



# Longitudinal Comparison of Bacterial Diversity and Antibiotic Resistance Genes in New York City Sewage

 Susan M. Joseph,<sup>a,b,\*</sup> Thomas Battaglia,<sup>a</sup> Julia M. Maritz,<sup>b</sup> Jane M. Carlton,<sup>b</sup> Martin J. Blaser<sup>a,c</sup>

<sup>a</sup>New York University School of Medicine, New York, New York, USA

<sup>b</sup>Center for Genomics and Systems Biology, Department of Biology, New York University, New York, New York, USA

<sup>c</sup>Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, New Jersey, USA

**ABSTRACT** Bacterial resistance to antibiotics is a pressing health issue around the world, not only in health care settings but also in the community and environment, particularly in crowded urban populations. The aim of our work was to characterize the microbial populations in sewage and the spread of antibiotic resistance within New York City (NYC). Here, we investigated the structure of the microbiome and the prevalence of antibiotic resistance genes in raw sewage samples collected from the fourteen NYC Department of Environmental Protection wastewater treatment plants, distributed across the five NYC boroughs. Sewage, a direct output of anthropogenic activity and a reservoir of microbes, provides an ecological niche to examine the spread of antibiotic resistance. Taxonomic diversity analysis revealed a largely similar and stable bacterial population structure across all the samples, which was found to be similar over three time points in an annual cycle, as well as in the five NYC boroughs. All samples were positive for the presence of the seven antibiotic resistance genes tested, based on real-time quantitative PCR assays, with higher levels observed for tetracycline resistance genes at all time points. For five of the seven genes, abundances were significantly higher in May than in February and August. This study provides characteristics of the NYC sewage resistome in the context of the overall bacterial populations.

**IMPORTANCE** Urban sewage or wastewater is a diverse source of bacterial growth, as well as a hot spot for the development of environmental antibiotic resistance, which can in turn influence the health of the residents of the city. As part of a larger study to characterize the urban New York City microbial metagenome, we collected raw sewage samples representing three seasonal time points spanning the five boroughs of NYC and went on to characterize the microbiome and the presence of a range of antibiotic resistance genes. Through this study, we have established a baseline microbial population and antibiotic resistance abundance in NYC sewage which can prove to be very useful in studying the load of antibiotic usage, as well as for developing effective measures in antibiotic stewardship.

**KEYWORDS** antibiotic resistance, microbiome, New York City, sewage

Antibiotic resistance (AR) in bacteria is the ability of a microorganism to survive in the presence of the inhibitory activity of antibiotics. Certain bacteria possess naturally occurring or intrinsic resistance, such as vancomycin resistance in strains of *Escherichia coli* or the presence of  $\beta$ -lactamases in *Streptomyces* (1). Alternatively, bacterial resistance to antibiotics may be acquired either via mutations that result in resistant survivors in a bacterial population (2) or via horizontal gene transfer through mechanisms including transformation or transfer of mobile elements such as plasmids, transposons (e.g., the *mecA* gene in methicillin-resistant *Staphylococcus aureus* [MRSA]) and bacteriophages (e.g., resistance transfer in strains of *S. aureus*) (3). This ability to


**Citation** Joseph SM, Battaglia T, Maritz JM, Carlton JM, Blaser MJ. 2019. Longitudinal comparison of bacterial diversity and antibiotic resistance genes in New York City sewage. *mSystems* 4:e00327-19. <https://doi.org/10.1128/mSystems.00327-19>.

**Editor** Zackery Bulman, University of Illinois at Chicago

**Copyright** © 2019 Joseph et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Susan M. Joseph, [susanj13@gmail.com](mailto:susanj13@gmail.com).

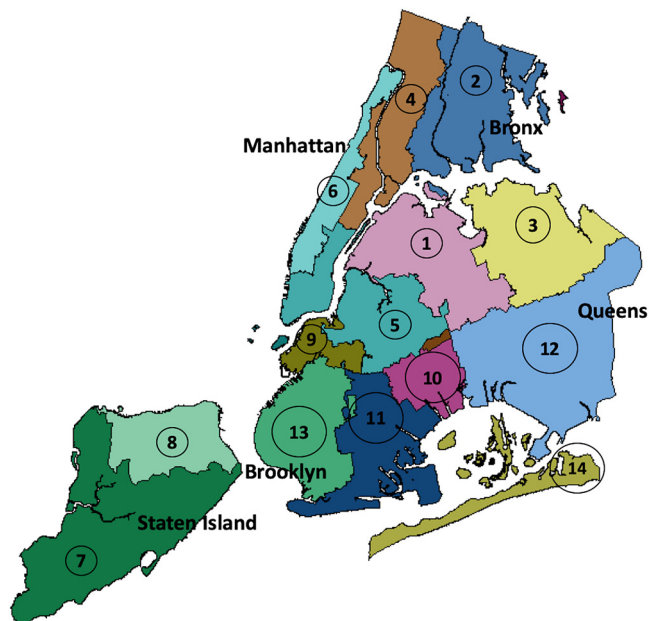
\* Present address: Susan M. Joseph, Centre for Host-Microbiome Interactions, Faculty of Dentistry, Oral and Craniofacial Sciences, King's College London, London, United Kingdom.

 Longitudinal study looking at bacterial diversity and antibiotic resistance gene abundance in the sewage of NYC boroughs

**Received** 18 May 2019

**Accepted** 19 July 2019

**Published** 6 August 2019



**FIG 1** Geographical distributions of NYC WWTPs. This New York City map indicates the drainage areas for the 14 WWTPs across the five NYC boroughs from which samples were obtained for this study. The map was plotted with the Maptools package in R using GIS data obtained from Open Sewer Atlas NYC (<http://openseweratlas.tumblr.com/data>).

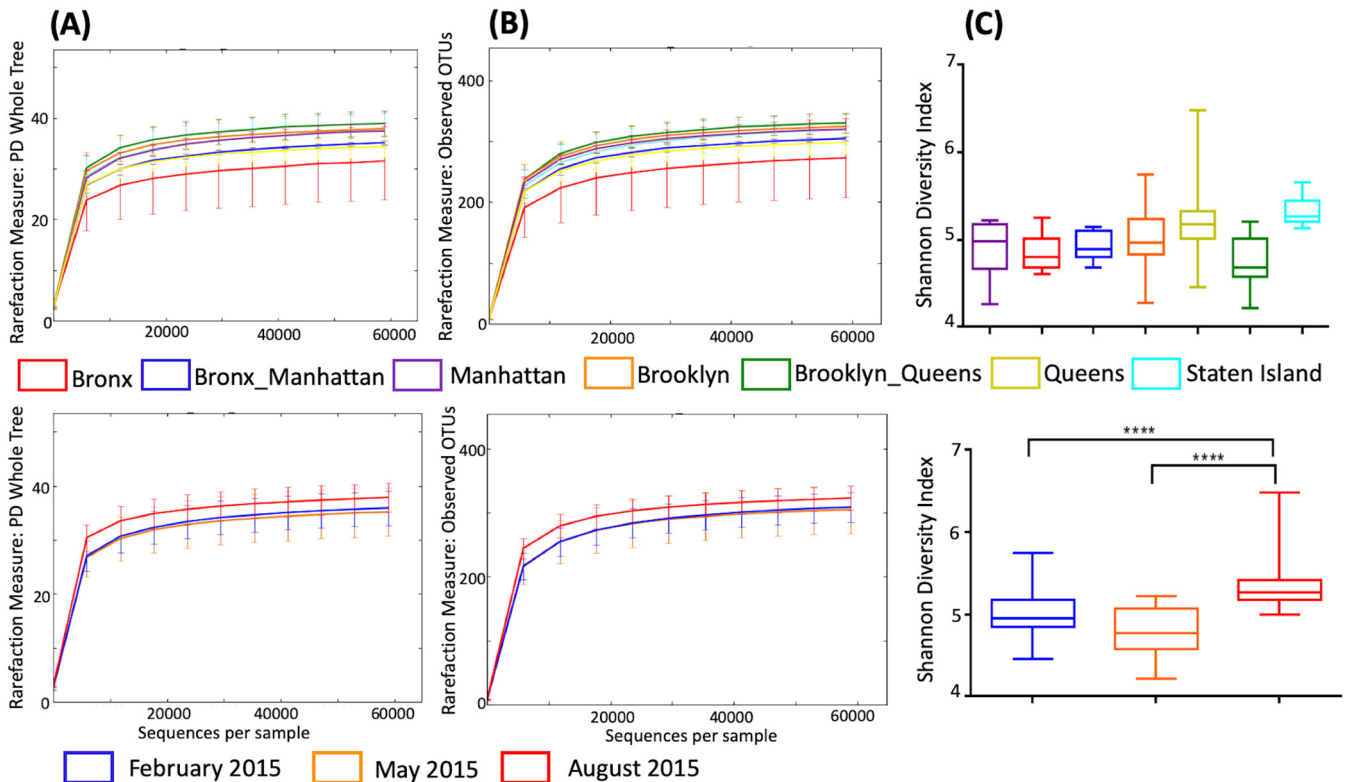
horizontally acquire resistance or resistance genes further magnifies the threats presented by AR.

