Restructuring of the Aquatic Bacterial Community by Hydric Dynamics Associated with Superstorm Sandy

Nikea Ulrich,^{a,b} Abigail Rosenberger,^a Colin Brislawn,^a Justin Wright,^{a,b} Collin Kessler,^a David Toole,^a Caroline Solomon,^a Steven Strutt,^a Erin McClure,^a Regina Lamendella^{a,b}

Juniata College, Department of Biology, Huntingdon, Pennsylvania, USA^a; WrightLabs LLC, Huntingdon, Pennsylvania, USA^b

ABSTRACT

Bacterial community composition and longitudinal fluctuations were monitored in a riverine system during and after Superstorm Sandy to better characterize inter- and intracommunity responses associated with the disturbance associated with a 100-year storm event. High-throughput sequencing of the 16S rRNA gene was used to assess microbial community structure within water samples from Muddy Creek Run, a second-order stream in Huntingdon, PA, at 12 different time points during the storm event (29 October to 3 November 2012) and under seasonally matched baseline conditions. High-throughput sequencing of the 16S rRNA gene was used to track changes in bacterial community structure and divergence during and after Superstorm Sandy. Bacterial community dynamics were correlated to measured physicochemical parameters and fecal indicator bacteria (FIB) concentrations. Bioinformatics analyses of 2.1 million 16S rRNA gene sequences revealed a significant increase in bacterial diversity in samples taken during peak discharge of the storm. Beta-diversity analyses revealed longitudinal shifts in the bacterial community structure. Successional changes were observed, in which *Betaproteobacteria* and *Gammaproteobacteria* decreased in 16S rRNA gene relative abundance, while the relative abundance of members of the *Firmicutes* increased. Furthermore, 16S rRNA gene sequences matching pathogenic bacteria, including strains of *Legionella*, *Campylobacter*, *Arcobacter*, and *Helicobacter*, as well as bacteria of fecal origin (e.g., *Bacteroides*), exhibited an increase in abundance after peak discharge of the storm. This study revealed a significant restructuring of in-stream bacterial community structure associated with hydric dynamics of a storm event.

IMPORTANCE

In order to better understand the microbial risks associated with freshwater environments during a storm event, a more comprehensive understanding of the variations in aquatic bacterial diversity is warranted. This study investigated the bacterial communities during and after Superstorm Sandy to provide fine time point resolution of dynamic changes in bacterial composition. This study adds to the current literature by revealing the variation in bacterial community structure during the course of a storm. This study employed high-throughput DNA sequencing, which generated a deep analysis of inter- and intracommunity responses during a significant storm event. This study has highlighted the utility of applying high-throughput sequencing for water quality monitoring purposes, as this approach enabled a more comprehensive investigation of the bacterial community structure. Altogether, these data suggest a drastic restructuring of the stream bacterial community during a storm event and highlight the potential of high-throughput sequencing approaches for assessing the microbiological quality of our environment.

More than 12,000 water bodies in the United States are considered to be impaired by fecal indicator bacteria (FIB), due to both point and nonpoint sources of pollution, according to the U.S. Environmental Protection Agency (USEPA) (1). Point sources of pollution are those that can be pinpointed to a specific location, such as discharges from wastewater treatment plants, operational wastes from industries, as well as combined and sanitary sewer overflows (1). Nonpoint sources are more diffuse and may be associated with a type of land use, such as agriculture or industry. Factors, including surface runoff, precipitation, drainage, and seepage, are all potential contributors to nonpoint source pollution (1). These sources of runoff can carry nutrients, toxins, and microorganisms into nearby water basins, specifically during a high-volume storm event.

Storm events have been shown to affect the composition of nutrients in the stream and cause acute disturbances in short-term stream health. Furthermore, storm events represent a significant contribution to stream impairments, as they represent a majority of the bacterial discharge and fecal loading in a watershed (2–5). Fecal contamination and pathogen concentrations above regula-

tory limits are predominant concerns of flood and runoff waters, and exposure to such waterborne pathogens has been shown to increase the potential for infection and accompanying health risks (6–15). Rain and floodwater harbor a variety of human enteric pathogens, including *Campylobacter*, *Cryptosporidium*, *Giardia*, adenoviruses, polyomaviruses, and enteroviruses (16–18). Feces

Received 22 February 2016 Accepted 30 March 2016

Accepted manuscript posted online 8 April 2016

Citation Ulrich N, Rosenberger A, Brislawn C, Wright J, Kessler C, Toole D, Solomon C, Strutt S, McClure E, Lamendella R. 2016. Restructuring of the aquatic bacterial community by hydric dynamics associated with Superstorm Sandy. *Appl Environ Microbiol* 82:3525–3536. doi:10.1128/AEM.00520-16.

Editor: P. D. Schloss, University of Michigan

Address correspondence to Regina Lamendella, lamendella@juniata.edu.

N.U. and A.R. contributed equally to this article.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.00520-16>.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.

from agricultural animals and wildlife are sources of a plethora of zoonotic pathogens (e.g., pathogenic *Escherichia coli*, *Salmonella*, and *Leptospira*) (9, 19). In addition to fecal pathogens, rainwater has been shown to harbor *Legionella pneumophila*, *Aeromonas hydrophila*, and *Clostridium perfringens* (20–22).

Traditional approaches and molecular assays have been used in an attempt to assess the microbiological quality of runoff-impacted environments. Previous studies have linked storm events to increases in fecal indicator bacteria (FIB) concentrations and turbidity (23, 24) as well as decreases in general water quality (2, 25). Various microbial source-tracking methods have been used in an attempt to track and quantify sources of fecal contamination in multiuse watersheds; however, research suggests these methods are inadequate at predicting and quantifying risk (26–28). The duration of storm events and influence of stormwater plumes are not easily evaluated, as contaminant origin is difficult to determine because of numerous potential contributions (26).

In order to better understand the microbial risks associated with freshwater environments during a storm event, a more comprehensive understanding of the spatial and temporal variations in aquatic bacterial diversity is needed. Previous studies have indicated that bacterial community structure in rivers varies seasonally and due to physical and chemical gradients (29–32). Afshinnekoo et al. (33) revealed that flooding events during Superstorm Sandy had long-lasting effects on the microbial communities of city surfaces. However, the variation in bacterial community structure during the course of a storm event within a stream environment has been understudied. High-throughput DNA sequencing has enabled deep analysis of inter- and intracommunity responses to environmental variables. In fact, recent research (34, 35) has indicated that rapid evaluation of the total bacterial community composition shows promise as a potential water quality monitoring tool. The investigation of successional changes with respect to bacterial community structure enables a more comprehensive investigation of water quality to detect rare populations that account for increased diversity in streams (36–38). High-throughput sequencing approaches have improved the identification of source-specific bacteria that serve as important indicators of fecal input in stream water (36, 39).

Here, we evaluated the bacterial community composition and successional changes in a second-order stream in Huntingdon, PA, during and after Superstorm Sandy (29 October to 3 November 2012). Superstorm Sandy was ranked a category 1 hurricane that covered 1.8 million square miles of the mid-Atlantic coast and into Canada and New England, according to the National Aeronautics and Space Administration (NASA) (40). Pennsylvania was among the most affected states, with heavy rainfall between 100 mm and 200 mm (40). High-throughput sequencing of the 16S rRNA gene was used to track temporal dynamics of inter- and intracommunity responses during the storm. This study provides fine time point resolution of bacterial composition during a 100-year storm event. Additionally, traditional fecal indicator organisms, including fecal coliforms and *E. coli* concentrations, were measured to evaluate water quality during the sampling period. Dramatic changes in the bacterial community structure and diversity were found during the sampling period. Several taxa fluctuated with discharge rate, suggesting successional changes occur within the bacterial assemblages in stream water ecosystems during storm events. In addition, these data highlight increases in the relative abundance of sequences matching fecal bacteria and po-

tentially pathogenic populations after the storm event. Altogether, these data suggest a drastic restructuring of the stream bacterial community and highlight the potential of high-throughput sequencing approaches in assessing the microbiological quality of our environment.

