfalls in coastal sub-tropical waters. *Environ. Int.* 116, 308–318. doi: 10.1016/j.envint.2018.04.005
- Alexander, J. R., and Groffman, P. M., and Rosi-Marshall, E. J. (2016). Nitrogen-cycling process rates across urban ecosystems. *FEMS Microbiol. Ecol.* 92:fiw198. doi: 10.1093/femsec/fiw198
- Andrews, S. C. (2019). *FastQC*. Available online at: <https://github.com/s-andrews/FastQC> (accessed October 2, 2019)
- Auld, H., MacIver, D., and Klaassen, J. (2004). Heavy rainfall and waterborne disease outbreaks: the Walkerton example. *J. Toxicol. Environ. Health A* 67, 1879–1887. doi: 10.1080/15287390490493475
- Ayling, M., Clark, M. D., and Leggett, R. M. (2020). New approaches for metagenome assembly with short reads. *Brief. Bioinform.* 21, 584–594. doi: 10.1093/bib/bbz020
- Brown, B. L., LePrell, R. V., Franklin, R. B., Rivera, M. C., Cabral, F. M., Eaves, H. L., et al. (2015). Metagenomic analysis of planktonic microbial consortia from a non-tidal urban-impacted segment of James River. *Stand. Genomic Sci.* 10:65. doi: 10.1186/s40793-015-0062-5
- Brumfield, K. D., Hasan, N. A., Leddy, M. B., Cotruvo, J., Rashed, S. M., Colwell, R. R., et al. (2020a). A comparative analysis of drinking water employing metagenomics. *PLoS ONE* 15:e0231210. doi: 10.1371/journal.pone.0231210
- Brumfield, K. D., Huq, A., Colwell, R. R., Olds, J. L., and Leddy, M. B. (2020b). Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. *PLoS ONE* 15:e0228899. doi: 10.1371/journal.pone.0228899
- Buchan, A., Alber, M., and Hodson, R. E. (2001). Strain-specific differentiation of environmental *Escherichia coli* isolates via denaturing gradient gel electrophoresis (DGGE) analysis of the 16S–23S intergenic spacer region. *FEMS Microbiol. Ecol.* 35, 313–321. doi: 10.1111/j.1574-6941.2001.tb00817.x
- Burcham, Z. M., Schmidt, C. J., Pechal, J. L., Brooks, C. P., Rosch, J. W., Benbow, M. E., et al. (2019). Detection of critical antibiotic resistance genes through routine microbiome surveillance. *PLoS ONE* 14:e0213280. doi: 10.1371/journal.pone.0213280
- Byappanahalli, M. N., Nevers, M. B., Korajkic, A., Staley, Z. R., and Harwood, V. J. (2012). *Enterococci* in the environment. *Microbiol. Mol. Biol. Rev.* 76, 685–706. doi: 10.1128/MMBR.00023-12
- Cao, Y., Sivaganesan, M., Kelly, C. A., Wang, D., Boehm, A. B., Griffith, J. F., et al. (2018). A human fecal contamination score for ranking recreational sites using the HF183/BacR287 quantitative real-time PCR method. *Water Res.* 128, 148–156. doi: 10.1016/j.watres.2017.10.071
- Chern, E. C., Siefring, S., Paar, J., Doolittle, M., and Haugland, R. A. (2011). Comparison of quantitative PCR assays for *Escherichia coli* targeting ribosomal RNA and single copy genes. *Lett. Appl. Microbiol.* 52, 298–306. doi: 10.1111/j.1472-765X.2010.03001.x
- Connelly, S., Fanelli, B., Hasan, N. A., Kaleko, M., and Colwell, R. R. (2019). Oral metallo-beta-lactamase protects the gut microbiome from carbapenem-mediated damage and reduces propagation of antibiotic resistance in pigs. *Front. Microbiol.* 10:101. doi: 10.3389/fmicb.2019.00101
- Cosmos ID. (2019). *CosmosID Metagenomics Cloud*. Available online at: <https://app.cosmosid.com/login> (accessed October 2, 2019).