Antibiotic resistance continues to grow and is now becoming a major global health crisis (4–6). Reported antibiotic resistance has led to excess mortality and failed treatment modes in hospitals, contributing to increased health care costs (7). These rising AR levels are fueled by the lack of rigid monitoring and control of the use of antibiotics in humans (8, 9) and in domesticated animals (10–12). Further exacerbating this global problem is that AR is not restricted to clinical settings but also occurs in both domestic and natural environments (13–16).

Raw sewage or wastewater is the product of many human activities in domestic, clinical, and industrial settings. Sewage represents a conglomerate of numerous microbes, including bacterial families found in human microbiota such as *Bacteroidaceae*, *Ruminococcaceae*, and *Lachnospiraceae* (17). Many of these bacteria enter the sewage environment as carriers of AR genes and exhibit AR phenotypes. Exposure to residual antibiotics in sewage wastes can further select for resistance traits, which can then be horizontally amplified within bacterial populations. As such, untreated sewage can increase the spread of AR (16) due to the presence of both residual antibiotics and AR genes, which also can survive wastewater treatment (18). As a consequence, AR genes are now considered to be “environmental contaminants” (19).

Citywide environmental microbiome studies can serve to monitor the health of the residents of the city. New York City (NYC), the largest U.S. city, with >8.6 million inhabitants, is comprised of five boroughs: Manhattan, Brooklyn, Queens, the Bronx, and Staten Island, each with ethnically diverse populations (20). The city is home to the country’s largest health care base with numerous busy hospitals. NYC’s enormous waste disposal needs are served by an extensive sewer system, including >7,000 miles of pipes leading to 14 wastewater treatment plants distributed across the five boroughs. The majority of these sewers are combined, carrying a mix of industrial and domestic wastewater, along with storm and rain water runoffs (21).

In this study, we collected 51 raw sewage samples from the 14 NYC wastewater treatment plants (WWTPs), which included serial specimens taken over a 6-month period from February to August 2015 (Fig. 1). Based on these samples, we conducted



**FIG 2** Alpha diversity in 102 sequenced sewage samples. Rarefaction was based on (A) phylogenetic diversity. (B) Observed number of OTUs. The x axes represent the numbers of sequences used in the rarefaction analysis. (C) Box-and-whisker plots depicting the species-level richness and diversity of OTUs based on the Shannon diversity index in the sewage samples. Within the boxes, the central line represents the median values, and the upper and lower boundaries indicate the 75th and 25th percentiles, respectively. The whiskers show the maximum and minimum values. Top panels show results based on location (NYC borough); bottom panels show results based on sampling season. Statistical significance was tested by one-way analysis of variance using the Kruskal-Wallis test, followed by multiple comparisons using Dunn's test (\*\*\*\*,  $P < 0.0001$ ).

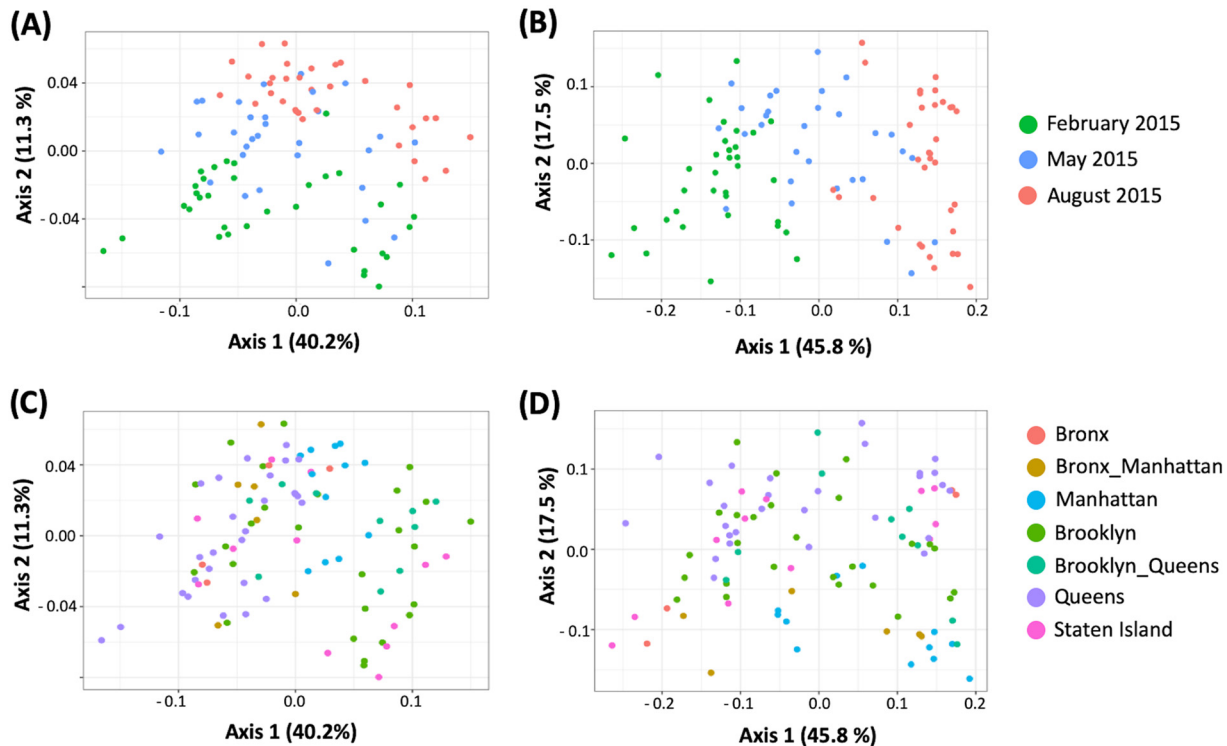
a culture-independent longitudinal investigation of the bacterial microbiome in NYC sewage and assessed the presence of clinically important AR genes. The two key questions that we sought to address were: (i) does the microbial population in NYC sewage vary between the boroughs over time; and (ii), can NYC sewage be used to monitor the spread of environmental antibiotic resistance in the city?

## RESULTS

**16S rRNA sequence analysis of NYC sewage samples.** The two MiSeq 16S rRNA sequencing runs of the raw sewage DNA samples yielded 39,010,014 (4,083 Mb) and 53,685,483 (5,619 Mb) sequences, with 92.1 and 85.5% showing Q30 values, respectively. After sequence assembly and quality filtering using QIIME, 15,984,186 paired-end reads remained for OTU picking. The 102 sequenced samples were represented by 384 unique OTUs from the generated paired-end reads. The number of reads in each sample ranged from a minimum of 58,892 to a maximum of 222,182, with a mean of  $134,364 \pm 31,164$  reads/sample. All further diversity analyses were performed on operational taxonomic unit (OTU) tables rarefied to 58,892 reads.

Alpha diversity analysis showed that this sequence depth provides an even sampling distribution across the population, based on both the phylogenetic diversity, as well as the observed number of OTUs (Fig. 2A and B). The Shannon diversity index had a mean value of 5.0 across the total population. Compared between sampling time points, the index was found to be significantly higher ( $P < 0.001$ ) during August compared to February and May (Fig. 2C).

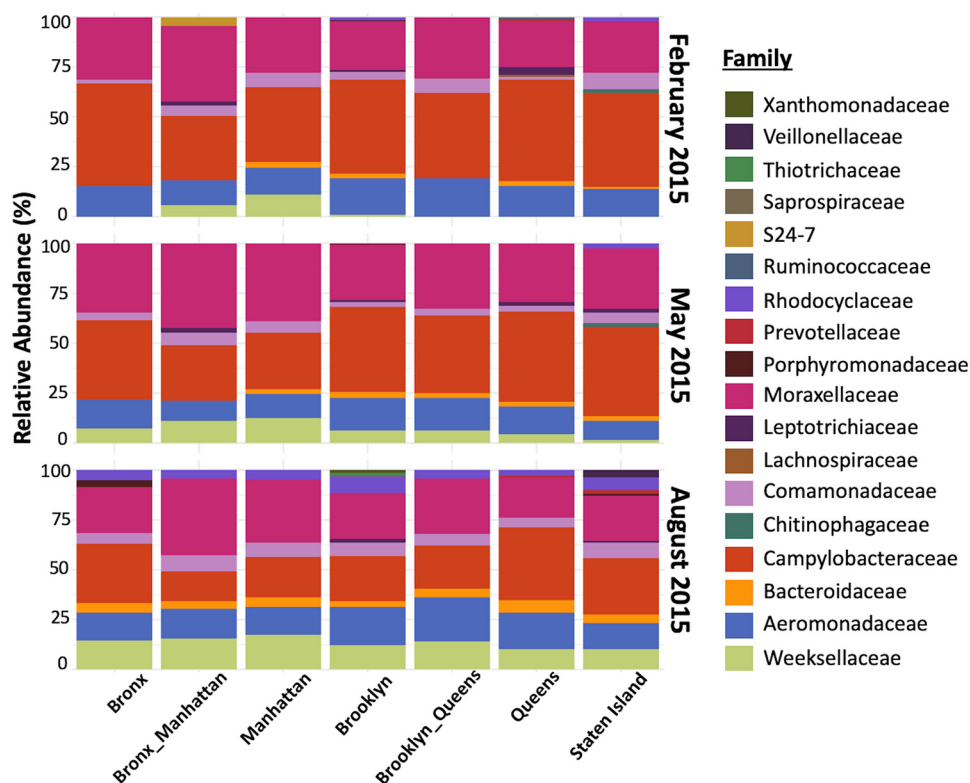
**Beta diversity analysis and batch effects.** Both weighted and unweighted UniFrac distances were used to determine the beta diversity within the sample populations.