MATERIALS AND METHODS

Site description and sampling. Sampling was performed on Muddy Run Creek, a 3.1-mile tributary of the Juniata River that travels through Huntingdon, PA (40°29'51.85"N, 78°0'51.02"W). The initial 1.5 miles of the second-order stream are fed mainly by surface runoff from residential housing developments and agricultural regions. Downstream of the sampling location, the tributary becomes subterranean and is fed by a multitude of sources, including groundwater, residential runoff, stormwater, and potentially septic sources (see Fig. S1 in the supplemental material). Water samples ($n = 25$) were collected in duplicate or triplicate during Superstorm Sandy at 12 different time points, starting on 29 October 2012 (day 1, 0.00 h; 20.81 m³/s), before peak discharge rates (270 m³/s) were measured. On 29 October 2012, the center of Superstorm Sandy lay in southeast Pennsylvania, with sustaining winds near 65 mph, and storm-force winds extended almost 500 miles from the center as it moved along the northeastern coast (41). Sampling was performed at various time points during the storm to enable robust coverage of stream dynamics. Water sampling was continued at the following time points: 30 October 2012 (day 2, 11.7 h, 14.9 h, 17.78 h, 21.12 h, and 23.92 h), 31 October 2012 (day 3, 26.42 h, 35.57 h, 42.3 h, and 46.78 h), 1 November 2012 (day 4, 67.11 h), and 3 November 2012 (day 5, 115.28 h). Samples of water (300 to 600 ml) were filtered immediately through 0.22- μ m-pore-size polyethersulfone filters (Millipore, Billerica, MA) and stored at -20°C until further processing. Additional sampling was performed in triplicate over a span of 4 days (8 to 11 October 2015) at three different time points to provide a baseline survey of bacterial community structure and stream chemistry under nonstorm conditions.

Stream water chemistry measurements (conductivity, pH, temperature, salinity, and total dissolved solids [TDS]) were taken on site at the time of sampling using a precalibrated PCTest 35 multiparameter probe (Oakton, Vernon Hills, IL) (see Table S1 in the supplemental material). *E. coli* and total coliforms were enumerated using the Colilert-18 test (Idexx, Westbrook, ME). Flow rate was approximated using data from the nearest U.S. Geological Survey (USGS) gauge station (station 01559000) for the Juniata River in Huntingdon (40°29'05"N, 78°01'09"W) to generate a hydrograph during the sampling period.

DNA extraction and 16S rRNA gene library preparation. Nucleic acid extractions were performed on water filters using a modified cetyltrimethylammonium bromide (CTAB) phenol-chloroform-isoamyl alcohol method, as described by Hazen et al. (42). The resulting pellet was resuspended in buffer EB (Qiagen, Germantown, MD), and the DNA was then subjected to the AllPrep DNA/RNA minikit (Qiagen), using the manufacturer's recommended protocol. DNA extracts were quantified using the Qubit 2.0 fluorometer double-stranded DNA (dsDNA) high-sensitivity DNA kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions and stored at -80°C .

Duplicate 25- μ l Illumina tag PCR mixtures from each sample contained final concentrations of 1 $\times$ PCR buffer, 0.8 mM dinucleoside triphosphates (dNTPs), 0.625 U of *Taq* polymerase, 0.2 μ M 515F forward primer, 0.2 μ M Illumina 806R reverse barcoded primer, and ~ 10 ng of template DNA per reaction. PCR was performed on an MJ Research PTC-200 thermocycler (Bio-Rad, Hercules, CA) using cycling conditions of 94°C for 3 min, followed by 35 cycles of 94°C for 45 s, 53°C for 60 s, and 72°C for 90 s, and ending with 72°C for 4 min; after, it was kept at 4°C. PCR products were visualized on a 2% agarose E-Gel (Invitrogen, Carlsbad, CA) stained with ethidium bromide. Positive products were pooled and purified with SPRI beads (Agencourt Bioscience Corporation, Beverly, MA) according to the manufacturer's instructions. Purified pools were then analyzed on the Agilent Bioanalyzer using a high-sensitivity

DNA kit (Agilent Technologies, Santa Clara, CA). Pooled libraries were stored at -20°C before transportation on dry ice for sequencing at the Children's Hospital DNA Sequencing Core (Cincinnati, OH).

Library pools were size verified using the Fragment Analyzer CE (Advanced Analytical Technologies, Inc., Ames, IA) and quantified using the Qubit high-sensitivity dsDNA kit (Life Technologies, Carlsbad, CA). Pools were diluted to a final concentration of 1 nM and a 10% spike of a PhiX V3 library (Illumina, San Diego, CA) was added, denatured for 5 min in an equal volume of 0.1 N NaOH, and further diluted to 12 pM in Illumina's HT1 buffer. The denatured and PhiX-spiked 12 pM pool was loaded on an Illumina MiSeq V2 500-cycle kit cassette with 16S rRNA library sequencing primers and set for 251-base paired-end reads.

Bioinformatics and statistical analyses. Sequence data for this project can be found in the NCBI Sequence Read Archive (SRP070501). Sequences were paired with a minimum overlap of 200 bp, trimmed at a length of 253 bp, and quality filtered at an expected error of $<1\%$ using USEARCH version 7 (43). Samples with a minimum of 5,000 reads were retained, resulting in 2,105,114 total sequences, encompassing 22,237 unique operational taxonomic units (OTUs) (97%). The sample size was reduced to 32 samples, and sequence data from day 4 did not pass quality filtering. Baseline microbial samples collected from 8 to 11 October 2015 ($n = 9$) underwent the same quality filtering, resulting in 279,215 sequences. Quality-filtered reads were analyzed using the QIIME 1.9.0 software package (44). OTUs were picked *de novo* using the UPARSE algorithm, and singleton reads were discarded, as recommended by Edgar (45). Taxonomy assignment was performed using the RDP Classifier and Greengenes 16S rRNA gene database (13-5 release) (a cluster of reads with 97% similarity was defined as an OTU) (46, 47).

Alpha-diversity multiple rarefactions were conducted using QIIME 1.9.0 on sequences across all samples from minimum depth of 100 sequences to a maximum depth of 5,000 sequences, with a step size of 500 sequences per sample for 30 iterations. Alpha rarefactions were collated using phylogenetic distance (PD) whole tree, Heip's evenness, Chao1, and observed species richness metrics. Alpha-diversity comparisons were conducted between bacterial communities corresponding to each sampling day and cumulative time using a nonparametric two-sample *t* test and nonparametric Monte Carlo permutations ($n = 999$). Comparisons between baseline and storm samples were also examined using nonparametric Monte Carlo permutations ($n = 999$). Visualization of trends in the alpha diversity of water samples was generated in R using the phyloseq package version 1.12.2 (48, 49).

Beta diversity was characterized with weighted UniFrac distances calculated between storm samples ($n = 32$), as well as with binary Jaccard indexes between storm and baseline ($n = 9$) data. OTU tables underwent cumulative sum scaling (CSS) normalization. Principal-coordinate analysis (PCoA) plots were generated in QIIME 1.9.0 to visualize the change in bacterial community structure during the storm and also to compare storm data to those of baseline bacterial communities. Adonis tests were performed on weighted UniFrac values to determine the significance of variation explained by cumulative time and environmental conditions. All statistical analyses were considered significant at an α value of 0.05.

Core microbiome analysis was visualized using a Venn diagram generated with Venny 2.0.2 (50) to reveal the number of unique and shared OTUs between the sampling days. Individual trends in OTU abundance with time were determined using Spearman rank correlations generated with an OTU normalized using phyloseq in R (47, 48). OTUs and meta-data categories with an R^2 value of >0.8 or less than -0.8 were retained. Correlations between OTUs were calculated with SparCC on log-transformed relative abundances using a bootstrap procedure and correlation threshold value of 0.3, as recommended by Friedman and Alm (51). Sample distances were computed with weighted UniFrac distances between day 1 water samples and all other samples. Visualization of trends in the relative abundance of taxa corresponding to sampling day was generated with an OTU table filtered to remove OTUs with $<0.005\%$ abundance ($n = 27$), as recommended by Bokulich et al. (52).