- Curriero, F. C., Patz, J. A., Rose, J. B., and Lele, S. (2001). The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948–1994. *Am. J. Public Health* 91, 1194–1199. doi: 10.2105/ajph.91.8.1194
- DCMR. (2014). *DCMR Water Quality Standards, 21-1104.8*. Available online at: <https://www.epa.gov/sites/production/files/2014-12/documents/dcwqs-2010.pdf> (accessed February 1, 2021).
- Di Cesare, A., Eckert, E. M., Rogora, M., and Corno, G. (2017). Rainfall increases the abundance of antibiotic resistance genes within a riverine microbial community. *Environ. Pollut.* 226, 473–478. doi: 10.1016/j.envpol.2017.04.036
- Dick, L. K., Bernhard, A. E., Brodeur, T. J., Santo Domingo, J. W., Simpson, J. M., Walters, S. P., et al. (2005). Host distributions of uncultivated fecal Bacteroidales bacteria reveal genetic markers for fecal source identification. *Appl. Environ. Microbiol.* 71, 3184–3191. doi: 10.1128/AEM.71.6.3184-3191.2005
- Dutilh, B. E., Cassman, N., McNair, K., Sanchez, S. E., Silva, G. G. Z., Boling, L., et al. (2014). A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* 5:4498. doi: 10.1038/ncomms5498
- EPA. (2002). *Federal Water Pollution Control Act*. Available online at: <https://www.epa.gov/sites/production/files/2017-08/documents/federal-water-pollution-control-act-508full.pdf> (accessed March 19, 2020).
- EPA. (2004). Microbial Source Tracking Guide Document. EPA/600/R-05/064.
- EPA. (2012). *2012 Recreational Water Quality Criteria*. EPA
- EPA. (2019). *Clean Water Act Analytical Methods*. EPA
- Epp Schmidt, D. J., Kotze, D. J., Hornung, E., Setälä, H., Yesilonis, I., Szlavetz, K., et al. (2019). Metagenomics reveals bacterial and archaeal adaptation to urban land-use: N catabolism, methanogenesis, and nutrient acquisition. *Front. Microbiol.* 10:2330. doi: 10.3389/fmicb.2019.02330
- Ferrera, I., and Sánchez, O. (2016). Insights into microbial diversity in wastewater treatment systems: how far have we come? *Biotechnol. Adv.* 34, 790–802. doi: 10.1016/j.biotechadv.2016.04.003
- Fisher, J. C., Newton, R. J., Dila, D. K., and McLellan, S. L. (2015). Urban microbial ecology of a freshwater estuary of Lake Michigan. *Elementa* 3:64. doi: 10.12952/journal.elementa.000064
- Ghai, R., Rodriguez-Valera, F., McMahon, K. D., Toyama, D., Rinke, R., Cristina Souza de Oliveira, T., et al. (2011). Metagenomics of the water column in the pristine upper course of the Amazon river. *PLoS ONE* 6:e23785. doi: 10.1371/journal.pone.0023785
- Gharpure, R., Perez, A., Miller, A. D., Wikswo, M. E., Silver, R., and Hlavsa, M. C. (2019). Cryptosporidiosis Outbreaks - United States, 2009–2017. *MMWR. Morb. Mortal. Wkly. Rep.* 68, 568–572. doi: 10.15585/mmwr.mm6825a3
- Gibb, K., Schobben, X., and Christian, K. (2017). Frogs host faecal bacteria typically associated with humans. *Can. J. Microbiol.* 63, 633–637. doi: 10.1139/cjm-2017-0119
- Green, H. C., Dick, L. K., Gilpin, B., Samadpour, M., and Field, K. G. (2012). Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. *Appl. Environ. Microbiol.* 78, 503–510. doi: 10.1128/AEM.05734-11
- Green, H. C., White, K. M., Kelly, C. A., and Shanks, O. C. (2014). Development of rapid canine fecal source identification PCR-based assays. *Environ. Sci. Technol.* 48, 11453–11461. doi: 10.1021/es502637b