**FIG 3** Beta diversity in sewage samples. PCoA results represent unweighted (A and C) and weighted (B and D) UniFrac distances by NYC major drainage area (lower panels) and sampling season (upper panels). Statistical significance was tested by the Adonis test using the Vegan package in R. Each of the data sets indicated distinctions between the tested groups at  $P < 0.001$ .

Initial analyses using the OTUs generated from the open reference data set revealed a strong clustering pattern based on the two Illumina MiSeq runs, especially in the unweighted analyses (see Fig. S1 in the supplemental material). Since the technical replicates from the May samples had been split between two sequencing runs (Table S1), the OTUs for this time point were studied in detail. The novel OTUs (94%) identified as part of the *de novo* clustering algorithm during the open reference OTU picking were found to be the major factor differentiating between the replicates. When the beta diversity of only the May samples was then replotted after the exclusion of these OTUs, the clustering pattern previously observed in the unweighted UniFrac analysis was found to diminish (Fig. S2A and C). Hence, the imbalance in the *de novo* OTU sequences between the runs could explain this artifact. Repeating the diversity analyses for the entire sample set using the OTUs generated from the closed-reference database showed elimination of this technical artifact. As such, to maintain analytical uniformity, all further analyses for this study were performed using the closed reference OTU database, despite the potential loss of unique bacterial diversities. The principal coordinate analysis (PCoA) plots that were generated using the unweighted UniFrac distances showed the 102 samples to be very tightly clustered, with no significant separation due to the sequencing run or sampling season, except for a single sample (ID S050.1) from the Bronx (Fig. S3A and C). When this outlier sample was excluded from the analysis, PCoA plots generated using both unweighted and weighted UniFrac distances showed a more distributed population (Fig. 3). By PERMANOVA analysis, 22.4 and 12.5% of the variation in the population could be accounted for by sampling month (Fig. 3A) and borough of origin, respectively (Fig. 3C); these relationships were significant ( $P < 0.01$ ) using the Adonis test for variance. Only 0.5% ( $P = 0.57$ ) of the population structure was explained by sequencing run differences, compared to 8.4% in the analysis when the open-reference OTU picking algorithm was used ( $P < 0.001$ ).

**Bacterial taxonomic diversity in the NYC sewage.** Taxonomic diversity analysis revealed a largely similar and stable population structure across all the samples and



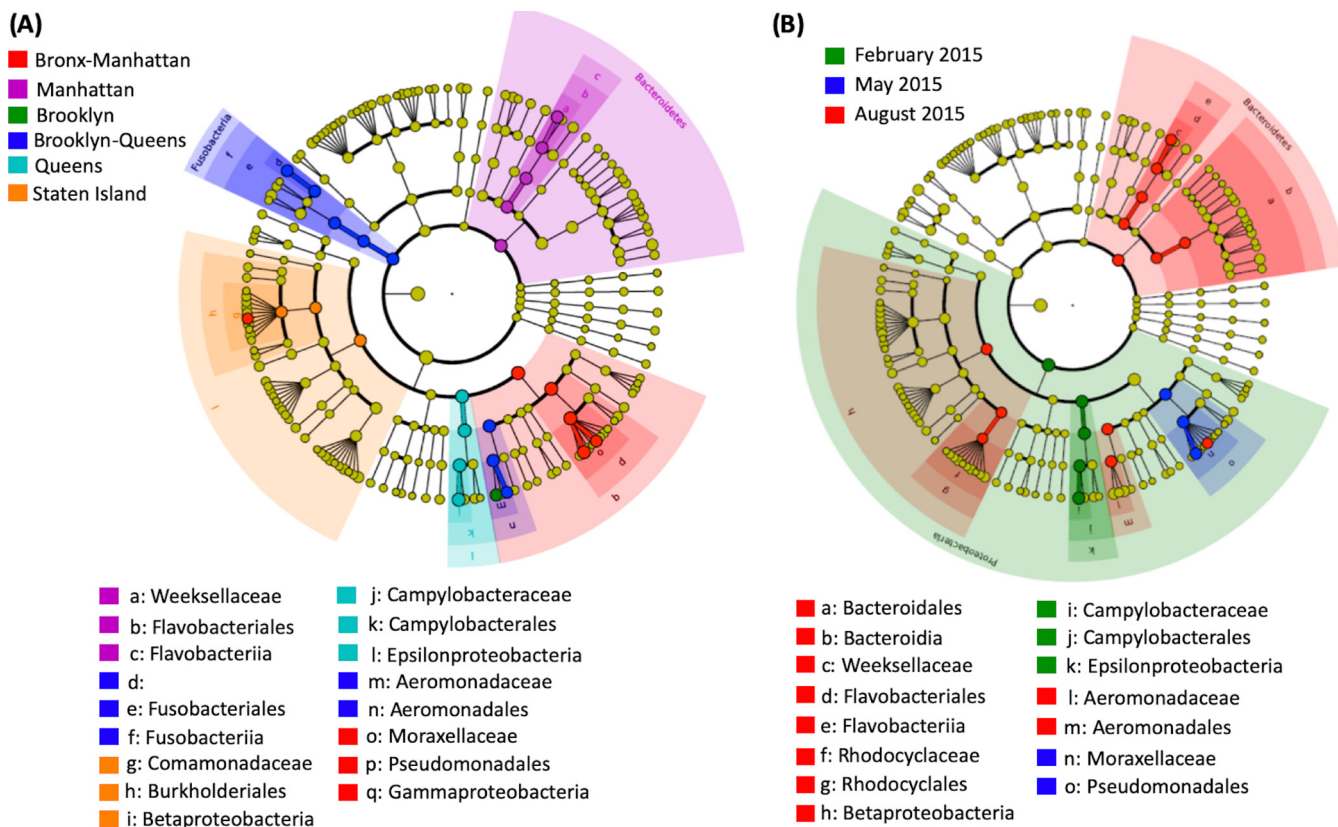
**FIG 4** Relative taxonomic abundances of the assignments in 102 sewage samples. Samples are grouped by family according to NYC borough and sampling season. Taxonomy was assigned to the identified OTUs based on Greengenes database v13.8. Only families represented by >3% abundance are listed.

across multiple time points, as well as across the NYC boroughs, with minor variation (Fig. 4). At the phylum level, *Proteobacteria* were found to be dominant (76.4%), followed by *Bacteroidetes* (14.1%), along with smaller proportions of *Fusobacteria* and *Firmicutes*; representatives of all four phyla being present in all samples. Phyla detected in some of the samples at low abundances included *Acidobacteria*, *Actinobacteria*, *Chlorobi*, *Spirochaetes*, *Verrucomicrobia*, and WWEI (an OTU previously identified in an anaerobic sludge digester [22]). At the family level, the OTUs were predominantly assigned to members of the families *Campylobacteraceae* (30.6%), *Moraxellaceae* (22.9%), and *Aeromonadaceae* (12.7%), all three of which were present in all boroughs at all three time points (Fig. 4). The family *Weekselaceae*, present at a relative abundance of 6.2% in the total population, was found in all the boroughs in May and August but only in 3 of 7 in February. Of the 73 families identified which comprised the 384 OTUs (indicated in Table S3), 23 were present in all 102 samples, representing the core microbiome, i.e., the organisms present in all sampled locations at all three time points, although most were present at very low abundances.

The OTUs with significantly higher relative abundances in each NYC borough, as well as each sampling time point, were calculated based on a linear discriminant analysis Effect Size algorithm (LEfSe; Fig. 5). Compared by borough, Manhattan showed a significantly higher abundance of *Bacteroidetes* members as biomarker taxa, while the other boroughs had biomarker taxa belonging to the phyla *Fusobacteria* and *Proteobacteria* (Fig. 5A). No significant enrichment of taxa was observed for the borough the Bronx. Members of *Bacteroidetes* also were significantly elevated in the August samples compared to the others. The August time point showed a higher number ( $n = 10$ ) of significantly different taxon abundances compared to February ( $n = 3$ ) and May ( $n = 2$ ) (Fig. 5B).