SourceTracker was used to investigate the presence of potential fecal contamination in the Muddy Run Creek water samples and to determine whether the source(s) was from human fecal contamination (53). Sequences from human fecal contamination sources included sewage influent ($n = 40$), activated sludge ($n = 40$), human stool ($n = 9$), and raw sewage ($n = 2$). Sequence data for sources studied were obtained through the NCBI Sequence Read Archive (SRA) from the projects PRJEB4688, PRJNA260846, PRJNA264400, PRJEB8668, and PRJNA292470 (<http://www.ncbi.nlm.nih.gov/sra>). These source sequences were chosen, as these samples were previously amplified for the 16S rRNA gene (V4 region) and sequenced on an Illumina MiSeq platform, consistent with library preparation and sequencing of the water samples. All source sequence data underwent the same filtering and quality measures as sequences from our study.

Nucleotide sequence accession number. Sequence data for this project were deposited in the NCBI Sequence Read Archive under accession number SRP070501.

RESULTS

Stream water chemistry during the storm event. Stream chemistry during the storm event covaried with time. Water chemistry measures, including conductivity, pH, salinity, and total dissolved solids (TDS), exhibited strong positive correlations with time ($r^2 > 0.71$, $P < 0.004$), while stream temperature correlated negatively with time ($r^2 = 0.75$, $P < 0.01$) (see Fig. S2 in the supplemental material). Because all abiotic measurements showed strong covariance with time and with other measures of water quality, a dissection of the source of variation related to changes in bacterial communities was prevented (see Fig. S2). Therefore, subsequent data analyses evaluated bacterial community variation with respect to time, rather than any individual or combination of metrics. However, the characteristics that were identified as covariants with time likely accompany most storm events, making time an effective proxy for tracking variation during the storm.

***E. coli* and total coliform concentrations for water quality assessment.** A hydrograph from 27 October to 3 November 2012 was created using data from the nearest USGS station (Juniata River at Huntingdon, PA) and shows that the peak discharge rate ($270 \text{ m}^3/\text{s}$) occurred on 30 October 2012 (day 2, 14.9 h) (Fig. 1). Muddy Run Creek was sampled at regular time points until the flow rate returned to baseline flow ($<50 \text{ m}^3/\text{s}$) on 3 November 2012 (day 5). Microbiological water quality was assessed during and after the storm event by measuring total fecal coliform and *E. coli* concentrations. Fecal coliforms ranged from 167 to 607 most probable number (MPN) per 100 ml over the course of sampling; however, the highest fecal concentration (607 MPN per 100 ml) was measured just after peak discharge rate (day 2, 17.78 h) (Fig. 1). *E. coli* concentrations at the first sampling point (day 1, 0.00 h) were 160 MPN per 100 ml and increased to the highest measured concentration (514 MPN per 100 ml) at the peak of the storm (day 2, 14.9 h) (Fig. 1). As the flow rate decreased after the storm peak, the *E. coli* concentration decreased (110 MPN per 100 ml) and stabilized over the next few time points to 225 MPN per 100 ml (day 2, 23.92 h; day 5, 115.28 h).

Bacterial community diversity. Alpha rarefaction analysis revealed that the observed number of OTUs changed significantly throughout the sampling period. The alpha diversity of the bacterial community structure dramatically changed over the course of the storm, as samples of days 1 and 2 compared to those of the final time points were significantly different (observed species, $P < 0.042$; Chao1, $P < 0.006$; phylogenetic distance, $P < 0.051$; Heip's

FIG 1 Hydrograph and microbiological water quality during sampling period. The flow rate (m^3/s) obtained from USGS station 01559000 Juniata River at Huntingdon, PA ($40^\circ 29' 05''\text{N}$, $78^\circ 01' 09''\text{W}$) and microbiological water quality data during Superstorm Sandy sampling period are shown. Sampling began on 29 October 2012 (day 1, 0.00 h), and subsequent time points are as follows, 30 October 2012 (day 2, 11.42 h, 14.54 h, 17.47 h, 21.07 h, and 23.55 h), 31 October 2012 (day 3, 42.3 h and 46.78 h), and 3 November 2012 (day 5, 115.28 h). Geometric mean densities of *E. coli* and fecal coliform derived from Colilert kits are reported as most probable numbers (MPN) for each sampling time point. *E. coli* and fecal coliform concentrations exhibited a pattern similar to that of the Muddy Run flow rate during the course of sampling.

evenness, $P < 0.033$). The observed richness and evenness were considerably higher during the peak flow rate than that at the final time points (Fig. 2). The number of observed species for day 1 samples was $1,961 \pm 197$ OTUs, while water samples from day 5 had an observed richness of $1,351 \pm 49$ OTUs at an even sampling depth. Differences in diversity between storm and baseline water samples revealed significantly less-abundant bacterial communities in baseline flow water samples ($P < 0.001$) (see Fig. S3 in the supplemental material).

Beta-diversity analyses displayed that the bacterial community composition changed over the course of the storm event. For example, principal-coordinate analysis (PCoA) showed distinct changes in bacterial community composition with respect to time (Adonis cumulative time, $r^2 = 0.63$, $F = 0.001$) (Fig. 3). Samples from the earlier time points clustered furthest right on axis 1, while samples taken during later time points were located to the left on axis 1 (Fig. 3). A total of 29.03% variation in the bacterial community was explained by axis 1, indicating that time shared a strong relationship with bacterial community structure in the water samples. Water samples also clustered by flow rate (Adonis flow rate, $r^2 = 0.79$, $F = 0.001$); however, initial and final storm samples were significantly different in bacterial composition (see Fig. S4 in the supplemental material). Baseline samples collected during seasonally matched time points were significantly different than storm samples (Adonis condition, $r^2 = 0.49$, $F = 0.001$); however, binary data of Jaccard indices at the family level revealed that baseline water samples were most similar to samples obtained during the final time points of the storm (see Fig. S5 in the supplemental material). Weighted UniFrac distances of averaged day 1 water samples compared to all other water samples acquired during the storm converged with baseline samples at the day 5 time points (Fig. 4). Day 1 water samples were most distant from day 5 samples (0.36) and baseline water samples (0.40) (Fig. 4).

While overall trends in beta diversity indicated shifts in the aquatic bacterial communities with respect to time, further investigation was performed to examine the unique and core OTUs within the water samples over the duration of the storm. Day 1

water samples were most abundant in unique OTUs (439 OTUs) compared to samples from the other sampling days (Fig. 5). Over the course of the storm, the number of unique taxa for each sampling day decreased (day 2, 121 unique OTUs; day 3, 10 unique OTUs). By day 5, the amount of unique OTUs increased to 175 in the water samples (Fig. 5). The initial and final sampling time points shared only 126 OTUs and were significantly different (nonparametric two-sample t test, $P = 0.048$). The bacterial communities within the day 2 and day 5 water samples were also significantly different (nonparametric two-sample t test, $P = 0.006$). Taxa abundant within all four sampling days (337 OTUs) were composed mainly of proteobacterial sequences (53%).