- Hamilton, K. A., Garner, E., Joshi, S., Ahmed, W., Ashbolt, N., Medema, G., et al. (2020). Antimicrobial-resistant microorganisms and their genetic determinants in stormwater: a systematic review. *Curr. Opin. Environ. Sci. Heal.* 16, 101–112. doi: 10.1016/j.coesh.2020.02.012
- Hamner, S., Brown, B. L., Hasan, N. A., Franklin, M. J., Doyle, J., Eggers, M. J., et al. (2019). Metagenomic profiling of microbial pathogens in the little bighorn river, Montana. *Int. J. Environ. Res. Public Health* 16:1097. doi: 10.3390/ijerph16071097
- Harwood, V., Shanks, O., Korajkic, A., Verbyla, M., Ahmed, W., and Iriate, M. (2017). “General and host-associated bacterial indicators of faecal pollution,” in *Global Water Pathogen Project*, eds J. B. Rose and B. Jiménez-Cisneros (A.Farnleitner, and A. Blanch (eds) Part 2 Indicators and Microbial Source Tracking Markers) (East Lansing, MI: Michigan State University; UNESCO). doi: 10.14321/waterpathogens.6. Available online at: <http://www.waterpathogens.org/book/bacterial-indicators> (accessed February 1, 2021).
- Harwood, V. J., Levine, A. D., Scott, T. M., Chivukula, V., Lukasik, J., Farrah, S. R., et al. (2005). Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Appl. Environ. Microbiol.* 71, 3163–3170. doi: 10.1128/AEM.71.6.3163-3170.2005
- Jebri, S., Muniesa, M., and Jofre, J. (2017). “General and host-associated bacteriophage indicators of faecal pollution,” in *Global Water Pathogen Project*, eds J. B. Rose and B. Jiménez-Cisneros (A.Farnleitner, and A. Blanch (eds) Part 2 Indicators and Microbial Source Tracking Markers) (East Lansing, MI: Michigan State University; UNESCO). doi: 10.14321/waterpathogens.7. Available online at: <http://www.waterpathogens.org/book/coliphage> (accessed February 1, 2021).
- Jenkins, S. N., Waite, I. S., Blackburn, A., Husband, R., Rushton, S. P., Manning, D. C., et al. (2009). Actinobacterial community dynamics in

- long term managed grasslands. *Antonie Van Leeuwenhoek* 95, 319–334. doi: 10.1007/s10482-009-9317-8
- Jin, D., Kong, X., Cui, B., Jin, S., Xie, Y., Wang, X., et al. (2018). Bacterial communities and potential waterborne pathogens within the typical urban surface waters. *Sci. Rep.* 8:13368. doi: 10.1038/s41598-018-31706-w
- Karkman, A., Pärnänen, K., and Larsson, D. G. J. (2019). Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments. *Nat. Commun.* 10:80. doi: 10.1038/s41467-018-07992-3
- Kongprajug, A., Mongkolsuk, S., and Sirikanchana, K. (2019). CrAssphage as a potential human sewage marker for microbial source tracking in Southeast Asia. *Environ. Sci. Technol. Lett.* 6, 159–164. doi: 10.1021/acs.estlett.9b00041
- Korajkic, A., McMin, B., Herrmann, M. P., Sivaganesan, M., Kelty, C. A., Clinton, P., et al. (2020). Viral and bacterial fecal indicators in untreated wastewater across the contiguous united states exhibit geospatial trends. *Appl. Environ. Microbiol.* 86, e02967–e02919. doi: 10.1128/AEM.02967-19
- Lax, S., Smith, D. P., Hampton-marcell, J., Owens, S. M., Shogan, B. D., Weiss, S., et al. (2012). Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science* 345, 1048–1052. doi: 10.1126/science.1254529
