


Microbial Community Patterns Associated with Automated Teller Machine Keypads in New York City

 Holly M. Bik,^{a*} Julia M. Maritz,^a Albert Luong,^b  Hakdong Shin,^{b*} Maria Gloria Dominguez-Bello,^b Jane M. Carlton^a

Center for Genomics and Systems Biology, Department of Biology, New York University, New York, New York, USA^a; Human Microbiome Program, New York University School of Medicine, New York, New York, USA^b

ABSTRACT In densely populated urban environments, the distribution of microbes and the drivers of microbial community assemblages are not well understood. In sprawling metropolitan habitats, the “urban microbiome” may represent a mix of human-associated and environmental taxa. Here we carried out a baseline study of automated teller machine (ATM) keypads in New York City (NYC). Our goal was to describe the biodiversity and biogeography of both prokaryotic and eukaryotic microbes in an urban setting while assessing the potential source of microbial assemblages on ATM keypads. Microbial swab samples were collected from three boroughs (Manhattan, Queens, and Brooklyn) during June and July 2014, followed by generation of Illumina MiSeq datasets for bacterial (16S rRNA) and eukaryotic (18S rRNA) marker genes. Downstream analysis was carried out in the QIIME pipeline, in conjunction with neighborhood metadata (ethnicity, population, age groups) from the NYC Open Data portal. Neither the 16S nor 18S rRNA datasets showed any clustering patterns related to geography or neighborhood demographics. Bacterial assemblages on ATM keypads were dominated by taxonomic groups known to be associated with human skin communities (*Actinobacteria*, *Bacteroides*, *Firmicutes*, and *Proteobacteria*), although SourceTracker analysis was unable to identify the source habitat for the majority of taxa. Eukaryotic assemblages were dominated by fungal taxa as well as by a low-diversity protist community containing both free-living and potentially pathogenic taxa (*Toxoplasma*, *Trichomonas*). Our results suggest that ATM keypads amalgamate microbial assemblages from different sources, including the human microbiome, eukaryotic food species, and potentially novel extremophilic taxa adapted to air or surfaces in the built environment. DNA obtained from ATM keypads may thus provide a record of both human behavior and environmental sources of microbes.

IMPORTANCE Automated teller machine (ATM) keypads represent a specific and unexplored microhabitat for microbial communities. Although the number of built environment and urban microbial ecology studies has expanded greatly in recent years, the majority of research to date has focused on mass transit systems, city soils, and plumbing and ventilation systems in buildings. ATM surfaces, potentially retaining microbial signatures of human inhabitants, including both commensal taxa and pathogens, are interesting from both a biodiversity perspective and a public health perspective. By focusing on ATM keypads in different geographic areas of New York City with distinct population demographics, we aimed to characterize the diversity and distribution of both prokaryotic and eukaryotic microbes, thus making a unique contribution to the growing body of work focused on the “urban microbiome.” In New York City, the surface area of urban surfaces in Manhattan far exceeds the geographic area of the island itself. We have only just begun to describe

Received 10 August 2016 **Accepted** 19 October 2016 **Published** 16 November 2016


Citation Bik HM, Maritz JM, Luong A, Shin H, Dominguez-Bello MG, Carlton JM. 2016. Microbial community patterns associated with automated teller machine keypads in New York City. *mSphere* 1(6):e00226-16. doi:10.1128/mSphere.00226-16.

Editor Timothy E. Mattes, University of Iowa

Copyright © 2016 Bik 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 Maria Gloria Dominguez-Bello, Maria.Dominguez-Bello@nyumc.org, or Jane M. Carlton, jane.carlton@nyu.edu.

*Present address: Holly M. Bik, Department of Nematology, University of California, Riverside, Riverside, California, USA; Hakdong Shin, Faculty of Food Science and Biotechnology, College of Life Science, Sejong University, Seoul, South Korea.

 The microbiome of ATM keypads derives from human body sites, food, and airborne sources #microBEnet @hollybik @Genome_Jane

the vast array of microbial taxa that are likely to be present across diverse types of urban habitats.

KEYWORDS: 16S rRNA, 18S rRNA, New York City, automated teller machine, environmental sequencing, urban microbiome

In recent years, the growing accessibility of high-throughput sequencing technologies has vastly expanded our knowledge of microbial communities on a global scale, encompassing both natural and human-made ecosystems. Environmental sequencing studies focusing on conserved, phylogenetically informative genetic loci (e.g., nuclear markers encoding ribosomal subunits, such as the 16S rRNA gene in bacteria/archaea and the 18S rRNA gene in eukaryotes [1]) have enabled rapid detection and description of uncultivated taxa across diverse ecosystems. The majority of studies have sought to describe microbial biodiversity and assess ecological patterns of “pristine” natural habitats such as oceans (2), lakes (3), soils (4, 5), sea ice and glaciers (6), and even clouds (7). More recently, however, a number of studies have aimed at directly assessing human contributions and their impacts on microbial communities. Studies of the built environment have sought to understand how building architecture and engineering, in conjunction with human behavior, may influence the microbes that we encounter during our time spent indoors (which represents 87% of our time on average [8]). Such studies have aimed at capturing the microbial communities associated with both air and surfaces in homes (9, 10), hospitals (11, 12), classrooms (13), offices (14, 15), and restrooms (16), as well as community assemblages within plumbing systems (17–19) and an expanding set of other indoor microhabitats.

In a similar but distinct vein, studies of urban ecosystems have aimed to understand the processes and factors governing microbial communities in metropolitan areas worldwide, where the density of humans is great but the ecosystems themselves are more sprawling and open to the elements. In urban ecosystems, the distribution and diversity of microbes may thus be shaped by a combination of both ecological and human-mediated processes. To date, urban environmental sequencing work has sought to describe microbial communities associated with urban air (20–24), rodents (25), and urban soils in green roofs, city parks, and road medians (26, 27), as well as surfaces within urban transit systems (28, 29). However, those studies captured only a small part of the microbial diversity that is likely to be present in urban environments, and there are many other types of locations, substrates, and surfaces which may serve to deepen our knowledge of microbial ecology and public health in metropolitan areas.