Longitudinal fluctuations in the presence and relative abundance of bacterial phyla were observed. The *Proteobacteria* dominated the stream bacterial community during days 1 and 2, comprising 62% of all 16S rRNA gene sequences (Fig. 6). By day 5, however, the relative abundance of proteobacterial sequences decreased to an average of 30% of the total bacterial sequences, with the exception at the 23.92-h time point on day 2 (61%) (Fig. 6). The relative abundance of *Firmicutes* sequences increased over the sampling period from an average relative abundance of 1.5% (day 1, 0.00 h) to 19% (day 5, 115.28 h) (Fig. 6). The *Bacilli* dominated the *Firmicutes* for all sampling days (see Fig. S6A in the supplemental material). *Bacilli* increased most significantly in abundance (29.4%) between the initial and final time points, with a peak in abundance at the 21.12-h time point (day 2) to 23% average relative abundance (see Fig. S6B in the supplemental material). The abundance of candidate phylum OD1 increased from approximately 1% (day 1, 0.00 h) to 13% (day 5, 115.28 h) (Fig. 5). In contrast, the relative abundance of *Verrucomicrobia* decreased from 5% (day 1, 0.00 h) to 1% (day 5, 115.28 h) (Fig. 5). Sequences belonging to the *Bacteroidetes* phylum were initially dominated by *Sphingobacteriia*, followed by an increase in *Cytophagia* (8.7%) during the 42.3-h (day 2) and 46.78-h (day 3) time points (see Fig. S6C in the supplemental material).

Betaproteobacteria and *Gammaproteobacteria* fluctuated in abundance over the course of the storm. For example, *Betaproteo-*

FIG 2 Alpha diversity of stream microbial communities during and after the storm event. Species richness was estimated by performing multiple rarefactions up to a depth of 5,600 sequences per sample ($n = 36$). The richness of OTUs from the maximum rarefaction depth was calculated using observed richness, Chao1, and abundance-based coverage estimator (ACE) and visualized using phyloseq in R. For each metric, species richness is colored by flow rate and separated by sampling day. Samples from days 1 and 2 possess higher α -diversity than that of samples taken after the peak of the storm.

bacteria were highest (49%) following the time of peak discharge rate (23.92 h) and then decreased in relative abundance (21%) by the 46.78-h time point (day 3) (Fig. 7A). *Gammaproteobacteria* decreased over time; however, at 23.92 h (day 2), *Gammaproteobacteria* exhibited a 6% increase in abundance before remaining at 10% relative abundance during the final three time points. Within the *Gammaproteobacteria* class, the *Pseudomonadaceae* and the *Enterobacteriaceae* families decreased throughout the sampling period (Fig. 7B). The *Pseudomonadaceae* abundance decreased considerably from 12% (day 1, 0.00 h) to 5% (day 2, 11.7 h), and by day 5 (115.28 h), *Pseudomonadaceae* comprised only 1.2% of the *Gammaproteobacteria* sequences (Fig. 7B). Similarly, the *Enterobacteriaceae* decreased 5% in average abundance between the first two time points of sampling before decreasing to 0.9% (day 5, 115.28 h). Baseline water samples revealed markedly different relative abundances of bacterial community members in comparison to samples acquired during the storm (see Fig. S7 in the supplemental material). *Proteobacteria* dominated baseline water samples, with the *Actinobacteria* phylum being second highest in abundance. *Firmicutes* were <10% abundant in all baseline water samples (see Fig. S7).

The relative abundance of several dominant bacterial taxa shared significant correlations with time. For example, 24 taxa, including specific taxa within the *Myxococcales* and *Pseudomonadales* orders, were found to have a strong negative correlation with time (Spearman's $\rho = -0.81$ to -0.92), while 30 taxa within the *Betaproteobacteria* and *Bacilli* classes were found to have a strong positive correlation with time (Spearman's $\rho = 0.8$ to 0.94) (see Fig. S8 in the supplemental material). Taxa associated with fecal waste showed an increase in abundance over the sampling period, such as sequences matching to the *Bacteroides* (Spearman's $\rho = 0.83$) (see Fig. S9 in the supplemental material). *Bacteroidetes* significantly correlated with members of the *Myxococcales* order ($P < 0.001$), in addition to certain taxa matching the *Clostridium* genus ($P < 0.001$). Furthermore, sequences matching to potentially pathogenic populations fluctuated in abundance within the water samples throughout the storm. 16S rRNA gene sequences matching the *Legionella* genus gradually increased throughout the sampling period to become the most abundant genus, comprising 46% of the bacterial sequences within the *Legionellaceae* family. Similarly, members of the *Campylobacteraceae* family showed patterns of temporal variation,

FIG 3 Beta diversity of stream microbial communities during and after Superstorm Sandy. Principal-coordinate analysis (PCoA) plot generated using weighted UniFrac distances of samples during and after Superstorm Sandy from a cumulative sum scaling (CSS)-normalized OTU table. Initial samples (red) cluster far right on PC1, and final samples (gray) are displayed farthest to the left. Data are from day 1 (0.00 h), day 2 (11.7 h, 14.9 h, 17.78 h, 21.12 h, and 23.92 h), day 3 (42.3 h and 46.78 h), and day 5 (115.28 h). Distinct clustering can be observed between water samples of different time points. The microbial community composition appears to undergo marked shifts during the storm event.

FIG 4 Sample distances of bacterial communities within storm and baseline samples. Shown are weighted UniFrac distances of averaged day 1 water samples ($n = 5$) compared to all other samples collected during the storm. A CSS-normalized OTU table of storm and baseline samples was used to calculate sample distance. The error bars show standard error of the mean values. Weighted UniFrac distances between initial and final initial samples converge with baseline samples, revealing a stabilization of bacterial community structure.

with *Arcobacter* having the highest relative abundance (63% of *Campylobacteraceae* sequences), followed by *Campylobacter*. The relative abundance of 16S rRNA gene sequences matching both genera increased even as the discharge rate decreased to baseline flow conditions. *Helicobacter* increased slightly to 3.7% of the *Helicobacteraceae* sequences by the final sampling time point (see Fig. S10 in the supplemental material).

SourceTracker analyses suggested that human-associated waste might be a potential source of fecal contamination in this study. Water samples exhibited a strong relationship with sewage influent sources (see Fig. S11 and Table S2 in the supplemental material). Other sources of human waste contamination, including active sludge, human stool, and raw sewage, indicated a minor source relationship. As much as 15% of the bacterial sequences from Muddy Run Creek were potentially associated with sources of human waste contamination (see Fig. S11). Sources classified as unknown by SourceTracker were likely due to other environmental, agricultural, or industrial sources.

DISCUSSION

The impact of storm events on bacterial community dynamics in streams has not been well studied. Here, we captured fine-resolution short-term temporal snapshots of the bacterial community during a 100-year storm event. Superstorm Sandy provided an opportunity to track changes in bacterial community structure through the progression of the storm and to monitor fluctuations in fecal indicators and specific taxa of interest. Furthermore, we investigated the relationship between successional changes in bac-

FIG 5 Core microbiome analysis between different sampling days during the storm event. A core microbiome diagram produced with Venny (50) illustrates the overlapping OTUs between sampling days based on OTU presence or absence in 100% of the water samples for each corresponding sampling day. An OTU table filtered to remove OTUs with $<0.005\%$ abundance and normalized for sequence depth with cumulative sum scaling (CSS) was used to generate the Venn diagram. Day 1 has significantly more unique taxa than the other sampling days.

terial community composition and physicochemical parameters. This is the first study to use high-throughput sequencing of the 16S rRNA gene, which revealed that stream bacterial communities may be restructured by hydric dynamics of storm events.

Bacterial communities within water samples collected from Muddy Run Creek exhibited patterns of temporal variation through the duration of Superstorm Sandy. Changes at the phylum level indicated a dramatic shift in the overall composition of the bacterial community, potentially due to contributions from sediment resuspension, stormwater runoff, and sewer overflow. Alpha-diversity metrics revealed that samples collected during a high flow rate ($>200 \text{ m}^3/\text{s}$) were the most diverse, as samples taken before and immediately following peak discharge harbored significantly more OTUs than samples taken under baseline flow conditions (Fig. 2; see also Fig. S3 in the supplemental material). Amount of rainfall has been shown to significantly correlate with bacterial influx (54–56), and greater bacterial abundance is related to extent of rainfall due to possible sediment resuspension (36). The presence of unique taxa was highest during the rising limb of the hydrograph (day 1 samples) and decreased after the storm (Fig. 5), suggesting the hydric dynamics of the storm generated a significant bacterial influx to the stream. An increase in nutrient load as a result of terrestrial, agricultural, and urban stormwater

FIG 6 Phylum abundance in water samples during the sampling period. Relative abundance of 16S rRNA gene sequences showing the seven most common phyla plotted by sample are shown. Plots were generated from an unrarified OTU table picked using the USEARCH sequence analysis tool. Day and time of sampling are displayed on the x axis. Cumulative time points are as follows: day 1 (0.00 h), day 2 (11.7 h, 14.9 h, 17.17 h, 21.12 h, and 23.92 h), day 3 (42.3 h and 46.78 h), and day 5 (115.28 h). The y axis displays the relative abundance of taxa per sample. Taxa with an average relative abundance of $<0.5\%$ across all samples were grouped together into the “other” category. *Proteobacteria* is the most abundant phylum in the water samples and decreased in abundance over the course of sampling.