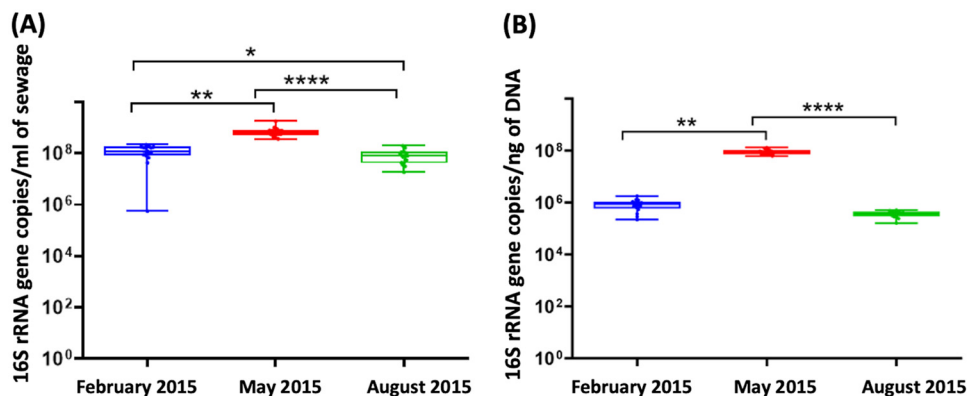
**Total 16S rRNA gene abundance in NYC sewage.** To assess the variation in total bacterial counts (absolute abundance), quantitative PCR (qPCR) assays were performed





**FIG 5** Differentially abundant OTUs identified in 102 sewage samples. LEfSe analyses were performed based on NYC borough (A) and sampling time point (B). Significantly abundant OTUs were determined based on an alpha value of <0.05 and a logarithmic LDA score (effect size) of >4.0.

to measure the number of 16S rRNA genes present in all 102 samples. The qPCR results showed  $10^8$  to  $10^9$  16S gene copies/ml in all the sewage samples across the three sampling points (Fig. 6). The abundances were highest (Fig. 6A) in the sewage samples from the May time point (mean,  $7.1 \times 10^8$  copies/sewage ml) compared to February ( $P < 0.01$ ) and August ( $P < 0.0001$ ). The biomass concentration of the samples calculated based on the ratio of the 16S rRNA copies to the DNA yields also was highest



**FIG 6** 16S rRNA gene copies in NYC sewage samples. (A) Abundances normalized against the volume of sample used. (B) Biomass concentrations of the sewage samples determined by the ratio of the 16S rRNA copies to the total DNA extracted (ng) for each sample. Each box represents a scatter of the mean values of the 17 replicated sewage samples collected at that particular time point. Error bars indicate the standard deviations at each time point. Statistical significance was determined by using a one-factor Friedman’s test, followed by Dunn’s multiple-comparison test (\*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ ).

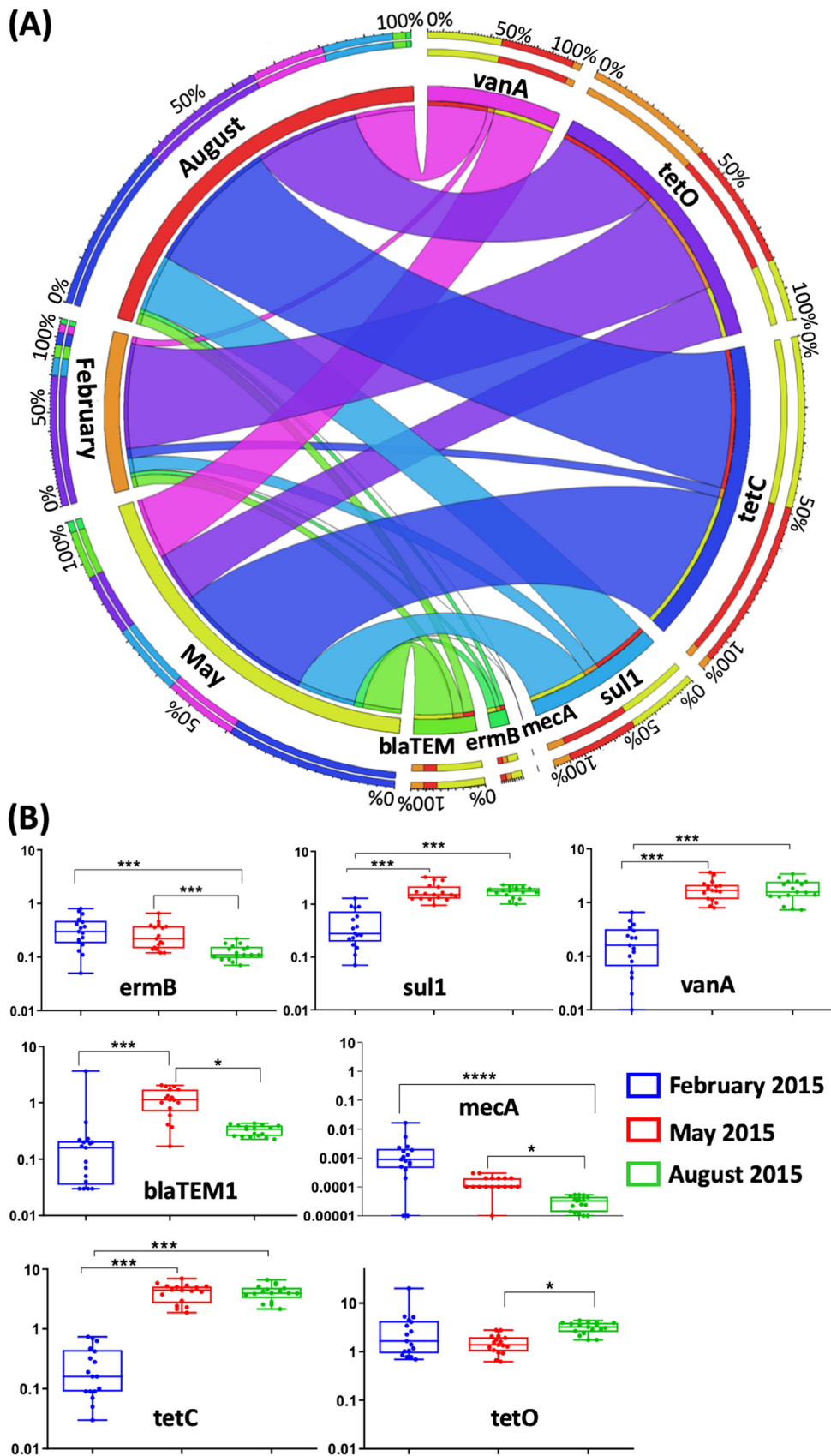
in May (Fig. 6B). Within each of the three testing time points, no significant differences were observed between the boroughs or the individual sewage samples.

**Abundance of antibiotic resistance genes.** Our main goal was to assess abundances of indicator antimicrobial resistance genes (ARGs). Using gene-specific qPCR assays, we examined seven ARGs (Table S2) that all confer resistance to commonly used antibiotics, including macrolides (*ermB*), sulfonamides (*sul1*), vancomycin (*vanA*), beta-lactams (*bla*<sub>TEM1</sub> and *mecA*), and tetracyclines (*tetC* and *tetO*) (8), as well as *mecA*, the most common determinant of MRSA (23). All seven of the resistance genes were present in each of the 102 sewage samples, with variations in concentrations (Fig. 7) and abundances (Fig. S4) across the seasons. For each resistance gene, no significant differences were observed between the NYC boroughs or between the individual sewage samples at each time point; as such, all further analyses were performed by comparing the values of the three seasonal time points. Apart from *mecA*, the other six resistance genes were present at abundances of  $10^6$  to  $10^8$  copies/sewage ml; *tetC* was detected at the highest mean concentrations ( $3.4 \times 10^7$  copies/sewage ml in May and  $3.4 \times 10^6$  copies/sewage ml during August (Table 1). During the February sampling, the highest detected mean concentration was for *tetO* ( $3.72 \times 10^6$  copies per ml of sewage). In parallel with the 16S rRNA gene abundances, five antibiotic resistance gene abundances were found to be significantly higher at the May time point compared to February and August. The exceptions were *vanA* and *mecA*, which showed declines in abundances from February to August. The *mecA* gene was detected at the lowest abundance, with a mean of  $3.7 \times 10^3$  copies/sewage ml in February, decreasing to  $2.5 \times 10^1$  copies/sewage ml in August. Samples from one WWTP in Queens had low concentrations of most of the genes, particularly in February and August. In general, the resistance gene abundances paralleled those of the 16S rRNA genes in the same samples, with the highest ratios for most in May, followed by August. Tetracycline resistance genes predominated in both months (Fig. 7A), and the ratios were lowest for *mecA*, decreasing from  $2 \times 10^{-5}$  (February) to  $3 \times 10^{-7}$  (August) (Fig. 7B). Linear regression analysis showed that the 16S rRNA gene abundances significantly correlated ( $P < 0.001$ ) with the abundances of most of the ARGs in the three sampling time points (Table S4).

## DISCUSSION

One of our priorities was to optimize the DNA yields from sewage samples to ensure optimum coverage of the microbially diverse populations in the samples. Since preliminary assays (24) indicated that for highest yield, genomic DNA extraction from samples should be performed on the day of collection, rather than after storage, we strictly followed that procedure in this study. We used replicate extraction samples in this study to reduce potential sampling and laboratory handling bias during the experimental procedures and, based on the taxonomic summary obtained post-sequencing, the replicates were found to be robust. Since sewage samples comprise domestic, clinical, and environmental waste, we expected the microbial diversity in these samples to be high. For this reason, the samples were split into two sequencing runs (Table S1) to maximize sequencing coverage and depth. The diversity analyses indicated that the sequencing run affected the distribution of the microbial population, causing an artifact involving the OTUs obtained from open reference OTU picking (Fig. S2). Our analysis of the entire study data set using a closed-reference OTU picking strategy avoided false conclusions arising from the batch sequencing run effect, although diminishing the sensitivity of the analysis to detect taxa at low abundances that an open-reference OTU picking strategy would provide. This is relevant to microbiome studies linked with environmental samples, highlighting the utility of combining comparable samples in a single sequencing run, as reported (25).