- Lee, C., Marion, J. W., and Lee, J. (2013). Development and application of a quantitative PCR assay targeting *Catellibacillus marimammalius* for assessing gull-associated fecal contamination at Lake Erie beaches. *Sci. Total Environ.* 454–455, 1–8. doi: 10.1016/j.scitotenv.2013.03.003
- Lee, D., Chang, H. H., Sarnat, S. E., and Levy, K. (2019). Precipitation and salmonellosis incidence in Georgia, USA: interactions between extreme rainfall events and antecedent rainfall conditions. *Environ. Health Perspect.* 127: 97005. doi: 10.1289/EHP4621
- Lee, S., Suits, M., Witusynski, D., Winston, R., Martin, J., and Lee, J. (2020). Residential urban stormwater runoff: a comprehensive profile of microbiome and antibiotic resistance. *Sci. Total Environ.* 723:138033. doi: 10.1016/j.scitotenv.2020.138033
- Lee, S., Suwa, M., and Shigemura, H. (2019). Metagenomic analysis of infectious F-specific RNA bacteriophage strains in wastewater treatment and disinfection processes. *Pathogens* 8:217. doi: 10.3390/pathogens8040217
- Li, X., Sivaganesan, M., Kelty, C. A., Zimmer-Faust, A., Clinton, P., Reichman, J. R., et al. (2019). Large-scale implementation of standardized quantitative real-time PCR fecal source identification procedures in the Tillamook Bay Watershed. *PLoS ONE* 14:e0216827. doi: 10.1371/journal.pone.0216827
- Ludwig, W., and Schleifer, K.-H. (2000). How quantitative is quantitative PCR with respect to cell counts? *Syst. Appl. Microbiol.* 23, 556–562. doi: 10.1016/S0723-2020(00)80030-2
- Mathai, P. P., Staley, C., and Sadowsky, M. J. (2020). Sequence-enabled community-based microbial source tracking in surface waters using machine learning classification: a review. *J. Microbiol. Methods* 177:106050. doi: 10.1016/j.mimet.2020.106050
- Meays, C. L., Broersma, K., Nordin, R., and Mazumder, A. (2004). Source tracking fecal bacteria in water: a critical review of current methods. *J. Environ. Manage* 73, 71–79. doi: 10.1016/j.jenvman.2004.06.001
- Mieszkis, S., Furet, J.-P., Corthier, G., and Gourmelon, M. (2009). Estimation of pig fecal contamination in a river catchment by real-time PCR using two pig-specific bacteroidales 16S rRNA genetic markers. *Appl. Environ. Microbiol.* 75, 3045–3054. doi: 10.1128/AEM.02343-08
- Mieszkis, S., Yala, J.-F., Joubrel, R., and Gourmelon, M. (2010). Phylogenetic analysis of Bacteroidales 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. *J. Appl. Microbiol.* 108, 974–984. doi: 10.1111/j.1365-2672.2009.04499.x
- Mohiuddin, M., and Schellhorn, H. (2015). Spatial and temporal dynamics of virus occurrence in two freshwater lakes captured through metagenomic analysis. *Front. Microbiol.* 6:960. doi: 10.3389/fmicb.2015.00960
- Newton, R. J., Bootsma, M. J., Morrison, H. G., Sogin, M. L., and McLellan, S. L. (2013). A microbial signature approach to identify fecal pollution in the waters off an urbanized coast of Lake Michigan. *Microb. Ecol.* 65, 1011–1023. doi: 10.1007/s00248-013-0200-9
- Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., and Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.* 75, 14–49. doi: 10.1128/MMBR.00028-10
- Ogilvie, L. A., Caplin, J., Dedi, C., Diston, D., Cheek, E., Bowler, L., et al. (2012). Comparative (meta)genomic analysis and ecological profiling of human gut-specific bacteriophage phiB124-14. *PLoS ONE* 7:e35053. doi: 10.1371/journal.pone.0035053