FIG 7 Relative abundance of the *Proteobacteria* phylum during and after the storm event. Relative abundance of 16S rRNA gene sequences for five classes within the *Proteobacteria* phylum (A) and three families within the *Gammaproteobacteria* class (B) are represented in the water samples ($n = 27$). The plots were generated using an unrarified OTU table and display the cumulative time points along the x axis (in hours [hrs]) and average relative abundance along the y axis. The number of replicates included in each and corresponding cumulative time point is also shown. Black bars display the standard error of the mean values for each cumulative time point. Fluctuations in the relative abundance of proteobacterial classes are prevalent throughout the sampling period.

inputs has been shown to affect bacterial abundance and diversity in stream water (57, 58).

Longitudinal changes in bacterial community structure were observed over the course of the storm. Clustering of samples was noted based on cumulative time and flow rate, indicating that a distinct transformation in bacterial presence and abundance occurred as a result of the storm. Water samples collected during the time of highest flow rate (day 2) distinctly clustered together, and the greatest dissimilarity was present between the day 1 and 5 water samples (Fig. 3). Stormwater influx and sediment in free-

flowing water during storm progression were possible factors causing such a shift in bacterial community structure. Spatial and temporal analyses of natural disturbance effect on coastal bacterial communities have shown differences in bacterial communities between storm and nonstorm conditions (59). Stream water conductivity, pH, salinity, and TDS measurements increased as the storm ended and baseline flow rate returned, indicating that different abiotic characteristics could influence bacterial composition. The observed covariance of water chemistry with time was likely due to the strong environmental forcing of large storm

events, which increases the amount of freshwater runoff and introduction of allochthonous material into receiving water bodies. While reports suggest that streams recover from storms 3 to 8 days after the storm event, the bacterial community may recover more slowly (59). No prestorm samples were taken in this study, preventing a comparison of poststorm to prestorm stream bacterial communities. However, baseline samples taken during October 2015 allowed for tracking the divergence of microbial communities during and after the storm. While baseline samples displayed distinct clustering, binary Jaccard indices revealed that bacterial community structure of baseline samples was most similar to that of the day 5 water samples (see Fig. S5 in the supplemental material). Furthermore, a convergence to baseline samples was apparent in a comparison of weighted UniFrac distances of day 1 samples with both day 5 and baseline water samples (Fig. 4). While there were differences in the relative abundances of taxa (see Fig. S7 in the supplemental material), community membership showed that poststorm samples more closely matched baseline samples (Fig. 4), suggesting that that within a few days after the storm, the bacterial community was beginning to resemble that under baseline conditions (see Fig. S5).

Proteobacteria, more specifically *Betaproteobacteria* and *Gammaproteobacteria*, 16S rRNA gene sequences dominated the stream water samples, as previously observed in other riverine studies (60, 61). Interestingly, *Proteobacteria* underwent the most drastic shifts in abundance over the course of the storm. At the onset of sampling, before peak discharge, *Proteobacteria* dominated the bacterial community, but the bacterial composition shifted following peak discharge toward a composition dominated by candidate phylum OD1 and *Firmicutes* (Fig. 6). Within the *Proteobacteria*, the *Betaproteobacteria* underwent most drastic fluctuation during and after the storm event, while the *Gammaproteobacteria* decreased steadily over the sampling period. Previous studies have noted similar trends with respect to fluctuation of the *Proteobacteria* during rainfall events (59, 62). The *Proteobacteria* are important in nutrient cycling within freshwater ecosystems, and specific subclasses (e.g., *Betaproteobacteria*) consist of bacterioplankton, which has been shown to readily fluctuate with varied nutrient concentrations (63).

In this study, fecal coliforms and *E. coli* concentrations decreased immediately after peak discharge of the storm (Fig. 1). Stream fecal coliforms have been found to fluctuate temporally and spatially, with greater loads appearing when the rate of rainfall is highest (64). The FIB concentrations in this study support this trend. Pachepsky and Shelton (65) and McBride et al. (66) have shown FIB to spike prior to reaching peak flow rate. *E. coli* has been found in high abundance in rainwater and is one of the most abundant potentially pathogenic bacterial rainwater samples (67), suggesting that the increased abundance of *E. coli* at the beginning of the storm might be a result of direct contributions from stormwater influx. Salinity negatively affects FIB persistence (68), suggesting one possible mechanism for decreased abundance during later sample time points. Other inputs, including contributions from sewer overflows in Muddy Run Creek, might have influenced the elevated levels of fecal bacteria in this stream.

While traditional fecal indicators decreased during the sampling period, sequences matching to known fecal bacterial targets increased in relative abundance after the peak of the storm. This was supported by SourceTracker results, which indicated that potential human fecal contribution was highest after peak flow of the

storm (see Fig. S11 and Table S2 in the supplemental material). For example, 16S rRNA gene sequences belonging to the *Clostridium* and *Blautia* genera, which are known to be of fecal origin, had higher relative abundance after the storm than before (see Fig. S9 in the supplemental material). While the *Bacteroidetes* members fluctuated with irregularity, the *Bacteroides* genus, known to dominate human fecal and sewage material (69), increased dramatically in abundance during later time points (see Fig. S9). Several *Bacteroides* and *Clostridium* spp. have been shown to have host-specific distributions and are indicators of more recent fecal contamination events; thus, they have become molecular targets for tracking sources of fecal contamination in the environment (70–73). Comprehensive sequencing studies such as this enable the simultaneous detection of these fecal bacterial targets and might provide a more holistic understanding of fecal inputs into aquatic environments. Increases in these fecal bacterial sequences might be a result of inputs from a sewer overflow in Muddy Run Creek, which were initially diluted by the stormwater discharge. Sewer overflows have been shown to pose a serious threat to the health and quality of urban streams (74, 75).

Increases in fecal bacterial sequences (e.g., *Bacteroides*) were mirrored by increases in sequences belonging to potentially pathogenic taxa during the 5-day sampling period. Previous research has documented increases in potential pathogens during storm events (36, 65). While both pathogens and FIB are released from sewage discharge, studies have shown no clear correlation between pathogens and fecal indicators (20, 76, 77). In this study, 16S rRNA gene sequences belonging to the *Campylobacter*, *Arco-bacter*, and *Helicobacter* genera increased in relative abundance throughout the sampling period (see Fig. S10 in the supplemental material). *Legionella* spp. were also present within water samples, comprising almost half of the *Legionellaceae* sequences by day 5. Previous literature has revealed that *Legionella* spp. can survive for several days in water (78, 79). The sequence data indicate that fecal bacteria and potential pathogens are higher days after the peak of the storm, suggesting that microbial risks may persist long after a storm has ended. Putative pathogenic bacteria, specifically *Campylobacter*, *Helicobacter*, and *Legionella*, were ubiquitously present in baseline samples at negligible abundance. However, bacteria of fecal origin were not present in baseline water samples compared to poststorm microbial communities. Although high-throughput sequencing of the 16S rRNA gene enabled us to identify fecal bacteria and potential pathogens, this approach should be employed with caution, due to the limited phylogenetic resolution when using the 16S rRNA gene as a target.