To characterize microbial diversity and biogeography in a unique (but insufficiently studied) component of the urban landscape, we carried out a baseline survey of microbial community diversity associated with automated teller machine (ATM) keypads in New York City (NYC). The geography, population density, and accessibility of demographic data in NYC provide a unique case study—in terms of floor space, the “indoor biome” of Manhattan is three times as large as the geographic area of the island itself (30). With this focus on NYC, we aim to facilitate the assessment of different factors that may govern microbial biodiversity and community structure in specific urban habitats. The ATM keypad can be considered a highly trafficked surface that routinely comes into contact with human inhabitants, similarly to railings, seats, and turnstiles in urban transit systems (28, 29). Several culture-based studies of ATMs around the world have previously been carried out (31–33), but no study to date has yet applied high-throughput sequencing methods to deeply characterize all taxa which may be present on ATM keypads. In the present study, we sampled ATM keypads across eight neighborhoods in three NYC boroughs, where US Census data indicated the presence of demographically distinct populations of local residents (Fig. 1). Our goal was to broadly characterize the microbial assemblages recovered from ATM keypads, using an environmental sequencing workflow that concurrently recovered bacterial/archaeal taxa (16S rRNA gene amplicons) as well as microbial eukaryote communities (18S rRNA gene amplicons). The parallel collection of sample metadata and neighborhood census

Neighborhood Tabulation Area (NTA)	NYC Borough	Dominant Race	Percent Foreign Born	Foreign Born Origin	Dominant Ancestry	Dominant Age Group	16S Samples	18S Samples
MN01 - Marble Hill Inwood	Manhattan	Hispanic/Latino (75.2%)	47.6%	Latin America	Irish	25 to 34	11	7
MN11 - Central Harlem South	Manhattan	Black (56.6%)	22.2%	Latin America	West Indian	25 to 34	7	7
MN17 - Midtown	Manhattan	White (68.2%)	53.6%	Asia	Irish	25 to 34	7	7
MN20 - Murray Hill Kips Bay	Manhattan	White (69.7%)	22.1%	Asia	Italian	25 to 34	2	0
MN27 - Chinatown	Manhattan	Asian (64.5%)	55.5%	Asia	Irish	25 to 34	11	7
QN22 - Flushing	Queens	Asian (66.7%)	69.7%	Asia	Italian	45 to 54	6	6
QN55 - South Ozone Park	Queens	Asian (24.5%)	58.5%	Latin America	West Indian	45 to 54	11	7
BK23 - West Brighton	Brooklyn	White (96.7%)	59.3%	Europe	Russian	65 to 74	11	7

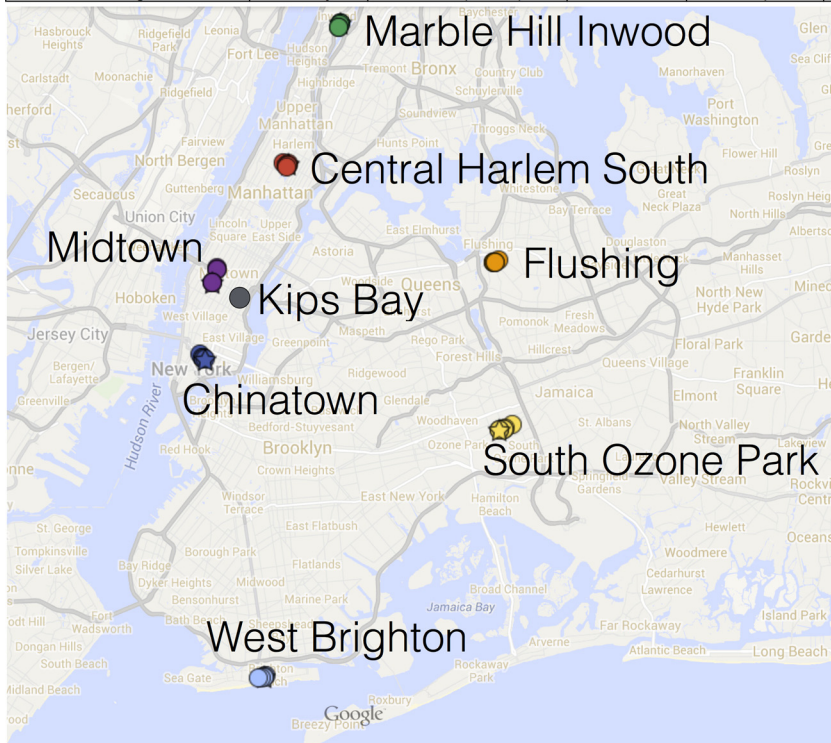


FIG 1 Map and population demographic metadata of sample sites in New York City. Microbial swab samples were collected at automated teller machines (ATMs) in eight neighborhood tabulation areas (NTAs), representing three boroughs of New York City (Manhattan, Queens, and Brooklyn). NTA population demographics, representing 5-year estimates from the United States Census Bureau’s American Community Survey (ACS) (2008 to 2012), were obtained from the NYC open data portal (<https://nycopendata.socrata.com/>). “ancestry” demographics represent write-in responses from a small subset of survey respondents, enabling respondents to report ethnic origins that are not otherwise captured in questions pertaining to race or foreign-born status in the ACS. Age data represent years. (Map data © 2016 Google.)

data also allowed us to assess whether microbial biogeography in NYC was correlated with ATM characteristics, local population demographics, or geographic factors. Statistical analyses were further implemented to determine the potential source of ATM microbial assemblages (e.g., the human microbiome, air, food, etc.) as well as potential biomarkers associated with different sample classes.

RESULTS

Influence of OTUs from control samples. We collected 66 samples from ATM keypads across New York City (8 neighborhoods in Manhattan, Queens, and Brooklyn; Fig. 1), including six control swabs that were exposed to ambient air at different sites. The sampling strategy was designed to target geographic areas with distinct ethnic and population demographics, known as neighborhood tabulation areas (NTAs), defined by the NYC Department of City Planning (see Materials and Methods). The majority of ATMs were sampled from indoor locations (62 samples were taken inside buildings or vestibules); however, a small subset of ATM keypads represented outdoor sample locations (4 samples [included in the 16S rRNA sequencing run only]). During initial analysis of 16S and 18S rRNA datasets, control samples formed a distinct grouping separate from the ATM samples in Unifrac principal-coordinate analyses (PCoAs) (see Fig. S1 in the supplemental material). Upon further investigation, our initial Source-

Tracker analysis revealed that a significant proportion of ATM sequences represented microbial operational taxonomic units (OTUs) that were also present in ambient air controls (Fig. S2 and S3). Examination of OTU tables suggested that the microbial OTUs present in control samples represent a mix of aerial microbes (e.g., fungal species and bacteria attached to dust particles, pollen, etc.), microbes present in the cotton swabs when purchased from the manufacturer, and “kit microbiomes” consisting of microbes derived from laboratory reagents (we recovered many known kit-associated bacterial genera such as *Acinetobacter*, *Pseudomonas*, *Deinococcus*, *Sphingobium*, and *Corynebacterium* [34]) or potential contamination introduced at some point during PCR and sequencing protocols. The OTUs sequenced from blank control samples most likely represent microbes from a combination of these sources; after assessing control samples and conducting SourceTracker analysis, we adhered to stringent data filtering protocols and subtracted all control sample OTUs from the entire data set.