- Ohad, S., Ben-Dor, S., Prilusky, J., Kravitz, V., Dassa, B., Chalifa-Caspi, V., et al. (2016). The development of a novel qPCR assay-set for identifying fecal contamination originating from domestic fowls and waterfowl in Israel. *Front. Microbiol.* 7:145. doi: 10.3389/fmicb.2016.00145
- Ondov, B. D., Bergman, N. H., and Phillippy, A. M. (2011). Interactive metagenomic visualization in a web browser. *BMC Bioinformatics* 12:385. doi: 10.1186/1471-2105-12-385
- Paruch, L., Paruch, A. M., Eiken, H. G., and Sørheim, R. (2019). Faecal pollution affects abundance and diversity of aquatic microbial community in anthropo-zoogenically influenced lotic ecosystems. *Sci. Rep.* 9:19469. doi: 10.1038/s41598-019-56058-x
- Poffé, R., and de Bree, E. O. (1991). Enumeration of *Aeromonas hydrophila* from domestic wastewater treatment plants and surface waters. *J. Appl. Bacteriol.* 71, 366–370. doi: 10.1111/j.1365-2672.1991.tb03802.x
- Ponnusamy, D., Kozlova, E. V., Sha, J., Erova, T. E., Azar, S. R., and Fitts, E. C. (2016). Cross-talk among flesh-eating *Aeromonas hydrophila* strains in mixed infection leading to necrotizing fasciitis. *Proc. Natl. Acad. Sci. U.S.A.* 113, 722–727. doi: 10.1073/pnas.1523817113
- Puig, A., Queralt, N., Jofre, J., and Araujo, R. (1999). Diversity of *Bacteroides fragilis* strains in their capacity to recover phages from human and animal wastes and from fecally polluted wastewater. *Appl. Environ. Microbiol.* 65, 1772–1776.
- Pusch, D., Oh, D.-Y., Wolf, S., Dumke, R., Schröter-Bobsin, U., Höhne, M., et al. (2005). Detection of enteric viruses and bacterial indicators in German environmental waters. *Arch. Virol.* 150, 929–947. doi: 10.1007/s00705-004-0467-8
- Ramírez-Castillo, F., Loera-Muro, A., Jacques, M., Garneau, P., Avelar-González, F., Harel, J., et al. (2015). Waterborne pathogens: detection methods and challenges. *Pathogens* 4, 307–334. doi: 10.3390/pathogens4020307
- Richardson, M., and Rautemaa-Richardson, R. (2019). Exposure to aspergillus in home and healthcare facilities' water environments: focus on biofilms. *Microorganisms* 7:7. doi: 10.3390/microorganisms7010007
- Rojas, A., Segev, G., Markovics, A., Aroch, I., and Baneth, G. (2017). Detection and quantification of *Spirocera lupi* by HRM qPCR in fecal samples from dogs with spirocercosis. *Parasit. Vectors* 10:435. doi: 10.1186/s13071-017-2374-3
- Roy, M. A., Arnaud, J. M., Jasmin, P. M., Hamner, S., Hasan, N. A., Colwell, R. R., et al. (2018). A metagenomic approach to evaluating surface water quality in haiti. *Int. J. Environ. Res. Public Health* 15:2211. doi: 10.3390/ijerph15102211
- Schilling, K. E., Zhang, Y.-K., Hill, D. R., Jones, C. S., and Wolter, C. F. (2009). Temporal variations of *Escherichia coli* concentrations in a large Midwestern river. *J. Hydrol.* 365, 79–85. doi: 10.1016/j.jhydrol.2008.11.029
- Shanks, O. C., Atikovic, E., Blackwood, A. D., Lu, J., Noble, R. T., Domingo, J. S., et al. (2008). Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. *Appl. Environ. Microbiol.* 74, 745–752. doi: 10.1128/AEM.01843-07
- Shanks, O. C., Kelty, C. A., Oshiro, R., Haugland, R. A., Madi, T., Brooks, L., et al. (2016). Data acceptance criteria for standardized human-associated fecal source identification quantitative real-time PCR methods. *Appl. Environ. Microbiol.* 82, 2773–2782. doi: 10.1128/AEM.03661-15