Beta diversity analysis revealed the bacterial populations to be essentially similar across the boroughs and the three time points, dominated by environmental bacterial taxa compared to human microbiota. This is not surprising since the sewage in these WWTPs is a mix of industrial waste and surface runoff mixed with domestic and clinical



**FIG 7** Antibiotic resistance gene representation in proportion to the 16S rRNA gene copies in the NYC sewage samples. Collections were made during February, May, and August 2015. (A) A Circos diagram represents the  
(Continued on next page)



**TABLE 1** Summary of the results obtained after quantitative monitoring of the 16S rRNA gene and seven ARGs in the sewage samples obtained from 14 NYC WWTPs over three sampling time points<sup>a</sup>

Gene tested	February		May		August	
	Mean gene copies/ml of sewage	Mean % resistance gene/16S gene copies	Mean gene copies/ml of sewage	Mean % resistance gene/16S gene copies	Mean gene copies/ml of sewage	Mean % resistance gene/16S gene copies
<b>16S</b>	<b>2.68E+08</b>	<b>NA</b>	<b>7.1E+08</b>	<b>NA</b>	<b>8.13E+07</b>	<b>NA</b>
<i>mecA</i>	3.65E+03	0.002	9.1E+02	0.0001	2.51E+01	0.00003
<i>bla</i> <sub>TEM-1</sub>	2.00E+05	0.31	7.47E+06	1.17	2.57E+05	0.3248
<i>ermB</i>	5.95E+05	0.13	2.27E+06	0.27	1.02E+05	0.13
<i>tetO</i>	3.72E+06	3.19	1.06E+07	1.53	2.35E+06	3.148
<i>tetC</i>	8.44E+05	0.3	3.04E+07	4.2	3.36E+06	4.15
<i>vanA</i>	4.05E+05	0.21	1.22E+07	1.83	1.70E+06	1.81
<i>sul1</i>	7.37E+05	0.44	1.2E+07	1.77	1.44E+06	1.72

<sup>a</sup>NA, not applicable. Boldfacing indicates data for the 16S gene.

wastes. These results are consistent with analysis of sewage from 71 U.S. cities showing that only 15% of the bacterial sequences were of human origin as determined based on their predominance in human gut microbial populations (17). At the family level, the microbial populations observed in this study shared several similarities, e.g., *Bacteroidaceae*, *Lachnospiraceae*, *Porphyromonadaceae*, and *Prevotellaceae*, among others. Both unweighted and weighted UniFrac analyses showed clustering by sampling time, consistent with seasonal variation. New York City is known to have major seasonal differences in temperature (ranging from mean 23.9°F in February 2015 to 79°F in August 2015) (26) that could strongly influence environmental microbial growth dynamics. The 16S rRNA qPCR assays showing significantly higher populations in May 2015 compared to February and August could not be explained by temperature alone, and these other factors, as yet undefined, might play a role. We were able to obtain from the NYC Department of Environmental Protection (DEP) information for other relevant factors such as atmospheric precipitation for the sampling days and the levels of suspended solids and nutrients in the sewage samples, as well as the biological oxygen demand levels (27), but no conclusive correlations were shown between these values and the results observed.

Similar qPCR assays were used to quantitate selected ARGs, based on their observed abundance in prior studies of aquatic and urban environments (28–30). Although all the sewage samples studied tested positive for all assayed ARGs, the May samples had the highest concentrations for five of the seven tested genes, reflecting the higher microbial load in the samples. Since the starting material used for the analysis in this study was quite low at 1 ml, extra precautions, such as the use of replicate samples at each stage, in-plate standards and controls for the qPCR, and accurate optimization of each primer pair, were applied to eliminate the possibility of any artifact formation that could be misleading.

Seasonal variation in antibiotic prescribing in NYC could explain some of the seasonal variation of ARGs, since this May peak in ARG could reflect a lag from the high reported prescription rates in the late winter months, especially in February and March (31, 32). Another factor that could explain the seasonal variation in the levels of antibiotic resistance genes is the influence of the environment and the dilution effect. There is a high level of snowfall in NYC, as well as storm events in the winter months, while the precipitation levels were very low for the spring and summer sampling time

#### FIG 7 Legend (Continued)

percent breakdown of the presence of the measured ARG concentrations in each month and vice versa indicated by the connecting ribbons. (B) Ratios were calculated by normalization of antibiotic resistance gene abundance against the 16S rRNA gene abundance for each sample at each time point. Each box represents a scatter of the mean values of the ratios for the 17 replicated sewage samples collected at that time point. The bars indicate the standard deviations at each time point. Statistical significance was determined using a one-factor Friedman's test, followed by Dunn's multiple-comparison test (\*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ ).

points of that year. Since these sampled sewage waters are also inclusive of storm and rainwater runoff, the dilution factors introduced by these environmental influences in winter may also play a role in differing levels of resistance genes in the samples tested. Since the sewer system drained the urban area of NYC, where there is not any farming or animal husbandry, we presume that there was no substantial agricultural contribution to the effluent.

Each of the two tested tetracycline resistance genes (*tetO* and *tetC*) showed high abundances for the three time points sampled. Tetracyclines are widely used in human and veterinary clinical medicine and in food animal husbandry, and resistance genes for tetracycline are also spread across environmental and human samples involving both Gram-positive and -negative pathogens and commensals (33). The high abundances of more than one tetracycline resistance gene in this study are consistent with this ubiquity of the *tet* gene group (16, 34), both as single genes and on mobile genetic elements with multiple resistance cassettes (33).

Although present at significantly lower concentrations compared to the other ARGs, *mecA* was detectable in all samples. In two separate studies of municipal and clinical wastewater in Germany, *mecA* was similarly found at much lower levels than other ARGs tested and was detected only in clinical wastewater samples (35, 36). Since all of the wastewater samples in the present study included sewage released from hospitals, their exact contributions to *mecA* abundances cannot be ascertained. *mecA* has been detected in other WWTP samples (37, 38), but abundances were not quantitated, precluding comparison with our study. A United States-based study of MRSA in wastewater reported 83% positivity of influent sewage samples and 8% positivity of effluent samples (39), reflecting the incomplete ability of wastewater treatment to completely clear this ARG. The presence of *mecA* in the sewage samples may both reflect the presence of community-acquired MRSA infections and potentially contribute to their spread.

Since the sewage samples used in these studies were a mix of different sources in the city, we cannot precisely identify the source of the antibiotic resistance genes detected. Antibiotic resistance in the environment is known to be especially high near urban areas because of the influence of human activities (13). Since the sewage samples in this study derive from both surface runoff and groundwater, and as much of the bacterial population reflects environmental bacteria (Fig. 4), we hypothesize that a substantial proportion of the observed ARGs is of environmental origin. Investigations of ARGs in Danish WWTPs indicated that most had low probability of being associated with human pathogens and represent what has been described as a “sewage resistome” (26).

It is pertinent to note here that qPCR provides measurement of only the resistance genes we targeted to study, making this a minimal estimate of the actual ARG presence. A number of recent sewage-based population studies have employed more sensitive methods, such as shotgun metagenomic sequencing, metagenomic cloning, and gene arrays, to estimate the presence of AMR gene populations in wastewater samples (40–43). Unlike most of these studies, we did not observe any significant geographically distinguishing factors between the samples; any variation observed in our study was primarily based on season (42, 43). Hendriksen et al. conducted a global comparison of AMR presence in untreated sewage samples using a shotgun metagenomic sequencing approach and found correspondence between public health factors such as sanitation and the levels of AMR genes and microbial composition in sewage (41). In another study focusing on sewage samples collected in Hong Kong, nanopore metagenomic sequencing helped reveal mobile genetic elements in bacteria as the predominant source of AMR genes (40). Since qPCR offers the advantage of targeting specific genes, a combination of methods using metagenomic sequencing and qPCR assays of genes of interest could be an important approach for global monitoring of antimicrobial resistance in aquatic environments.

Sewage also can contain residual antibiotic compounds (44) which exert further selective pressure on the environmental bacterial populations. The closed WWTP niche

also may select for horizontal gene transfer (45) among the sewage bacteria, especially since many ARGs are present on mobile elements (36) and may be coselected by factors conferring resistance to heavy metals (46), biocides (including triclosan and chlorhexidine) (47), and pharmaceutical compounds. Toxicological data for antibiotic levels in these samples that correspond with sampling times were not available. However, several studies have shown the presence of high levels of pharmaceuticals, including some types of antibiotics, in the aquatic environment of the NYC area that potentially correspond to wastewater discharge and contamination (48, 49).

Although this study focuses only on the microbial and ARG populations in raw, untreated sewage, resistance may also persist after wastewater treatment, with ARGs present in the effluents of sewage treatment plants (44, 50). Since the treated effluents are often released into rivers and other groundwater sources and also being reused in recent times (51), such selective pressures could add to the environmental antibiotic resistance reservoir, highlighting the importance of strict regulations in this regard. Recent studies have reported the unexpected presence of ARGs in locations as isolated as the Arctic, a testament to the ease of widespread dissemination of antibiotic resistance through human activities (52, 53). There are currently no regulations in the United States for monitoring the level of antibiotics or ARGs in aquatic reservoirs or even in drinking water; such rules should be considered as part of broader stewardship efforts.

In conclusion, we report the results from a citywide bacterial diversity and antibiotic resistance study conducted on New York City sewage, showing a stable, generally conserved, bacterial population, and indicator ARGs that were similarly conserved. With increasing global urbanization, citywide studies such as these can help to establish baseline characteristics of the urban microbiome and resistome. These data also can serve as a reference point for further longitudinal studies, comparisons with other locales, and future major extreme weather events.