Alpha and beta diversity analyses. Assessment of alpha diversity suggested that the phylogenetic diversity of microbial communities on ATM keypads had been adequately captured by the sequencing workflows in this study (Fig. S4). Rarefaction curves calculated from stringently filtered OTU tables (subjected to abundance-based OTU filtering and subtraction of all control OTUs) were observed to be almost flat (for eukaryotic 18S rRNA data rarefied at 8,900 sequences per sample; Fig. S4B) or to be beginning to level off (for 16S rRNA data rarefied at 1,700 sequences per sample; Fig. S4A) within each of the eight NYC neighborhoods sampled. The differences in the shapes of the rarefaction curves were most likely due to the increased sequencing effort per sample for 18S rRNA amplicons and to the putatively lower phylogenetic diversity of microbial eukaryote taxa in urban environments.

Taxonomy summaries for 16S and 18S rRNA showed that the major taxa recovered from ATM samples were largely consistent across datasets (Fig. 2 and 3). However, the presence or absence and relative abundances of other minor taxonomic groups were much more variable across samples. In the 16S rRNA data set, the most abundant bacterial phyla across most samples were *Actinobacteria*, *Bacteroides*, *Firmicutes*, and *Proteobacteria* (Fig. 2A); these taxa are representative of human skin communities and have been previously shown to dominate urban surfaces in the Boston subway system (28). At the class level, *Actinobacteria*, *Bacilli*, *Clostridia*, *Alphaproteobacteria*, and *Gammaproteobacteria* showed the highest relative abundances across most samples (Fig. 2B). The *Alphaproteobacteria* are considered a widespread and metabolically diverse group of environmental bacteria and have also been shown to be associated with urban transit system surfaces (28). Only nine archaeal OTUs were observed at low relative abundances and restricted to a few samples in the final abundance-filtered OTU tables. Further work is needed to determine whether *Archaea* can truly be considered “rare taxa” on ATM keypads or, alternatively, whether the 16S rRNA primer set used in this study prevented recovery of the majority of archaeal taxa; alternative primer sets or shotgun metagenomic sequencing is needed to provide further insight. In the 18S rRNA data set, fungal OTUs represented the largest taxonomic proportion in most samples, with metazoa and unassigned OTUs (those with no BLAST hit) representing two other taxonomic categories with high relative abundances across most samples (Fig. 3). Protist lineages (*Amoebozoa*, *Alveolata*, *Rhizaria*) had low diversity, were present at much lower relative abundances, and showed more variability across samples. The majority of protist sequences were derived from *Alveolata* and represent free-living ciliates, particularly *Oligohymenophorea* and *Colpodea*, which are taxa commonly found in freshwater and soil habitats (35, 36). Three samples contained a >1% relative abundance of *Entamoeba* species (samples 632, 637, and 661 from ATM keypads in West Brighton and midtown), a genus of protists classed within the phylum *Amoebozoa* and generally associated with the intestinal tract (37). One sample contained a >10% relative abundance of *Silicofilosea* protists (sample 633 from West Brighton), an amoebal member of the *Rhizaria* group known to be associated with bacterivory and fungivory in soil ecosystems (38). Free-living trichomonads (*Monotrichomonas carabina*

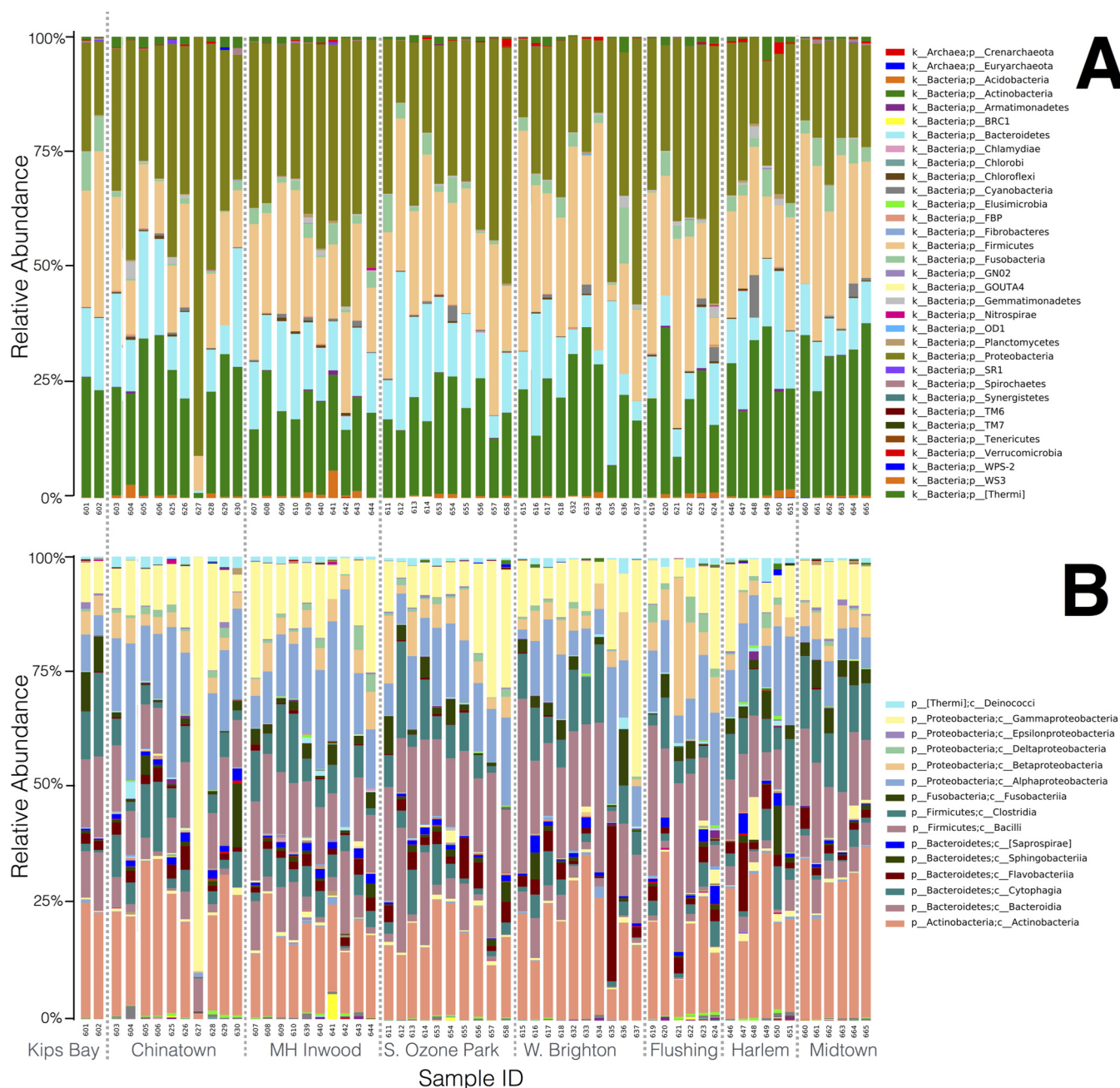


FIG 2 Relative abundances of bacterial/archaeal groups in 16S rRNA data set. (A) Microbial taxonomy summarized at phylum level. (B) Microbial taxonomy summarized at the class level; the legend displays only the top 15 most abundant taxa in the bar chart. Plots were generated in QIIME using abundance-filtered OTU tables with control OTUs subtracted. MH, Marble Hill; S., South; W., West.