This study captures beta fluctuations of a bacterial community structure during a 100-year storm event and has provided a simultaneous view of successional changes in total bacterial community structure, as well as an in-depth investigation of temporal dynamics of fecal bacteria and potential pathogens during a storm event. Use of the 16S rRNA gene as a genetic marker allows for holistic bacterial community assessment by using multiple indicators to assess microbiological water quality over large spatial and temporal scales (34, 39, 54). Tracking bacterial community composition might prove to be an informative tool for long-term temporal studies, as other factors, including seasonal changes and climate change, cause shifts in bacterial community dynamics and associated risks (19, 32, 62, 80). Further studies should examine beyond the 16S rRNA gene and probe functional genetic markers, such as genes involved in host-microbial interactions, which might prove

to be more temporally stable indicators for assessing microbial risks. With the reduced cost of high-throughput sequencing and improved sensitivity, shotgun metagenomics and metatranscriptomics approaches might enable the direct measurement of functional targets in the environment. Evaluation of *in situ* total bacterial community composition and function might lead to more comprehensive water quality assessments and enable an evaluation of the magnitude and distribution of microbial contamination in the environment.

ACKNOWLEDGMENTS

We thank Idexx for providing the Colilert enzymatic assays, and we also thank Mehdi Keddache for providing sequencing data.

This research was supported by a grant to Juniata College from the Howard Hughes Medical Institute (<http://www.hhmi.org>) through the Precollege and Undergraduate Science Education Program, the National Science Foundation (www.nsf.gov), NSF award DBI-1248096

FUNDING INFORMATION

This work, including the efforts of Regina Lamendella, was funded by National Science Foundation (NSF) (DBI-1248096). This work, including the efforts of Regina Lamendella, was funded by Howard Hughes Medical Institute (HHMI).

The funders had no role in the study design, data collection and interpretation, or the decision to submit the work for publication

REFERENCES

- United States Environmental Protection Agency (USEPA). 2009. Review of published studies to characterized relative risks from different sources of fecal contamination in recreational water. U.S. Environmental Protection Agency, Washington, DC. <https://www.deq.idaho.gov/media/910372-relative-risks-from-different-sources-fecal-contamination-recreational-water-2009.pdf>.
- Marsalek J, Rochfort Q. 2004. Urban wet-weather flows: sources of fecal contamination impacting on recreational waters and threatening drinking-water sources. *J Toxicol Environ Health Part A* 67:1765–1777. <http://dx.doi.org/10.1080/1528739049042430>.
- Reeves RL, Grant SB, Mrse RD, Copil Oancea CM, Sanders BF, Boehm AB. 2004. Scaling and management of fecal indicator bacteria in runoff from a coastal urban watershed in southern California. *Environ Sci Technol* 38:2637–2648. <http://dx.doi.org/10.1021/es034797g>.
- Surbeck CQ, Jiang SC, Ahn JH, Grant SB. 2006. Flow fingerprinting fecal pollution and suspended solids in stormwater runoff from an urban coastal watershed. *Environ Sci Technol* 40:4435–4441. <http://dx.doi.org/10.1021/es060701h>.
- Sidhu JPS, Ahmed W, Gernjak W, Aryal R, McCarthy D, Palmer A, Kolotelo P, Toze S. 2013. Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. *Sci Total Environ* 463–464:488–496.
- Geldreich EE. 1996. Pathogenic agents in freshwater resources. *Hydrol Process* 10:315–333. [http://dx.doi.org/10.1002/\(SICI\)1099-1085\(199602\)10:2<315::AID-HYP361>3.0.CO;2-H](http://dx.doi.org/10.1002/(SICI)1099-1085(199602)10:2<315::AID-HYP361>3.0.CO;2-H).
- Haile RW, Witte JS, Gold M, Cressey R, McGee C, Millikan RC, Glasser A, Harawa N, Ervin C, Harmon P, Harper J, Dermand J, Alamillo J, Barrett K, Nides M, Wang G. 1999. The health effects of swimming in ocean water contaminated by storm drain runoff. *Epidemiology* 10:355–363. <http://dx.doi.org/10.1097/00001648-199907000-00004>.
- Ferguson C, Husman AM, Altavilla N, Deere DA, Ashbolt NJ. 2003. Fate and transport of surface water pathogens in watersheds. *Crit Rev Environ Sci Technol* 33:299–361. <http://dx.doi.org/10.1080/10643380390814497>.
- Arnold RD, Walling JP. 2007. Waterborne pathogens in urban watersheds. *J Water Health* 5:149–162. <http://dx.doi.org/10.2166/wh.2006.001>.
- Davies TJ, Pedersen AB. 2008. Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proc Biol Sci* 275:1695–1701. <http://dx.doi.org/10.1098/rspb.2008.0284>.
- Viau EJ, Goodwin KD, Yamahara KM, Layton BA, Sassoubre LM, Burns SL, Tong HI, Wong SHC, Lu Y, Boehm AB. 2011. Bacterial pathogens in Hawaiian coastal streams—associations with fecal indicators, land cover, and water quality. *Water Res* 45:3279–3290. <http://dx.doi.org/10.1016/j.watres.2011.03.033>.
- ten Veldhuis JAE, Clemens FHRL, Sterk G, Berends BR. 2010. Microbial risks associated with exposure to pathogens in contaminated urban flood water. *Water Res* 44:2910–2918. <http://dx.doi.org/10.1016/j.watres.2010.02.009>.
- Teng J, Vaze J, Chiew FHS, Wang B, Perraud J-M. 2012. Estimating the relative uncertainties sourced from GCMs and hydrological models in modeling climate change impact on runoff. *J Hydrometeorol* 13:122–139. <http://dx.doi.org/10.1175/JHM-D-11-058.1>.
- Shapiro K, Miller WA, Silver MW, Odagiri M, Largier JL, Conrad PA, Mazet JAK. 2013. Research commentary: association of zoonotic pathogens with fresh, estuarine, and marine macroaggregates. *Microb Ecol* 65:928–933. <http://dx.doi.org/10.1007/s00248-012-0147-2>.
- de Man H, Bouwknecht M, van Heijnsbergen E, Leenen EJTM, van Knapen F, de Roda Husman AM. 2014. Health risk assessment for splash parks that use rainwater as source water. *Water Res* 54:254–261. <http://dx.doi.org/10.1016/j.watres.2014.02.010>.
- Puig M, Jofre J, Lucena F, Allard A, Wadell G, Girones R. 1994. Detection of adenoviruses and enteroviruses in polluted waters by nested PCR amplification. *Appl Environ Microbiol* 60:2963–2970.
- Lipp EK, Farrah SA, Rose JB. 2001. Assessment and impact of microbial fecal pollution and human enteric pathogens in a coastal community. *Mar Pollut Bull* 42:286–293. [http://dx.doi.org/10.1016/S0025-326X\(00\)00152-1](http://dx.doi.org/10.1016/S0025-326X(00)00152-1).
- Katukiza AY, Ronteltap M, van der Steen P, Foppen JWA, Lens PNL. 2014. Quantification of microbial risks to human health caused by waterborne viruses and bacteria in an urban slum. *J Appl Microbiol* 116:447–463. <http://dx.doi.org/10.1111/jam.12368>.
- Hofstra N. 2011. Quantifying the impact of climate change on enteric waterborne pathogen concentrations in surface water. *Curr Opin Environ Sustain* 3:471–479. <http://dx.doi.org/10.1016/j.cosust.2011.10.006>.
- Ahmed W, Goonetilleke A, Gardner T. 2010. Implications of faecal indicator bacteria for the microbiological assessment of roof-harvested rainwater quality in southeast Queensland, Australia. *Can J Microbiol* 56:471–479. <http://dx.doi.org/10.1139/W10-037>.
- Schets FM, Italiaander R, van den Berg HHJL, de Roda Husman AM. 2010. Rainwater harvesting: quality assessment and utilization in The Netherlands. *J Water Health* 8:224–235. <http://dx.doi.org/10.2166/wh.2009.037>.
- Schalk JAC, Docters van Leeuwen AE, Lodder WJ, de Man H, Euser S, den Boer JW, de Roda Husman AM. 2012. Isolation of *Legionella pneumophila* from pluvial floods by amoebal coculture. *Appl Environ Microbiol* 78:4519–4521. <http://dx.doi.org/10.1128/AEM.00131-12>.
- Dorner SM, Anderson WB, Gaulin T, Candon HL, Slawson RM, Payment P, Huck PM. 2007. Pathogen and indicator variability in a heavily impacted watershed. *J Water Health* 5:599. <http://dx.doi.org/10.2166/wh.2007.010>.
- Rowny JG, Stewart JR. 2012. Characterization of nonpoint source microbial contamination in an urbanizing watershed serving as a municipal water supply. *Water Res* 46:6143–6153. <http://dx.doi.org/10.1016/j.watres.2012.09.009>.
- Crabill C, Donald R, Snelling J, Foust R, Southam G. 1999. The impact of sediment fecal coliform reservoirs on seasonal water quality in Oak Creek, Arizona. *Water Res* 33:2163–2171. [http://dx.doi.org/10.1016/S0043-1354\(98\)00437-0](http://dx.doi.org/10.1016/S0043-1354(98)00437-0).
- Simpson JM, Santo Domingo JW, Reasoner DJ. 2002. Microbial source tracking: state of the science. *Environ Sci Technol* 36:5279–5288. <http://dx.doi.org/10.1021/es026000b>.
- Roy AH, Wenger SJ, Fletcher TD, Walsh CJ, Ladson AR, Shuster WD, Thurston HW, Brown RR. 2008. Impediments and solutions to sustainable, watershed-scale urban stormwater management: lessons from Australia and the United States. *Environ Manag* 42:344–359. <http://dx.doi.org/10.1007/s00267-008-9119-1>.
- Ashbolt NJ, Schoen ME, Soller JA, Roser DJ. 2010. Predicting pathogen risks to aid beach management: the real value of quantitative microbial risk assessment (QMRA). *Water Res* 44:4692–4703. <http://dx.doi.org/10.1016/j.watres.2010.06.048>.
- Allison SD, Martiny JBH. 2008. Resistance, resilience, and redundancy in microbial communities. *Proc Natl Acad Sci U S A* 105(Suppl 1):11512–11519. <http://dx.doi.org/10.1073/pnas.0801925105>.
- Logue JB, Bürgmann H, Robinson CT. 2008. Progress in the ecological