## MATERIALS AND METHODS

**Sewage collection and DNA extraction.** For this study, raw sewage samples were obtained from DEP WWTPs at three time points: February, May, and August 2015. For each time point, 17 samples were collected from the 14 NYC WWTPs, representing the five NYC boroughs (Fig. 1); samples from three of the WWTPs were split to represent smaller drainage areas within these plants, resulting in a total of 17 samples (27). Two of these samples also were from drainage areas shared across boroughs and thus have been indicated as “Bronx\_Manhattan” and “Brooklyn\_Queens” in the data (see Table S1 in the supplemental material). Each sample was a composite of sewage obtained over a 24-h period from each plant. Additional details of the raw sewage collected from the WWTPs are described in Maritz et al. (27).

Samples were provided in 200-ml containers, from which 1-ml volumes were used for DNA extraction. During earlier protocol optimizations, the majority of the samples were found to be densely particulate, which clogged the filter membranes during passage of small sewage volumes. After yield comparisons with centrifugation processes (24), the protocol was optimized to use 1 ml of sewage sample. Each sample was then subjected to rigorous bead beating for ~20 min using 0.7-mm garnet beads, followed by nucleic acid extraction using Powersoil DNA extraction kits (Mobio, Carlsbad, CA), according to the manufacturer’s instructions. Duplicate extractions were performed simultaneously, and all processing and extraction procedures were completed on the day of sampling for standardization and to ensure maximum DNA yields. Nucleic acid quantitation was carried out with the Qubit 2.0 (Thermo Fisher Scientific, Waltham, MA), and the samples were stored at –20°C.

**16S rRNA library preparation and sequencing.** For each DNA sample, the V4 region of the 16S rRNA gene was amplified using the primer set 515F/806R (54) with 12-bp barcodes in triplicate reactions. These standard bacterial primer sets have been used in other large-scale community studies (55). Each 25- $\mu$ l PCR reaction mixture contained 10  $\mu$ l of 5 Prime Hot Master mix (5-PRIME), 1  $\mu$ l each of the forward and reverse primers (5  $\mu$ M), 10  $\mu$ l of molecular-biology-grade distilled water, and 3  $\mu$ l of each genomic DNA template. Reaction negative controls, as well as kit extraction controls, were amplified simultaneously. The PCR conditions were as follows: 94°C for 3 min; 35 cycles of 94°C for 45 s, 50°C for 1 min, and 72°C for 1.5 min; followed by a final extension step for 10 min at 72°C. The amplified triplicate reactions were pooled and quantified using a Quant-iT PicoGreen dsDNA assay kit (Invitrogen, Waltham, MA), and then equimolar concentrations of each amplified sample were combined into a pool that was purified by using a QIAquick PCR purification kit (Qiagen, MD). The pooled amplicons were sequenced on an Illumina MiSeq 2500 instrument using 2  $\times$  150 paired-end sequencing at the NYU Genome Technology Center. The samples were split into two batches to obtain a higher sequencing depth to maximize capturing diversity. The distribution of the samples across the two sequencing runs and PCR is indicated in Table S1. PCR controls were included in each run.

**Bioinformatics analysis.** Bioinformatic processing was performed using QIIME v1.9.1 (56) with parameters as described previously (57). Briefly, mate-pair reads were joined using fastq-join (58) of the fastq\_join.py command, and reads with at least a 40-nucleotide (nt) overlap between mate-pairs were retained for further quality filtering. The remaining sequences were truncated at any site containing >3 consecutive bases receiving a quality score of <20, and any read containing one or more ambiguous base call was discarded, as were reads <190 nt in length. Samples were then demultiplexed, and OTUs were picked first using the open reference and then the closed-reference OTU picking methods in QIIME. Reads were clustered into 97% identity OTUs using UCLUST (v1.2.22), followed by taxonomy assignment using the RDP classifier (v2.2) against the Greengenes database version 13.8 (59) with a 0.5 confidence (60). All singleton OTUs were removed and chimeric sequences were then removed using ChimeraSlayer (61). The resulting OTU table was filtered to remove OTUs with a relative abundance of <0.01% across all samples to eliminate minor taxa (57) and was used in all subsequent analyses.

Diversity analyses were performed using the QIIME script core\_diversity\_analyses.py. Alpha diversity was calculated within QIIME using the Shannon-index, PD (phylogenetic diversity) tree, and observed OTUs. Rarefaction plots were subsampled to the lowest sample sequencing depth of 58,892 sequences/sample, and significance was determined using a two-sample nonparametric *t* test with 999 Monte Carlo permutations and corrected for multiple comparisons using Bonferroni correction. Beta diversity was assessed using unweighted and weighted UniFrac distances, and ordinations were plotted using PCoA in the Phyloseq and ggplot2 packages in R (62, 63). Significance was determined using the nonparametric permutational MANOVA (PerMANOVA) Adonis test with 999 permutations using the Vegan package (v2.5-2) in R (64). Linear discriminant analysis effect size (LEfSe) analysis (65) was performed at the family level on the filtered OTU table using an alpha value of <0.05 and a Linear Discriminant Analysis (LDA) score greater than 4.0. Taxonomic abundance plots were generated using the R packages, Phyloseq, and ggplot2. OTUs were summarized at the family level with a minimum relative abundance of 3% in each sample.

**Preparation of standards for qPCR.** The antibiotic resistance genes for qPCR were selected based on resistance to some of the more commonly used antibiotics and their abundances observed in prior studies of aquatic and urban environments (28–30). The reference standards for the qPCR assays were constructed by cloning a large construct of each resistance gene into a suitable vector backbone, followed by confirmation of sequence and orientation. The standards for the 16S rRNA and *mecA* genes were cloned on a pGEM-T Easy (Promega) vector backbone and obtained from Zhan Gao (66). Commercially purchased pBR322 (NEB) was used as a reference standard for the *bla*<sub>TEM</sub> and *tetC* genes. The standards for the other resistance genes were constructed for this study using the pGEM-T Easy vector for cloning the reference genes. The full-length genes for *bla*<sub>S<sub>HV</sub></sub> and *sul1* were isolated from one of the sewage samples used in this study. The *ermB* gene was amplified from an NYU in-house strain of *S. aureus* (B. Shopsis, unpublished data). The *vanA* gene was amplified from *E. faecium* DSMZ 17050, while the *tetO* gene was amplified from the fecal DNA from mice treated with tetracycline (67). Each PCR contained 5  $\mu$ l of the 10 $\times$  buffer, 1  $\mu$ l of a deoxynucleoside triphosphate mix, 5  $\mu$ l each of the forward and reverse primers, 0.25  $\mu$ l of *Taq* polymerase, and 1  $\mu$ l (~10 ng) of genomic DNA sample adjusted to a final volume of 50  $\mu$ l using molecular-biology-grade dH<sub>2</sub>O. All the primers for the full-length amplification of these genes (Table S2) were synthesized by Integrated DNA Technologies (Coralville, IA). The PCR conditions were as follows: (i) initial denaturation at 95°C for 2 min; (ii) 35 cycles of 95°C for 30 s, an annealing temperature as listed in Table S2 for 45 s, and 72°C for 1 min; and (iii) final elongation at 72°C for 7 min. The amplified PCR products were confirmed using 2% agarose gel electrophoresis and then purified using the QIAquick PCR purification kit. The purified products were then used to construct standards by cloning using the pGEM-T Easy vector system (Promega, WI).

**Quantitation of resistance genes in the sewage samples.** The quantitation of the 16S rRNA genes and antibiotic resistance genes in the sewage samples were performed by PCR amplification on a LightCycler 480 Instrument II (Roche Diagnostics, IN). qPCR reactions of 20- $\mu$ l volumes each were set up using the LightCycler 480 SYBR green I Master hot-start reaction mix. The primers used and their annealing temperatures are listed in Table S2 and were synthesized by Integrated DNA Technologies. The amplification conditions for the qPCR program were as follows: 40 cycles of 95°C for 10 s, respective annealing temperatures for 20 s, and the fluorescence acquisition step at 72°C for 20 s. Tenfold dilutions of the standards (with copy numbers ranging from 10<sup>8</sup> to 10<sup>0</sup>) for each gene were used to first plot the standard curve and then to optimize the primers to ensure that acceptable efficiencies of 90 to 110% were maintained for each set of primers. Baseline and melting curve analyses were performed using the LightCycler 480 software (Roche). Each reaction was carried out in duplicate wells, and each run was performed with replicate samples and standards on the same plate to maintain uniformity. Since some of the data sets did not exhibit normal distribution, a nonparametric one-factor Friedman's test was used to estimate any statistically significant differences within the means of the groups in the population for each gene tested. This was followed by the Dunn's multiple-comparison tests to compare the *P* values and statistical significances between the three time points in each gene set. Linear regression analysis was performed to calculate the relationship between the abundances of the 16S rRNA genes and AR genes using log-transformed values. All statistical analyses were performed using GraphPad Prism 7. Visualization of the ARG qPCR data was performed using GraphPad Prism 7 and Circos Table Viewer v0.63-9 (52, 68).

**Data availability.** All the MiSeq sequencing output fastq files and metadata are publicly available via the NCBI Sequence Read Archive (SRA) under accession no. [PRJNA542507](https://www.ncbi.nlm.nih.gov/sra/PRJNA542507).

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mSystems.00327-19>.