and *Ditrichomonas honigbergii* [39, 40]), as well as a gut-associated commensal (*Pentatrachomonas hominis*) typically found in humans and other mammals (41) and a species closely related to the human parasite *Trichomonas vaginalis* that was originally isolated from avian sources (*Trichomonas* sp. strain 5 AP-2012; GenBank accession no. JX512960), were also recovered from ATM keypads. Recent studies suggest that both of these host-associated trichomonads may exhibit zoonotic characteristics (i.e., transmission between humans, domesticated animals, and wildlife) (42). *Toxoplasma*, another zoonotic protist taxon (43), was also detected at >3% relative abundance on one ATM keypad (sample 632 from West Brighton).

Beta diversity analyses of microbial communities revealed a lack of clear patterns across ATM keypads in New York City, and this absence of any obvious groupings was

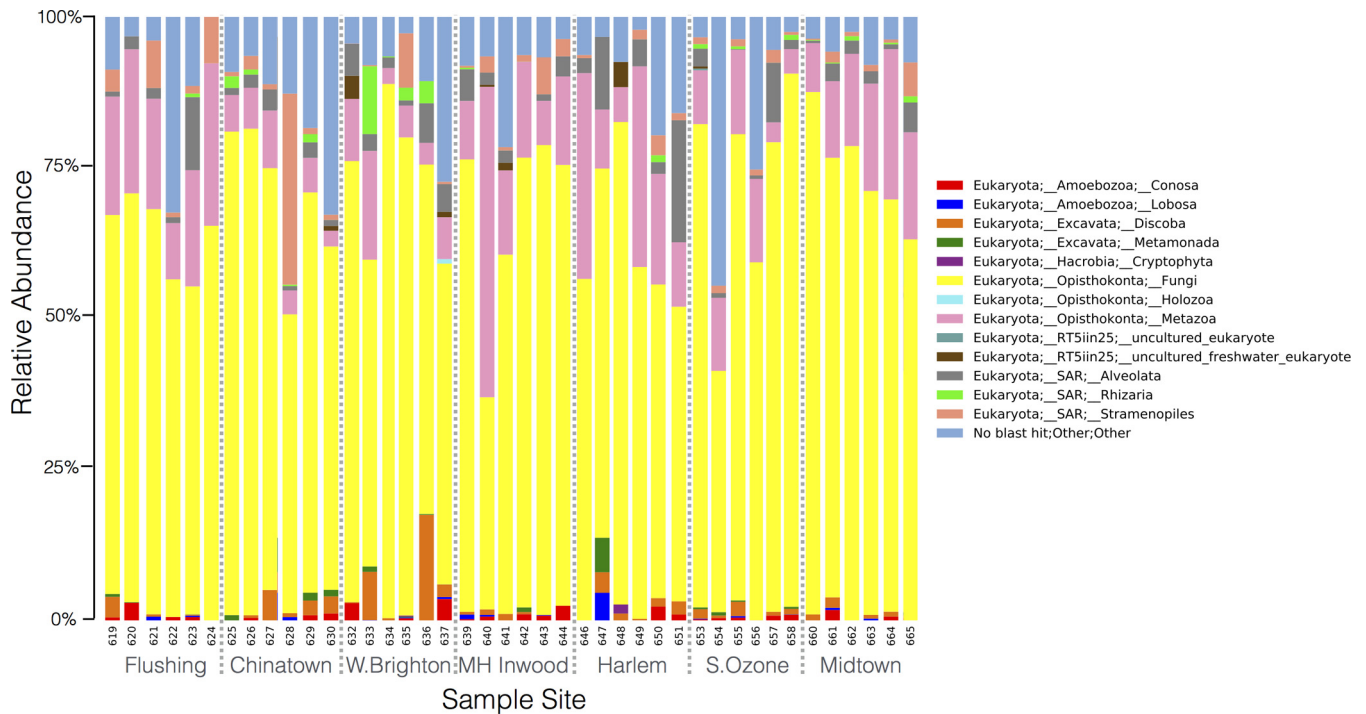


FIG 3 Relative abundances of eukaryotic groups in 18S rRNA data set. Summary of level 3 taxonomy data from the SILVA database, showing higher-level eukaryotic ranks observed in the ATM keypad data set. The plot was generated in QIIME using abundance-filtered OTU tables with control OTUs subtracted.

consistent across both prokaryotic (16S rRNA) and eukaryotic (18S rRNA) datasets (Fig. 4). In weighted and unweighted principal-coordinate analysis (PCoA) using Unifrac distances in QIIME, ATM samples showed no obvious clustering according to geography (neighborhood or borough; Fig. 4A and D), type of site where an ATM was located (bank, store, gas station, etc.; Fig. 4F), or local population demographic metadata obtained from online sources (predominant race group, age group, etc., in each NTA; Fig. 4B and E). Other factors such as date and time of sampling and material of ATM keypad (metal or plastic) also did not reveal any clear clustering of microbial communities (data not shown). Four outdoor ATMs were included in our sample set but were not included in 18S rRNA sequencing; while these four ATMs clustered together in 16S rRNA PCoAs, the corresponding outdoor samples were obtained from the same neighborhood (Chinatown, Manhattan). Small groups of samples from other neighborhoods were also observed to cluster together in our 16S rRNA data set (Fig. 4A), making it impossible to separate the influences of neighborhood and indoor/outdoor ATM location.

Datasets were assessed using various bioinformatic filtering strategies (abundance-based OTU filtering, differing levels of rarefaction), including approaches that both included (Fig. S1) and subtracted (Fig. 4) the microbial OTUs present in blank control samples. None of these methods produced strong groupings in PCoAs, suggesting that the lack of sample clustering across NYC ATMs represents a biologically valid result. However, permutational multivariate analysis of variance (PERMANOVA) tests revealed that the majority of sample groupings in unweighted Unifrac PCoAs are nonetheless statistically significant (Table 1). Borough and neighborhood were found to be statistically significant for both 16S and 18S rRNA datasets; additionally, ATM location (indoor/outdoor), population demographics (race), and Illumina run were statistically significant only in the 16S rRNA data set, and site type was statistically significant only in the 18S rRNA data set. These PERMANOVA results suggest that there may be some subtle differences in microbial community fingerprints across sample groups (e.g., that are revealed only by comparisons of Unifrac phylogenetic distances, as in this statistical