- genetics and biodiversity of freshwater bacteria. *Bioscience* 58:103–113. <http://dx.doi.org/10.1641/B580205>.
31. Wanjugi P, Harwood VJ. 2013. The influence of predation and competition on the survival of commensal and pathogenic fecal bacteria in aquatic habitats. *Environ Microbiol* 15:517–526. <http://dx.doi.org/10.1111/j.1462-2920.2012.02877.x>.
 32. Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B, Huse S, McHardy AC, Knight R, Joint I, Somerfield P, Fuhrman JA, Field D. 2012. Defining seasonal marine microbial community dynamics. *ISME J* 6:298–308. <http://dx.doi.org/10.1038/ismej.2011.107>.
 33. Afshinnikoo E, Meydan C, Chowdhury S, Jaroudi D, Boyer C, Bernstein N, Maritz JM, Reeves D, Gandara J, Chhangawala S, Ahsanuddin S, Simmons A, Nessel T, Sundaresh B, Pereira E, Jorgensen E, Kolokotronis SO, Kirchberger N, Garcia I, Gandara D, Dhanraj S, Nawrin T, Saletore Y, Alexander N, Vijay P, Hénaff EM, Zumbo P, Walsh M, O'Mullan GD, Tighe S, Dudley JT, Dunaif A, Ennis S, O'Halloran E, Magalhaes TR, Boone B, Jones AL, Muth TR, Paolantonio KS, Alter E, Schadt EE, Garbarino J, Prill RJ, Carlton JM, Levy S, Mason CE. 2015. Geospatial resolution of human and bacterial diversity with city-scale metagenomics. *Cell Syst* 1:72–87. <http://dx.doi.org/10.1016/j.cels.2015.01.001>.
 34. Unno T, Jang J, Han D, Kim JH, Sadowsky MJ, Kim OS, Chun J, Hur HG. 2010. Use of barcoded pyrosequencing and shared OTUs to determine sources of fecal bacteria in watersheds. *Environ Sci Technol* 44:7777–7782. <http://dx.doi.org/10.1021/es101500z>.
 35. Newton RJ, Bootsma MJ, Morrison HG, Sogin ML, McLellan SL. 2013. A microbial signature approach to identify fecal pollution in the waters off an urbanized coast of Lake Michigan. *Microb Ecol* 65:1011–1023. <http://dx.doi.org/10.1007/s00248-013-0200-9>.
 36. Noble RT, Griffith JF, Blackwood AD, Fuhrman JA, Gregory JB, Hernandez X, Liang X, Bera AA, Schiff K. 2006. Multitiered approach using quantitative PCR to track sources of fecal pollution affecting Santa Monica Bay, California. *Appl Environ Microbiol* 72:1604–1612. <http://dx.doi.org/10.1128/AEM.72.2.1604-1612.2006>.
 37. Staley C, Unno T, Gould TJ, Jarvis B, Phillips J, Cotner JB, Sadowsky MJ. 2013. Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *J Appl Microbiol* 115:1147–1158. <http://dx.doi.org/10.1111/jam.12323>.
 38. Staley C, Gould TJ, Wang P, Phillips J, Cotner JB, Sadowsky MJ. 2014. Bacterial community structure is indicative of chemical inputs in the Upper Mississippi River. *Front Microbiol* 5:524. <http://dx.doi.org/10.3389/fmicb.2014.00524>.
 39. McLellan SL, Eren AM. 2014. Discovering new indicators of fecal pollution. *Trends Microbiol* 22:697–706. <http://dx.doi.org/10.1016/j.tim.2014.08.002>.
 40. Kunz M, Mühr B, Kunz-Plapp T, Daniell JE, Khazai B, Wenzel F, Vannieuwenhuysse M, Comes T, Elmer F, Schröter K, Fohringer J, Münzberg T, Lucas C, Zschau J. 2013. Investigation of superstorm Sandy 2012 in a multi-disciplinary approach. *Nat Hazards Earth Syst Sci* 13:2579–2598. <http://dx.doi.org/10.5194/nhess-13-2579-2013>.
 41. National Aeronautics and Space Administration (NASA). 2013. Hurricane Sandy (Atlantic Ocean). National Aeronautics and Space Administration, Washington, DC. http://www.nasa.gov/mission_pages/hurricanes/archives/2012/h2012_Sandy.html.
 42. Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N, Jansson JK, Probst A, Borglin SE, Fortney JL, Stringfellow WT, Bill M, Conrad ME, Tom LM, Chavarria KL, Alusi TR, Lamendella R, Joyner DC, Spier C, Baelum J, Auer M, Zemla ML, Chakraborty R, Sonnenthal EL, D'haeseleer P, Holman H-YN, Osman S, Lu Z, Van Nostrand JD, Deng Y, Zhou J, Mason OU. 2010. Deep-sea oil plume enriches oil-degrading bacteria. *Science* 330:204–208. <http://dx.doi.org/10.1126/science.1195979>.
 43. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461. <http://dx.doi.org/10.1093/bioinformatics/btq461>.
 44. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <http://dx.doi.org/10.1038/nmeth.f.303>.
 45. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998. <http://dx.doi.org/10.1038/nmeth.2604>.
 46. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–5072. <http://dx.doi.org/10.1128/AEM.03006-05>.
 47. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267. <http://dx.doi.org/10.1128/AEM.00062-07>.
 48. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. <http://dx.doi.org/10.1371/journal.pone.001217>.
 49. R Core Development Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
 50. Oliveros JC. 2007. VENN. An interactive tool for comparing lists with Venn Diagrams. BioinfoGP, CNB-CSIC, Madrid, Spain. <http://bioinfo.cnb.csic.es/tools/venny/>.
 51. Friedman J, Alm EJ. 2012. Inferring correlation networks from genomic survey data. *PLoS Comput Biol* 8:e1002687. <http://dx.doi.org/10.1371/journal.pcbi.1002687>.
 52. Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA, Caporaso JG. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 10:57–59. <http://dx.doi.org/10.1038/nmeth.2276>.
 53. Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R, Kelley ST. 2011. Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* 8:761–763. <http://dx.doi.org/10.1038/nmeth.1650>.
 54. Stumpf CH, Piehler MF, Thompson S, Noble RT. 2010. Loading of fecal indicator bacteria in North Carolina tidal creek headwaters: Hydrographic patterns and terrestrial runoff relationships. *Water Res* 44:4704–4715. <http://dx.doi.org/10.1016/j.watres.2010.07.004>.
 55. Cho KH, Pachepsky YA, Kim JH, Guber AK, Shelton DR, Rowland R. 2010. Release of *Escherichia coli* from the bottom sediment in a first-order creek: experiment and reach-specific modeling. *J Hydrol* 391:322–332. <http://dx.doi.org/10.1016/j.jhydrol.2010.07.033>.
 56. Cho KH, Cha SM, Kang JH, Lee SW, Park Y, Kim JW, Kim JH. 2010. Meteorological effects on the levels of fecal indicator bacteria in an urban stream: a modeling approach. *Water Res* 44:2189–2202. <http://dx.doi.org/10.1016/j.watres.2009.12.051>.
 57. Krometis LAH, Characklis GW, Simmons OD, III, Dilts MJ, Likirdopoulos CA, Sobsey MD. 2007. Intra-storm variability in microbial partitioning and microbial loading rates. *Water Res* 41:506–516. <http://dx.doi.org/10.1016/j.watres.2006.09.029>.
 58. Ibekwe AM, Leddy MB, Bold RM, Graves AK. 2012. Bacterial community composition in low-flowing river water with different sources of pollutants. *FEMS Microbiol Ecol* 79:155–166. <http://dx.doi.org/10.1111/j.1574-6941.2011.01205.x>.
 59. Yeo SK, Huggett MJ, Eiler A, Rappé MS. 2013. Coastal bacterioplankton community dynamics in response to a natural disturbance. *PLoS One* 8:e56207. <http://dx.doi.org/10.1371/journal.pone.0056207>.
 60. Glöckner FO, Fuchs BM, Amann R. 1999. Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence *in situ* hybridization. *Appl Environ Microbiol* 65:3721–3726.
 61. Poretsky R, Rodriguez-R LM, Luo C, Tsementzi D, Constantinidis KT. 2014. Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *PLoS One* 9:e93827. <http://dx.doi.org/10.1371/journal.pone.0093827>.
 62. Staley C, Gould TJ, Wang P, Phillips J, Cotner JB, Sadowsky MJ. 2015. Species sorting and seasonal dynamics primarily shape bacterial communities in the Upper Mississippi River. *Sci Total Environ* 505:435–445. <http://dx.doi.org/10.1016/j.scitotenv.2014.10.012>.
 63. Salcher MM, Posch T, Pernthaler J. 2013. *In situ* substrate preferences of abundant bacterioplankton populations in a prealpine freshwater lake. *ISME J* 7:896–907. <http://dx.doi.org/10.1038/ismej.2012.162>.
 64. Lewis DJ, Atwill ER, Lennox MS, Hou L, Karle B, Tate KW. 2005. Linking on-farm dairy management practices to storm-flow fecal coliform loading for California coastal watersheds. *Environ Monit Assess* 107:407–425. <http://dx.doi.org/10.1007/s10661-005-3911-7>.
 65. Pachepsky YA, Shelton DR. 2011. *Escherichia coli* and fecal coliforms in