**FIG S1**, TIF file, 1 MB.

**FIG S2**, TIF file, 0.9 MB.

**FIG S3**, TIF file, 0.9 MB.

**FIG S4**, TIF file, 1.1 MB.

**TABLE S1**, XLSX file, 0.01 MB.

**TABLE S2**, XLSX file, 0.01 MB.

**TABLE S3**, XLSX file, 0.03 MB.

**TABLE S4**, XLSX file, 0.01 MB.

## ACKNOWLEDGMENTS

We thank the staff of the NYC Department of Environmental Protection for sewage samples.

Funding was provided by an NYU Grand Challenge project, and by the C&D and Zlinkoff Funds.

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

J.M.M. is now an employee of Merck Research Laboratories.

## REFERENCES

- D'Costa VM, King CE, Kalan L, Morar M, Sung WW, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, Golding GB, Poinar HN, Wright GD. 2011. Antibiotic resistance is ancient. *Nature* 477:457–461. <https://doi.org/10.1038/nature10388>.
- Martinez JL, Baquero F. 2000. Mutation frequencies and antibiotic resistance. *Antimicrob Agents Chemother* 44:1771–1777. <https://doi.org/10.1128/AAC.44.7.1771-1777.2000>.
- Schaeffler S. 1982. Bacteriophage-mediated acquisition of antibiotic resistance by *Staphylococcus aureus* type 88. *Antimicrob Agents Chemother* 21:460–467. <https://doi.org/10.1128/AAC.21.3.460>.
- Laxminarayan R, Duse A, Wattal C, Zaidi AKM, Wertheim HFL, Sumpradit N, Vlieghe E, Hara GL, Gould IM, Goossens H, Greko C, So AD, Bigdeli M, Tomson G, Woodhouse W, Ombaka E, Peralta AQ, Qamar FN, Mir F, Kariuki S, Bhutta ZA, Coates A, Bergstrom R, Wright GD, Brown ED, Cars O. 2013. Antibiotic resistance: the need for global solutions. *Lancet Infect Dis* 13:1057–1098. [https://doi.org/10.1016/S1473-3099\(13\)70318-9](https://doi.org/10.1016/S1473-3099(13)70318-9).
- Laxminarayan R, Sridhar D, Blaser M, Wang M, Woolhouse M. 2016. Achieving global targets for antimicrobial resistance. *Science* 353:874–875. <https://doi.org/10.1126/science.aaf9286>.
- Rees T, Blaser M. 2016. Waking up from antibiotic sleep. *Perspect Public Health* 136:202–204. <https://doi.org/10.1177/1757913916643449>.
- CDC. 2013. Antibiotic resistance threats in the United States. Centers for Disease Control and Prevention, Atlanta, GA.
- Hicks LA, Bartoces MG, Roberts RM, Suda KJ, Hunkler RJ, Taylor TH, Jr, Schrag SJ. 2015. US outpatient antibiotic prescribing variation according to geography, patient population, and provider specialty in 2011. *Clin Infect Dis* 60:1308–1316.
- Blaser MJ. 2016. Antibiotic use and its consequences for the normal microbiome. *Science* 352:544–545. <https://doi.org/10.1126/science.aad9358>.
- Allen HK. 2014. Antibiotic resistance gene discovery in food-producing animals. *Curr Opin Microbiol* 19:25–29. <https://doi.org/10.1016/j.mib.2014.06.001>.
- Chang Q, Wang W, Regev-Yochay G, Lipsitch M, Hanage WP. 2014. Antibiotics in agriculture and the risk to human health: how worried should we be? *Evol Appl* . <https://doi.org/10.1111/eva.12185>.
- Brüssow H. 2017. Adjuncts and alternatives in the time of antibiotic resistance and in-feed antibiotic bans. *Microb Biotechnol* 10:674–677. <https://doi.org/10.1111/1751-7915.12730>.
- Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davies J, Handelsman J. 2010. Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* 8:251–259. <https://doi.org/10.1038/nrmicro2312>.
- Perron GG, Whyte L, Turnbaugh PJ, Hanage WP, Dantas G, Desai MM. 2013. Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. *PLoS One* 10:e0069533. <https://doi.org/10.1371/journal.pone.0069533>.
- Riesenfeld CS, Goodman RM, Handelsman J. 2004. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ Microbiol* 6:981–989. <https://doi.org/10.1111/j.1462-2920.2004.00664.x>.
- Rodríguez-Mozaz S, Chamorro S, Martí E, Huerta B, Gros M, Sánchez-Melsió A, Borrego CM, Barceló D, Balcázar JL. 2015. Occurrence of antibiotics and antibiotic resistance genes in hospital and urban wastewaters and their impact on the receiving river. *Water Res* 69:234–242. <https://doi.org/10.1016/j.watres.2014.11.021>.
- Newton RJ, McLellan SL, Dila DK, Vineis JH, Morrison HG, Eren AM, Sogin ML. 2015. Sewage reflects the microbiomes of human populations. *mBio* 6:e02574. <https://doi.org/10.1128/mBio.02574-14>.
- Czekalski N, Berthold T, Caucci S, Egli A, Burgmann H. 2012. Increased levels of multiresistant bacteria and resistance genes after wastewater treatment and their dissemination into Lake Geneva, Switzerland. *Front Microbiol* 3:106. <https://doi.org/10.3389/fmicb.2012.00106>.
- Sanderson H, Fricker C, Brown RS, Majury A, Liss SN. 2016. Antibiotic resistance genes as an emerging environmental contaminant. *Environ Rev* 2016:1–14. <https://doi.org/10.1139/er-2015-0069>.
- Logan JR, Zhang W, Alba RD. 2002. Immigrant enclaves and ethnic communities in New York and Los Angeles. *American Sociological Rev* 67:299–322. <https://doi.org/10.2307/3088897>.
- NYC Environmental Protection. 2018. New York City's wastewater. NYC Environmental Protection, New York, NY. <https://www1.nyc.gov/html/dep/html/wastewater/index.shtml>. Accessed 5 September 2018.
- Rivière D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissenbach J, Li T, Camacho P, Sghir A. 2009. Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *ISME J* 3:700. <https://doi.org/10.1038/ismej.2009.2>.
- Antonanzas F, Lozano C, Torres C. 2015. Economic features of antibiotic resistance: the case of methicillin-resistant *Staphylococcus aureus*. *Pharmacoeconomics* 33:285–325. <https://doi.org/10.1007/s40273-014-0242-y>.
- Maritz JM, Rogers KH, Rock TM, Liu N, Joseph S, Land KM, Carlton JM. 2017. An 18S rRNA workflow for characterizing protists in sewage, with a focus on zoonotic trichomonads. *Microb Ecol* 74:923–936. <https://doi.org/10.1007/s00248-017-0996-9>.
- Chase J, Fouquier J, Zare M, Sonderegger DL, Knight R, Kelley ST, Siegel J, Caporaso JG. 2016. Geography and location are the primary drivers of office microbiome composition. *mSystems* 11:e00022-16. <https://doi.org/10.1128/mSystems.00022-16>.
- National Weather Service. 2019. Average monthly and annual temperatures at Central Park. U.S. National Weather Service, Washington, DC. <https://www.weather.gov/media/okx/Climate/CentralPark/monthlyannualtemp.pdf>. Accessed 25 August 2019AQ.
- Maritz JM, Ten Eyck TA, Elizabeth Alter S, Carlton JM. 2019. Patterns of