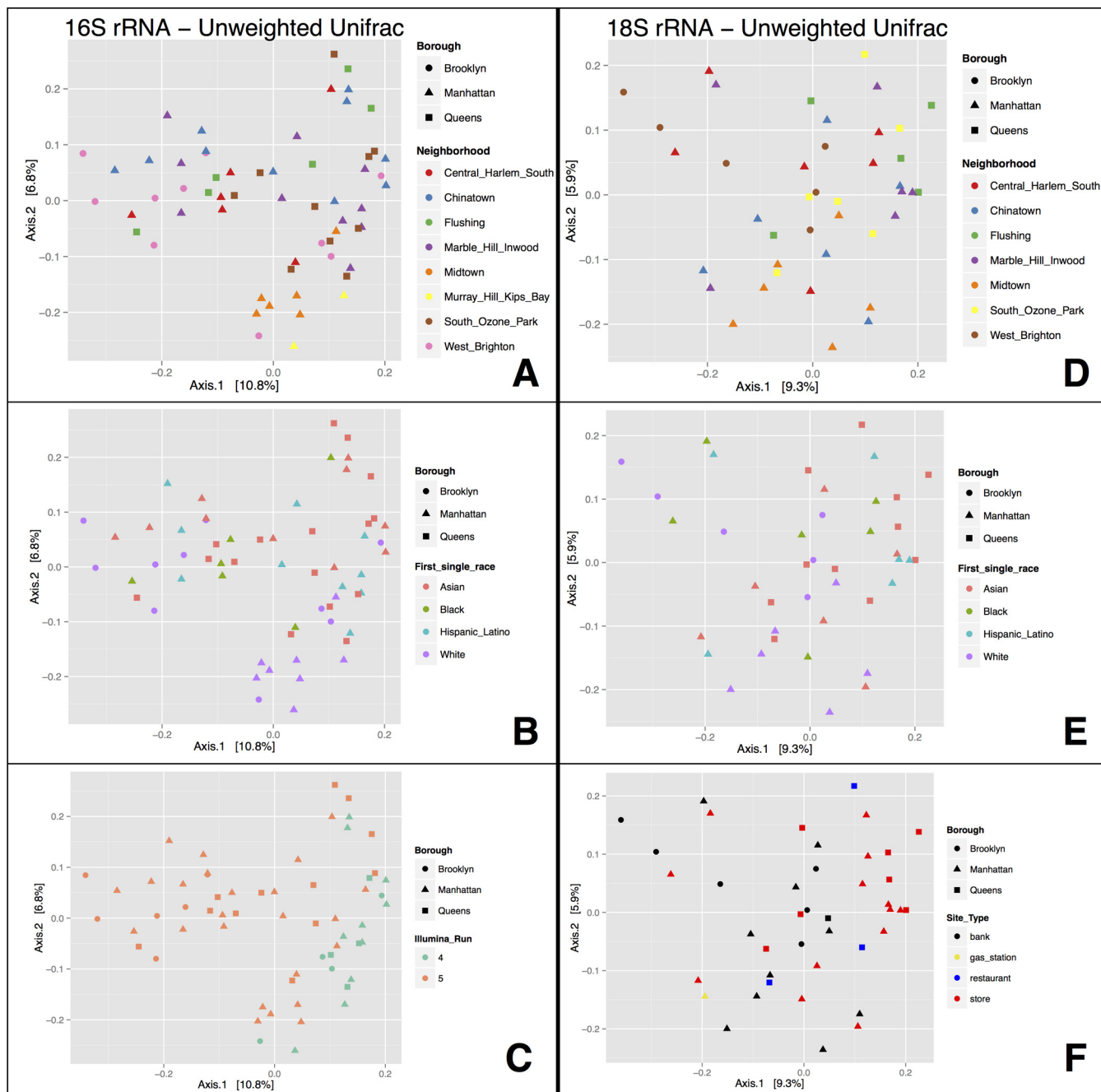


FIG 4 Beta-diversity analyses of microbial taxa recovered from ATM keypads. Data represent results of unweighted Unifrac PCoAs for 16S rRNA for bacteria/archaea (A to C) and 18S rRNA for eukaryotes (D to F), showing no obvious clustering of microbial assemblages according to NYC neighborhood (A and D), census population demographics (race group with highest proportion in each neighborhood) (B and E), or type of site where ATM was located (F). The strongest clustering pattern in the data set was a technical artifact observed for 16S rRNA samples sequenced across two Illumina MiSeq runs (C).

test) that are not otherwise apparent in broader community comparisons such as those performed by PCoAs (Fig. 2).

During analysis, the strongest clustering pattern observed in our data set was a putative technical artifact resulting from 16S rRNA samples being split across two Illumina MiSeq runs (Fig. 4C). Data filtering and rarefaction did not effectively reduce or eliminate this technical artifact, and the Illumina run was found to be a strongly statistically significant sample grouping in PERMANOVA tests of Unifrac distances ($P = 0.0001$; Table 1). The persistence of such a technical artifact has been similarly reported

TABLE 1 PERMANOVA test for statistical significance of sample groupings^a

Category	16S rRNA (Bacteria/archaea)		18S rRNA (eukaryotes)	
	Pseudo-F value	P value	Pseudo-F value	P value
Borough	1.2744	0.0279	1.3959	0.0028
Neighborhood	1.3184	0.0003	1.1880	0.0043
Site type (Bank, store, etc.)	1.0497	0.2593	1.1942	0.0257
ATM location (indoor/outdoor)	1.7337	0.0048	NA	NA
Population demographics (race)	1.3559	0.0038	1.1141	0.1040
Illumina run	3.2439	0.0001	NA	NA

^aStatistical tests were performed on unweighted Unifrac distance matrices (where PCoAs were generated from abundance-filtered OTU tables with control OTUs subtracted), using 10,000 permutations per test. Bold numbers represent significant *P* values of <0.05. Pseudo-F numbers represent *F* values by permutation.

in other recent studies (14). However, in our case, the 18S rRNA amplicons from eukaryotic communities were sequenced on a single Illumina run and thus provided an independent assessment of PCoA patterns. Eukaryotic PCoAs did not show any large differences in the microbial patterns by site (Fig. 4D and F).

Source of microbial communities on ATM keypads. SourceTracker analysis carried out on the 16S rRNA data set indicated that the majority of microbes on each ATM keypad were derived from an unknown source (Fig. 5). For the majority of ATM samples, <25% of the microbial community was assigned to an identified source, although four samples (samples 627, 635, 646, and 649) were shown to have >50% of microbes assigned to known sources. Although this study included only four ATMs located outdoors (gold stars in Fig. 5), ATM location did not seem to influence the proportion of “unknown” sources for microbial OTUs, which was high across most samples. The most common identified sources of microbes on ATM keypads appeared to be household surfaces such as televisions, restrooms, kitchens, and pillows. In our SourceTracker analysis, we included data previously obtained from human hands and palms (44) as well as from other body sites such as the nose, ear, and gut (45). However, human skin and other body sites were not identified as dominant sources of ATM microbes in our SourceTracker analysis, despite the inclusion of 46 samples from two studies representing human skin (44, 45).