- freshwater and estuarine sediments. *Crit Rev Environ Sci Technol* 41: 1067–1110. <http://dx.doi.org/10.1080/10643380903392718>.
66. McBride G, Till D, Ryan T, Ball A, Lewis G, Palmer S, Weinstein P. 2002. Freshwater microbiology research programme report: pathogen occurrence and human health risk assessment analysis. Ministry for the Environment, Wellington, New Zealand. <https://www.mfe.govt.nz/sites/default/files/freshwater-microbiology-nov02.pdf>.
 67. Kaushik R, Balasubramanian R. 2012. Assessment of bacterial pathogens in fresh rainwater and airborne particulate matter using real-time PCR. *Atmos Environ* 46:131–139. <http://dx.doi.org/10.1016/j.atmosenv.2011.10.013>.
 68. Anderson KL, Whitlock JE, Harwood VJ. 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Appl Environ Microbiol* 71:3041–3048. <http://dx.doi.org/10.1128/AEM.71.6.3041-3048.2005>.
 69. Koskey AM, Fisher JC, Eren AM, Ponce-Terashima R, Reis MG, Blanton RE, McLellan SL. 2014. *Blautia* and *Prevotella* sequences distinguish human and animal fecal pollution in Brazil surface waters. *Environ Microbiol Rep* 6:696–704. <http://dx.doi.org/10.1111/1758-2229.12189>.
 70. Ahmed W, Huygens F, Goonetilleke A, Gardner T. 2008. Real-time PCR detection of pathogenic microorganisms in roof-harvested rainwater in Southeast Queensland, Australia. *Appl Environ Microbiol* 74:5490–5496. <http://dx.doi.org/10.1128/AEM.00331-08>.
 71. Dick LK, Bernhard AE, Brodeur TJ, Santo Domingo JW, Simpson JM, Walters SP, Field KG. 2005. Host distributions of uncultivated fecal *Bacteroidales* bacteria reveal genetic markers for fecal source identification. *Appl Environ Microbiol* 71:3184–3191. <http://dx.doi.org/10.1128/AEM.71.6.3184-3191.2005>.
 72. Field KG, Samadpour M. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. *Water Res* 41:3517–3538. <http://dx.doi.org/10.1016/j.watres.2007.06.056>.
 73. Vogel JR, Stoeckel DM, Lamendella R, Zelt RB, Santo Domingo JW, Walker SR, Oerther DB. 2007. Identifying fecal sources in a selected catchment reach using multiple source-tracking tools. *J Environ Qual* 36:718–729. <http://dx.doi.org/10.2134/jeq2006.0246>.
 74. Seager J, Abrahams RG. 1990. The impact of storm sewage discharges on the ecology of a small urban river. *Water Sci Technol* 22:163–171.
 75. Deng H. 2012. A review of diversity-stability relationship of soil microbial community: what do we not know? *J Environ Sci* 24:1027–1035. [http://dx.doi.org/10.1016/S1001-0742\(11\)60846-2](http://dx.doi.org/10.1016/S1001-0742(11)60846-2).
 76. Hörman A, Rimhanen-Finne R, Maunula L, Von Bonsdorff CH, Torvela N, Heikinheimo A, Hänninen ML. 2004. *Campylobacter* spp., *Giardia* spp., *Cryptosporidium* spp., noroviruses, and indicator organisms in surface water in southwestern Finland, 2000–2001. *Appl Environ Microbiol* 70:87–95. <http://dx.doi.org/10.1128/AEM.70.1.87-95.2004>.
 77. Savichtcheva O, Okabe S. 2006. Alternative indicators of fecal pollution: relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. *Water Res* 40:2463–2476. <http://dx.doi.org/10.1016/j.watres.2006.04.040>.
 78. Grimes DJ. 1991. Ecology of estuarine bacteria capable of causing human disease: a review. *Estuaries* 14:345–360. <http://dx.doi.org/10.2307/1352260>.
 79. Thomas C, Gibson H, Hill DJ, Mabey M. 1998. *Campylobacter* epidemiology: an aquatic perspective. *J Appl Microbiol* 85(Suppl 1):168S–177S.
 80. Hunter PR. 2003. Climate change and waterborne and vector-borne disease. *J Appl Microbiol* 94(Suppl):37S–46S.