- protist diversity associated with raw sewage in New York City. *ISME J* <https://doi.org/10.1038/s41396-019-0467-z>.
28. Bockelmann U, Dorries HH, Ayuso-Gabella MN, Salgot de Marçay M, Tandoi V, Levantesi C, Masciopinto C, Van Houtte E, Szewzyk U, Wintgens T, Grohmann E. 2009. Quantitative PCR monitoring of antibiotic resistance genes and bacterial pathogens in three European artificial groundwater recharge systems. *Appl Environ Microbiol* 75:154–163. <https://doi.org/10.1128/AEM.01649-08>.
  29. Munir M, Wong K, Xagorarakis I. 2011. Release of antibiotic-resistant bacteria and genes in the effluent and biosolids of five wastewater utilities in Michigan. *Water Res* 45:681–693. <https://doi.org/10.1016/j.watres.2010.08.033>.
  30. Proia L, von Schiller D, Sánchez-Melsió A, Sabater S, Borrego CM, Rodríguez-Mozaz S, Balcázar JL. 2016. Occurrence and persistence of antibiotic resistance genes in river biofilms after wastewater inputs in small rivers. *Environ Pollut* 210:121–128. <https://doi.org/10.1016/j.envpol.2015.11.035>.
  31. Sun L, Klein EY, Laxminarayan R. 2012. Seasonality and temporal correlation between community antibiotic use and resistance in the United States. *Clin Infect Dis* 55:687–694. <https://doi.org/10.1093/cid/cis509>.
  32. Durkin MJ, Jafarzadeh SR, Hsueh K, Sallah YH, Munshi KD, Henderson RR, Fraser VJ. 2018. Outpatient antibiotic prescription trends in the United States: a national cohort study. *Infect Control Hosp Epidemiol* 39:584–589. <https://doi.org/10.1017/ice.2018.26>.
  33. Chopra I, Roberts M. 2001. Tetracycline antibiotics: mode of action applications molecular biology and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev* 65:232–260. <https://doi.org/10.1128/MMBR.65.2.232-260.2001>.
  34. Auerbach EA, Seyfried EE, McMahon KD. 2007. Tetracycline resistance genes in activated sludge wastewater treatment plants. *Water Res* 41:1143–1151. <https://doi.org/10.1016/j.watres.2006.11.045>.
  35. Volkman H, Schwartz T, Bischoff P, Kirchen S, Obst U. 2004. Detection of clinically relevant antibiotic-resistance genes in municipal wastewater using real-time PCR (TaqMan). *J Microbiol Methods* 56:277–286. <https://doi.org/10.1016/j.mimet.2003.10.014>.
  36. Schwartz T, Kohlen W, Jansen B, Obst U. 2003. Detection of antibiotic-resistant bacteria and their resistance genes in wastewater surface water and drinking water biofilms. *FEMS Microbiol Ecol* 43:325–335. <https://doi.org/10.1111/j.1574-6941.2003.tb01073.x>.
  37. Thompson JM, Gündoğdu A, Stratton HM, Katouli M. 2013. Antibiotic-resistant *Staphylococcus aureus* in hospital wastewaters and sewage treatment plants with special reference to methicillin-resistant *Staphylococcus aureus* (MRSA). *J Appl Microbiol* 114:44–54. <https://doi.org/10.1111/jam.12037>.
  38. Naquin A, Shrestha A, Sherpa M, Nathaniel R, Boopathy R. 2015. Presence of antibiotic resistance genes in a sewage treatment plant in Thibodaux, Louisiana, USA. *Bioresour Technol* 188:79–83. <https://doi.org/10.1016/j.biortech.2015.01.052>.
  39. Rosenberg Goldstein RE, Micallef SA, Gibbs SG, Davis JA, He X, George A, Kleinfelder LM, Schreiber NA, Mukherjee S, Sapkota A, Joseph SW, Sapkota AR. 2012. Methicillin-resistant *Staphylococcus aureus* (MRSA) detected at four U.S. wastewater treatment plants. *Environ Health Perspect* 120:1551–1558. <https://doi.org/10.1289/ehp.1205436>.
  40. Che Y, Xia Y, Liu L, Li A-D, Yang Y, Zhang T. 2019. Mobile antibiotic resistance in wastewater treatment plants revealed by Nanopore metagenomic sequencing. *Microbiome* 7:44. <https://doi.org/10.1186/s40168-019-0663-0>.
  41. Hendriksen RS, Munk P, Njage P, van Bunnik B, McNally L, Lukjancenko O, Röder T, Nieuwenhuijse D, Pedersen SK, Kjeldgaard J, Kaas RS, Clausen PTL, Vogt JK, Leekitcharoenphon P, van de Schans MGM, Zuidema T, de Roda Husman AM, Rasmussen S, Petersen B, Amid C, Cochrane G, Sicheritz-Ponten T, Schmitt H, Alvarez JRM, Aidara-Kane A, Pamp SJ, Lund O, Hald T, Woolhouse M, Koopmans MP, Vigre H, Petersen TN, Aarestrup FM. 2019. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat Commun* 10:1124. <https://doi.org/10.1038/s41467-019-08853-3>.
  42. Pärnänen KMM, Narciso-da-Rocha C, Kneis D, Berendonk TU, Cacace D, Do TT, Elpers C, Fatta-Kassinos D, Henriques I, Jaeger T, Karkman A, Martinez JL, Michael SG, Michael-Kordatou I, O'Sullivan K, Rodriguez-Mozaz S, Schwartz T, Sheng H, Sorum H, Stedtfeld RD, Tiedje JM, Giustina SVD, Walsh F, Vaz-Moreira I, Virta M, Manaia CM. 2019. Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Sci Adv* 5:eaau9124. <https://doi.org/10.1126/sciadv.aau9124>.
  43. Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, Calderon M, Cabrera L, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G. 2016. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* 533:212. <https://doi.org/10.1038/nature17672>.
  44. Xu J, Xu Y, Wang H, Guo C, Qiu H, He Y, Zhang Y, Li X, Meng W. 2015. Occurrence of antibiotics and antibiotic resistance genes in a sewage treatment plant and its effluent-receiving river. *Chemosphere* 119:1379–1385. <https://doi.org/10.1016/j.chemosphere.2014.02.040>.
  45. Shakibaie MR, Jalilzadeh KA, Yamakanamardi SM. 2009. Horizontal transfer of antibiotic resistance genes among gram negative bacteria in sewage and lake water and influence of some physico-chemical parameters of water on conjugation process. *J Environ Biol* 30:45–49.
  46. Di Cesare A, Eckert EM, D'Urso S, Bertoni R, Gillan DC, Wattiez R, Corno G. 2016. Co-occurrence of integrase 1 antibiotic and heavy metal resistance genes in municipal wastewater treatment plants. *Water Res* 94:208–214. <https://doi.org/10.1016/j.watres.2016.02.049>.
  47. Jutkina J, Marathe N, Flach C-F, Larsson D. 2018. Antibiotics and common antibacterial biocides stimulate horizontal transfer of resistance at low concentrations. *Science of the Total Environment* 616:172–178. <https://doi.org/10.1016/j.scitotenv.2017.10.312>.
  48. Cantwell MG, Katz DR, Sullivan JC, Shapley D, Lipscomb J, Epstein J, Juhl AR, Knudson C, O'Mullan GD. 2018. Spatial patterns of pharmaceuticals and wastewater tracers in the Hudson River Estuary. *Water Res* 137:335–343. <https://doi.org/10.1016/j.watres.2017.12.044>.
  49. Young S, Juhl A, O'Mullan GD. 2013. Antibiotic-resistant bacteria in the Hudson River Estuary linked to wet weather sewage contamination. *J Water Health* 11:297–310. <https://doi.org/10.2166/wh.2013.131>.
  50. Bengtsson-Palme J, Hammaren R, Pal C, Ostman M, Bjorlenius B, Flach CF, Fick J, Kristiansson E, Tysklind M, Larsson DG. 2016. Elucidating selection processes for antibiotic resistance in sewage treatment plants using metagenomics. *Sci Total Environ* 572:697–712. <https://doi.org/10.1016/j.scitotenv.2016.06.228>.
  51. Chiemchaisri C, Chiemchaisri W, Prasertkulsak S, Hamjinda NS, Kootatep T, Itonaga T, Yamamoto K. 2015. Evaluation of treated sewage reuse potential and membrane-based water reuse technology for the Bangkok Metropolitan area. *Water Sci Technol* 72:1954–1961. <https://doi.org/10.2166/wst.2015.420>.
  52. McCann CM, Christgen B, Roberts JA, Su J-Q, Arnold KE, Gray ND, Zhu Y-G, Graham DW. 2019. Understanding drivers of antibiotic resistance genes in High Arctic soil ecosystems. *Environment International* 125:497. <https://doi.org/10.1016/j.envint.2019.01.034>.
  53. Hernández J, González-Acuña D. 2016. Anthropogenic antibiotic resistance genes mobilization to the polar regions. *Infect Ecol Epidemiol* 6:32112. <https://doi.org/10.3402/iee.v6.32112>.
  54. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 108:4516–4522. <https://doi.org/10.1073/pnas.1000080107>.
  55. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vázquez-Baeza Y, González A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, et al. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551:457. <https://doi.org/10.1038/nature24621>.
  56. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena 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. <https://doi.org/10.1038/nmeth.f.303>.
  57. 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. <https://doi.org/10.1038/nmeth.2276>.
  58. Aronesty E. 2013. Comparison of sequencing utility programs. *Tobioij* 7:1. <https://doi.org/10.2174/1875036201307010001>.
  59. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6:610–618. <https://doi.org/10.1038/ismej.2011.139>.
  60. Bokulich NA, Rideout JR, Kopylova E, Bolyen E, Patnode J, Ellett Z, McDonald D, Wolfe B, Maurice CF, Dutton RJ. 2015. A standardized

- extensible framework for optimizing classification improves marker-gene taxonomic assignments; 2167-9843. PeerJ PrePrints 3:e1502.
61. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E, Methe B, DeSantis TZ, Petrosino JF, Knight R, Birren BW. 2011. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 21:494–504. <https://doi.org/10.1101/gr.112730.110>.
  62. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. <https://doi.org/10.1371/journal.pone.0061217>.
  63. Wickham H. 2011. ggplot2. *Wires Comp Stat* 3:180–185. <https://doi.org/10.1002/wics.147>.
  64. Dixon P. 2003. VEGAN: a package of R functions for community ecology. *J Vegetation Sci* 14:927–930. <https://doi.org/10.1111/j.1654-1103.2003.tb02228.x>.
  65. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.
  66. Horton JM, Gao Z, Sullivan DM, Shopsin B, Perez-Perez GI, Blaser MJ. 2015. The cutaneous microbiome in outpatients presenting with acute skin abscesses. *J Infect Dis* 211:1895–1904. <https://doi.org/10.1093/infdis/jiv003>.
  67. Cho I, Yamanishi S, Cox L, Methe BA, Zavadil J, Li K, Gao Z, Mahana D, Raju K, Teitler I, Li H, Alekseyenko AV, Blaser MJ. 2012. Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature* 488:621–626. <https://doi.org/10.1038/nature11400>.
  68. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res* 19:1639–1645. <https://doi.org/10.1101/gr.092759.109>.