Microbial biomarkers from LEfSe analysis. Linear discriminant analysis (LDA) effect size (LEfSe) analysis suggested the presence of a number of significant microbial biomarker taxa across different sample groupings (Table 2). In 16S rRNA datasets, geographic location (borough/neighborhood) represented the only sample grouping exhibiting no significant enrichment or depletion of microbial taxa. However, this pattern was not consistent for eukaryotes, where 7 to 36 biomarker taxa were attributed to geographic location in the 18S rRNA data set. Among all sample groups, the highest numbers of biomarker taxa were reported for “location” of ATMs within the 16S rRNA data set (keypads located indoors versus outdoors; Fig. S5); however, the type of site (bank, restaurant, gas station, etc.) and the population demographics associated with each NTA (predominant race group) were also associated with a low number of biomarker taxa in both 18S and 16S rRNA data (Fig. 6). Fungi comprised the majority of eukaryotic biomarkers identified in the 18S rRNA data set (Fig. 6A) and included both common species and specialized taxa. For example, the fungal species *Aspergillus niger* and *Occultifur externus* were both found to be enriched on ATM keypads sampled in Central Harlem South. *Aspergillus* species are ubiquitous and widespread fungal species associated with outdoor and indoor air (46, 47), and *A. niger* is a species heavily utilized in industrial processes (48) and the cause of black mold disease in many fruit and vegetable crops (49). In contrast, *O. externus* is a recently described novel species originally isolated from plant litter in Portugal (50). Furthermore, the xerophilic food-borne mold *Xeromyces bisporus* was reported as another fungal biomarker in NYC neighborhoods with predominately white population demographics (Fig. 6A); this

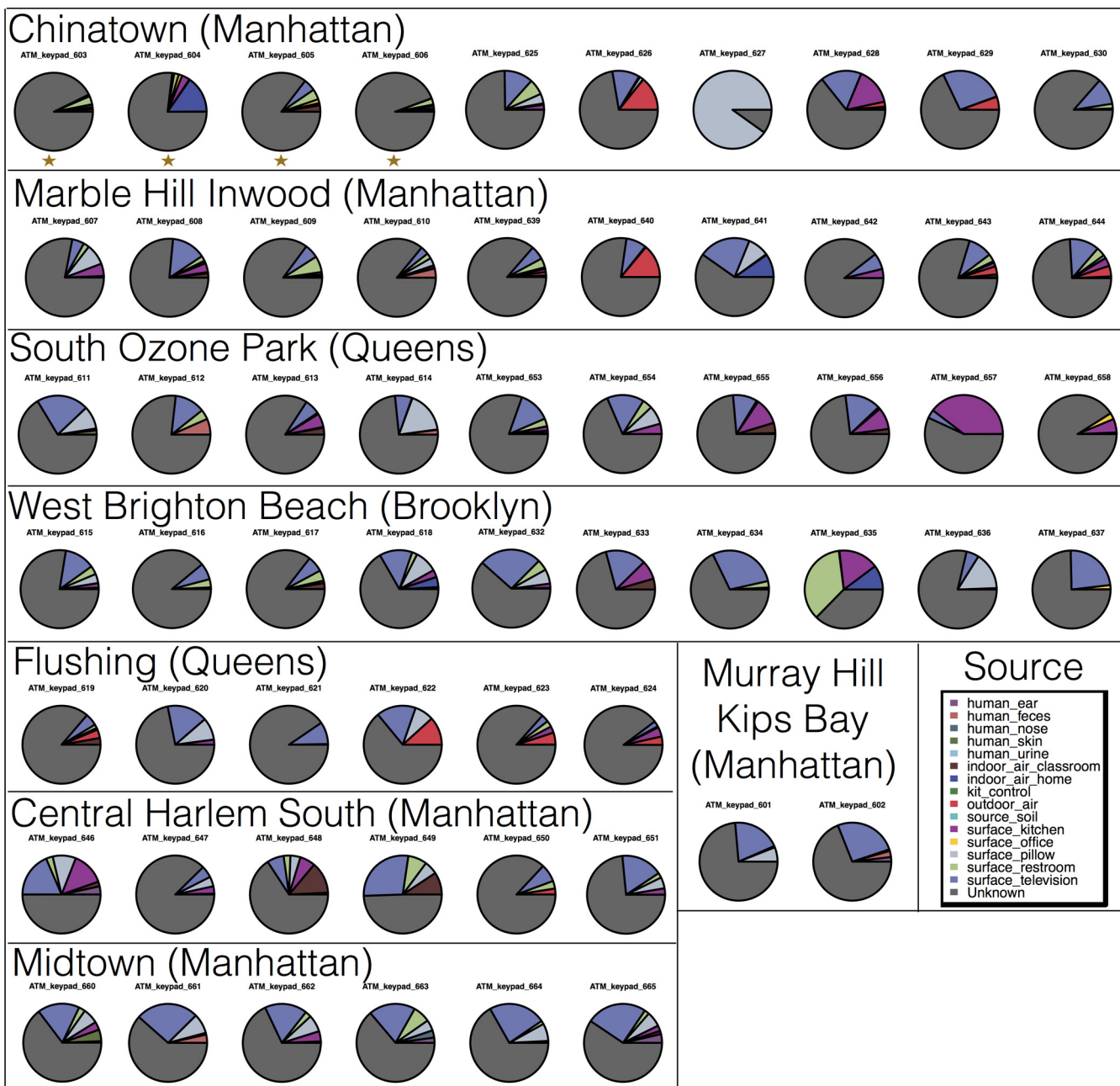


FIG 5 SourceTracker analysis of bacterial/archaeal assemblages on ATM keypads. Closed-reference OTUs (16S rRNA only) from this study were compared to 12 published datasets representing a range of potential source habitats (human body, building surfaces, indoor/outdoor air). The majority of microbes on each ATM keypad were derived from an unknown source. The most common identified source across all ATMs appeared to be household surfaces (rest room, kitchen, pillows, and televisions) and outdoor air. Gold stars denote the four ATMs in this study located at outdoor sites.

fungal species has been reported to grow at extremely low water activity levels that are lower than those seen with any other known organism (51).