- Shanks, O. C., Kelty, C. A., Sivaganesan, M., Varma, M., and Haugland, R. A. (2009). Quantitative PCR for genetic markers of human fecal pollution. *Appl. Environ. Microbiol.* 75, 5507–5513. doi: 10.1128/AEM.00305-09
- Shanks, O. C., Newton, R. J., Kelty, C. A., Huse, S. M., Sogin, M. L., and McLellan, S. L. (2013). Comparison of the microbial community structures of untreated wastewaters from different geographic locales. *Appl. Environ. Microbiol.* 79, 2906–2913. doi: 10.1128/AEM.03448-12
- Shannon, K. E., Lee, D.-Y., Trevors, J. T., and Beaudette, L. A. (2007). Application of real-time quantitative PCR for the detection of selected bacterial pathogens during municipal wastewater treatment. *Sci. Total Environ.* 382, 121–129. doi: 10.1016/j.scitotenv.2007.02.039
- Shrestha, A., Kelty, C. A., Sivaganesan, M., Shanks, O. C., and Dorevitch, S. (2020). Fecal pollution source characterization at non-point source impacted

- beaches under dry and wet weather conditions. *Water Res.* 182:116014. doi: 10.1016/j.watres.2020.116014
- Siefring, S., Varma, M., Atikovic, E., Wymer, L., and Haugland, R. A. (2008). Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. *J. Water Health* 6, 225–237. doi: 10.2166/wh.2008.022
- Sivaganesan, M., Haugland, R. A., Chern, E. C., and Shanks, O. C. (2010). Improved strategies and optimization of calibration models for real-time PCR absolute quantification. *Water Res.* 44, 4726–4735. doi: 10.1016/j.watres.2010.07.066
- Stachler, E., Keltly, C., Sivaganesan, M., Li, X., Bibby, K., and Shanks, O. C. (2017). Quantitative CrAssphage PCR assays for human fecal pollution measurement. *Environ. Sci. Technol.* 51, 9146–9154. doi: 10.1021/acs.est.7b02703
- Staley, C., Gould, T. J., Wang, P., Phillips, J., Cotner, J. B., and Sadowsky, M. J. (2014). Core functional traits of bacterial communities in the Upper Mississippi river show limited variation in response to land cover. *Front. Microbiol.* 5:414. doi: 10.3389/fmicb.2014.00414
- Staley, C., Unno, T., Gould, T. J., Jarvis, B., Phillips, J., Cotner, J. B., et al. (2013). Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *J. Appl. Microbiol.* 115, 1147–1158. doi: 10.1111/jam.12323
- Staley, Z. R., Boyd, R. J., Shum, P., and Edge, T. A. (2018). Microbial source tracking using quantitative and digital PCR To identify sources of fecal contamination in stormwater, river water, and beach water in a Great Lakes area of concern. *Appl. Environ. Microbiol.* 84, e01634–e01618. doi: 10.1128/AEM.01634-18
- Stamps, B. W., Leddy, M. B., Plumlee, M. H., Hasan, N. A., Colwell, R. R., and Spear, J. R. (2018). Characterization of the microbiome at the World's largest potable water reuse facility. *Front. Microbiol.* 9:2435. doi: 10.3389/fmicb.2018.02435
- Su, J.-Q., An, X.-L., Li, B., Chen, Q.-L., Gillings, M. R., Chen, H., et al. (2017). Metagenomics of urban sewage identifies an extensively shared antibiotic resistome in China. *Microbiome* 5:84. doi: 10.1186/s40168-017-0298-y
- United Nations. (2018). *Sustainable Development Goal 6: Synthesis Report on Water and Sanitation*. New York, NY: United Nations. Available online at: https://sustainabledevelopment.un.org/content/documents/19901SDG6_SR2018_web_3.pdf (accessed February 1, 2021).