A number of eukaryotic metazoan taxa appeared to be associated with population demographics in different NYC neighborhoods. Notably, bony fish (*Teleostei*) and molluscs (*Bivalvia*) were significantly enriched in ATM samples obtained from predominantly Asian neighborhoods (Flushing/Chinatown), while chickens (*Gallus gallus*) were significantly enriched in ATM samples obtained from a predominantly black neighborhood (Central Harlem South). In the 16S rRNA data set, ATM keypads located in laundromats and stores exhibited the highest number of biomarker taxa, with *Lactobacillales* significantly enriched across both site types (Fig. 6B). Overall, the number of

TABLE 2 Number of significantly discriminative taxa reported in LefSe analysis (absolute LDA score, >2.0)^a

Category	No. of significantly discriminative taxa					
	16S rRNA (Bacteria/archaea)			18S rRNA (eukaryotes)		
	L5 taxa (family)	L6 taxa (genus)	OTUs	L5 taxa (family)	L6 taxa (genus)	OTUs
Borough/neighborhood	0	0	0	7	11	36
Site type (Bank, store, etc.)	2	9	3	3	5	25
ATM location (indoor/outdoor)	93	235	148	NA	NA	NA
Population demographics (race)	1	2	1	3	6	16

^aLefSe analyses were performed on normalized BIOM tables from open reference OTU picking, following abundance-based filtering and removal of OTUs present in kit control samples. LefSe analyses were performed on OTU tables summarized at the L5 (family) and L6 (genus) taxonomy levels, as well as on unsummarized OTU tables. NA, ATM location comparisons were not possible for 18S rRNA, as only indoor ATMs were included in the eukaryotic data set.

significantly discriminative taxa was very low in LefSe results from the 16S rRNA data set, with oftentimes only one enriched taxon reported per metadata class (Table 2).

DISCUSSION

Here we present the first broad assessment of microbial communities associated with ATM keypads in New York City, characterizing assemblages of bacteria/archaea (66 samples) and microbial eukaryotes (48 samples) from eight NYC neighborhoods across Brooklyn, Queens, and Manhattan. This data set represents an important addition to the growing body of research focused on urban microbial ecology and specifically complements work performed in New York City which to date has focused on green roofs and park/median soils (26, 27), sewage (J. M. Maritz, K. H. Rogers, T. M. Rock, N. Liu, S. Joseph, K. M. Land, and J. M. Carlton, submitted for publication), rodents (25), and both air and surfaces within the NYC subway system (23, 24, 29). The results are of particular

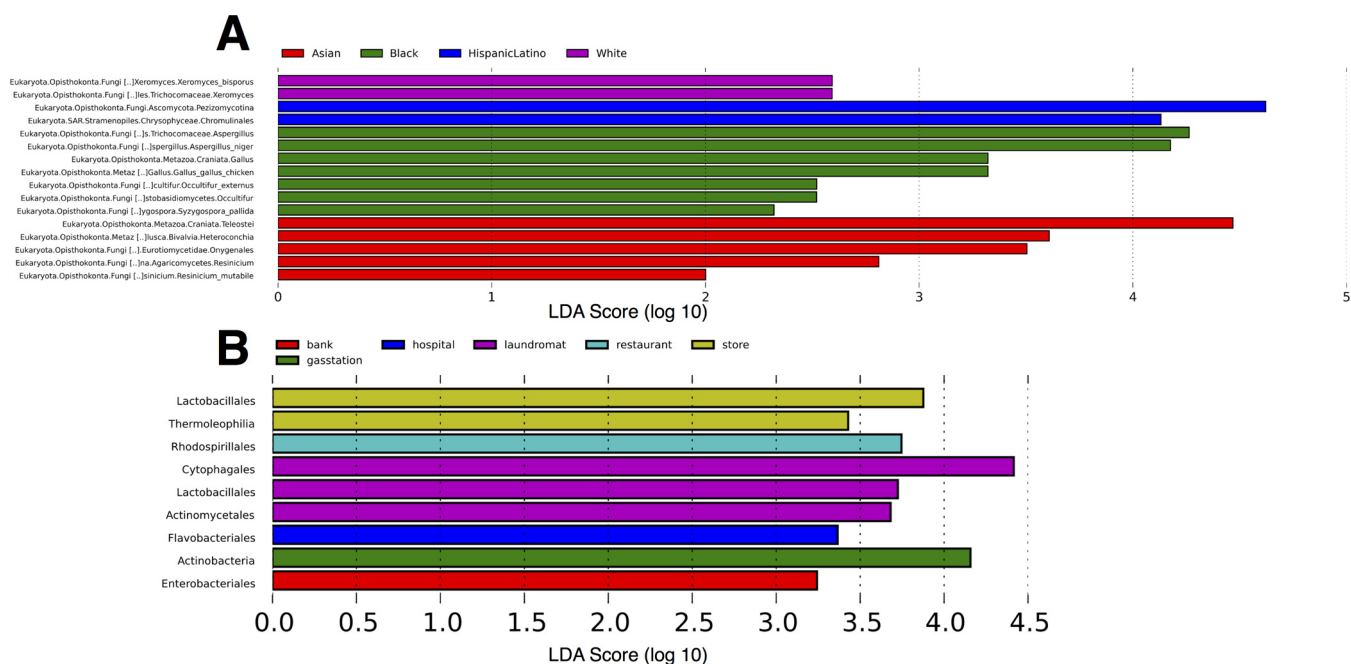


FIG 6 Linear discriminant analysis (LDA) effect size (LefSe) analysis to determine microbial biomarker taxa across sample groups. (A) Eukaryotic 18S rRNA OTUs significantly enriched across census population demographics (predominant race group in each NTA). (B) Bacterial/archaeal genera significantly enriched across different ATM site types in 16S rRNA data set.

relevance with respect to humans, since the surfaces studied are touched by people and could potentially mediate interpersonal transfer of microbes or microbial DNA.

Unifrac principal-coordinate analysis indicated an overall lack of biogeographic patterns structuring microbial communities on ATM keypads in NYC (Fig. 4). This lack of any obvious pattern was consistent across both bacteria/archaea (16S rRNA) and microbial eukaryote (18S rRNA) datasets, suggesting that microbial community structure is not governed by any of the broad metadata categories (e.g., geography, population demographics, site type, data/time of sample collection, etc.) that we assessed during this study. The absence of biogeographic patterns could be explained by a number of factors. ATMs are subject to high use in urban areas such as NYC and could be subject to human-driven homogenization of the microbial communities present on keypads. In any given neighborhood, transient users of ATMs (tourists, commuters, visitors from other NTAs, etc.) might be common and might reduce/eliminate any specific microbial community signatures which might be associated with the population demographics in a given NTA (e.g., those related to age group, ethnicity, etc.). Furthermore, periodic cleaning or disinfection of ATMs, if implemented, may severely reduce the microbial diversity and prevent unique assemblages from accumulating or differentiating on ATM keypads across space and time.

Despite the lack of distinct clusters in Unifrac PCoAs (Fig. 4), LEfSe analyses identified a number of microbial biomarkers indicative of certain metadata classes in both 16S and 18S rRNA datasets. Across all metadata categories in the 18S rRNA data set, LEfSe analyses reported a large number of fungal biomarker taxa (e.g., Fig. 6). This suggests that localized enrichment of some fungal taxa—for example, enrichment of specific taxa that might represent a small fraction of the microbial community and thus might not represent strong enough enrichment to allow differentiation of overall microbial communities on Unifrac PCoAs—may be driving subtle biogeographic patterns in microbial eukaryote communities on ATM keypads. Previous studies have reported geographic partitioning and localization of urban fungal assemblages in NYC soils (27). It is unclear whether the fungal biomarkers on ATM keypads represent truly localized fungal diversity, or alternatively, stochastic enrichment of human-transported or air-borne taxa.