- University of Cambridge. (2020). *WinBUGS*. Available online at: <https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/> (accessed June 5, 2020).
- USEPA. (2012). *Water: monitoring and assessment-5.11 fecal bacteria*. Available at: <https://archive.epa.gov/water/archive/web/html/vms511.html> (accessed March 19, 2020).
- USEPA. (2019). *USEPA Method 1699: Characterization of Human Fecal Pollution in Water by HF183/BacR287 TaqMan Quantitative Polymerase Chain Reaction (qPCR) Assay*. Washington, DC: USEPA.
- Uyaguari-Diaz, M. I., Chan, M., Chaban, B. L., Croxen, M. A., Finke, J. F., Hill, J. E., et al. (2016). A comprehensive method for amplicon-based and metagenomic characterization of viruses, bacteria, and eukaryotes in freshwater samples. *Microbiome* 4:20. doi: 10.1186/s40168-016-0166-1
- Wade, T. J., Sams, E., Brenner, K. P., Haugland, R., Chern, E., Beach, M., et al. (2010). Rapidly measured indicators of recreational water quality and swimming-associated illness at marine beaches: a prospective cohort study. *Environ. Heal.* 9:66. doi: 10.1186/1476-069X-9-66
- Wang, P., Chen, B., Yuan, R., Li, C., and Li, Y. (2016). Characteristics of aquatic bacterial community and the influencing factors in an urban river. *Sci. Total Environ.* 569–570, 382–389. doi: 10.1016/j.scitotenv.2016.06.130
- Wang, Y., Hong, Y., Ma, M., Wu, S., Op den Camp, H. J. M., Zhu, G., et al. (2019). Anthropogenic pollution intervenes the recovery processes of soil archaeal community composition and diversity from flooding. *Front. Microbiol.* 10:2285. doi: 10.3389/fmicb.2019.02285
- WHO. (2017). *Potable Reuse: Guidance for Producing Safe Drinking-Water*. Geneva. Available online at: <https://apps.who.int/iris/bitstream/handle/10665/258715/9789241512770-eng.pdf?sequence=1&isAllowed=y> (accessed February 1, 2021).
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer. Available online at: <https://ggplot2.tidyverse.org>
- Wu, H., Oun, A., Kline-Robach, R., and Xagorarakis, I. (2018). Microbial pollution characterization at a TMDL site in Michigan: source identification. *J. Great Lakes Res.* 44, 412–420. doi: 10.1016/j.jglr.2018.02.006
- Yang, Y., Xu, C., Cao, X., Lin, H., and Wang, J. (2017). Antibiotic resistance genes in surface water of eutrophic urban lakes are related to heavy metals, antibiotics, lake morphology and anthropic impact. *Ecotoxicology* 26, 831–840. doi: 10.1007/s10646-017-1814-3

Disclaimer: Information has been subjected to USEPA peer and administrative review and has been approved for external publication. Any opinions expressed in this paper are those of the authors and do not necessarily reflect the official positions and policies of the USEPA. Any mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Conflict of Interest: RC is the founder of CosmosID, Inc., Rockville, MD and also a Professor at the University of Maryland. NH was an employee of CosmosID, Inc. JC is the founder of Joseph Cotruvo and Associates LLC, Washington, DC. ML is the founder of Essential Environmental and Engineering Systems, Huntington Beach, CA. The specific roles of these authors have been articulated in the Author Contributions section.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Brumfield, Cotruvo, Shanks, Sivaganesan, Hey, Hasan, Huq, Colwell and Leddy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.