In the eukaryotic data set, the most striking biomarkers appear to indicate a “molecular echo” (52) of food species on ATM keypads in certain neighborhoods (the domestic chicken *Gallus gallus* in Central Harlem South and bony fish [*Teleostei*] and mollusk [*Bivalvia*] species in Chinatown/Flushing; Fig. 6), potentially reflecting the concentrations of specialized restaurants in different areas of NYC. These food species appeared in the LEfSe results as significantly enriched OTUs within each respective neighborhood. While our study design does not allow us to pinpoint the exact source of such DNA, one reasonable explanation is that residual DNA from a recent local meal may persist on a person’s hands and be transferred to the ATM keypad upon use.

In addition to obvious species, another potential food biomarker is the fungal species *X. bisporus*, which was significantly enriched in midtown and other NYC neighborhoods with predominantly white population demographics. *X. bisporus* is a foodborne mold originally isolated from licorice and associated with spoilage of high-sugar foods such as cakes and confectionaries (51). Although we cannot confirm the original source of *X. bisporus* OTUs on ATM keypads, it seems plausible that this fungus had been transferred from people who had recently handled baked goods, particularly in a commuter-heavy area such as midtown Manhattan, where there are many nearby convenience stores and cafés selling this type of food product to business workers. A previous metagenomic analysis of the NYC subway system also detected food signatures across surfaces in the urban transit system (29), suggesting that genomic material from meals is routinely transferred around NYC by human inhabitants and may thus represent a common component of the urban microbiome.

The human microbiome represents an obvious source of microbial communities on ATM keypads; however, SourceTracker analysis did not pinpoint human skin or any other body site as a primary source of microbes (Fig. 5). In contrast, outdoor air and

household surfaces—kitchens, restrooms, pillows, and televisions—were the most commonly identified source habitats. Household surfaces may effectively collect microbial communities from various sources (food, family members, pets, dust) and thus represent a pool of microbes originating from different habitats. This is in contrast to swab samples collected directly from human body sites (see, e.g., the published datasets [44, 45] used as human sources in this study), which represent the personalized microbiome associated with a single person. Since each ATM keypad in New York City is most likely utilized by hundreds of people each day (and may come into contact with air, water, and microbes from different urban surfaces), the microbial communities obtained in this study may represent an “average” community that is effectively pooled from vastly different sources (14). An alternative and potentially more plausible hypothesis is that the SourceTracker algorithm may be highly sensitive to the primer region and sequencing platform used to generate the “source” training sequences. The samples from the household surfaces representing the majority of assigned sources for ATM keypads were generated using the same primer set and sequencing technology utilized in this study (Illumina HiSeq/MiSeq data using the 515F/806R primers to amplify the 16S rRNA gene [10, 53]), which may explain why these household surfaces were identified as source habitats. In contrast, human microbiome “source” samples were generated using different primer sets and sequencing platforms (e.g., Roche 454), and the distinct rRNA region and lower-throughput nature of the sequencing technology may confound the ability of SourceTracker to assess source/sink habitats between these samples. Overall, the vast majority of ATM microbial communities were derived from an “unknown” source, and it is also possible that this unassigned community fraction represents human-associated microbes that SourceTracker was unable to recognize.

Caveats. In the present study, we encountered a number of issues that confounded data analysis and were challenging to circumvent. Technical artifacts in the 16S rRNA data set were obvious and persistent (Fig. 4C); future studies should aim to sequence all samples on a single Illumina run, in order to avoid the introduction of technical artifacts that may confound data analysis. Low-diversity microbial communities were unexpectedly recovered from “control” samples, requiring stringent data filtering to remove all potential contaminant OTUs. The collection protocol for ambient air control samples may have inadvertently collected species from airborne dust; alternatively, the manufactured swab samples we utilized may not have been entirely sterile. Regardless, our stringent data filtering protocol ultimately resulted in a significant reduction of sequences per sample—representing a level of coverage that was sufficient for microbial ecology analyses but a sequencing depth that was far from ideal (particularly for the 16S rRNA amplicons which were not sequenced as exhaustively as those of the eukaryotic 18S rRNA). Thus, our analyses may have failed to detect some microbial diversity and community patterns that might only have become apparent with extremely deep sequencing of ATM keypads (e.g., rare biosphere biomarker taxa), especially in lower-coverage bacterial/archaeal samples.

Future work. The present study aimed to solely characterize microbial diversity (living or dead) on ATM keypads; we had no way of measuring levels of active versus dormant microbes or of assessing what could be considered “residual” or “transitory” species aside from obvious food species (e.g., microbes/fungi transported through air/dust). Additional, complementary analyses will be needed to determine whether the detected species are metabolically active and whether some species can survive on ATM keypads for extended periods of time. Although we detected DNA signatures from potential pathogens on ATM keypads, more-targeted studies will be needed to confirm the source, viability, and distribution of such pathogenic species on urban surfaces. For example, the human pathogen *Trichomonas vaginalis* cannot be differentiated from closely related zoonotic species using 18S rRNA loci alone (Maritz et al., submitted) and in this study we could not confirm the likely source of eukaryotic pathogen species. Future studies of the urban microbiome should also include an expanded examination of ATMs in both indoor and outdoor locations. LEfSe biomarkers in our 16S rRNA data

set (see Fig. S5 in the supplemental material), as well as previous studies of urban transit systems (28), have indicated that indoor/outdoor location is a significant factor structuring microbial communities. Depletion/enrichment of indoor and outdoor biomarkers may be especially relevant for eukaryotic taxa, especially for pollen and airborne fungal spores. Unfortunately, we were unable to assess such patterns in the present study, as the subset of eukaryotic 18S rRNA data included only samples from indoor ATMs. In addition, there is a need for expanded studies that investigate shifts in ATM microbial communities over time, as well as for more-intense sampling covering a wider geographic area (including sample collection across other cities worldwide). Our study provided only a small snapshot of microbial communities, in terms of time points and the number of ATMs sampled. Nonetheless, the present data set has provided a significant baseline for microbial diversity on ATM keypads, broadly encompassing two taxonomic domains. The detection of microbial biomarker taxa in this study provides a jumping-off point for future study design and hypothesis testing, hinting at potential large-scale trends that may influence the distribution and persistence of microbes on highly trafficked urban surfaces.

MATERIALS AND METHODS

Sampling and metadata collection. Microbial swab samples were collected during June and July 2014 at automated teller machines (ATMs) in eight neighborhood tabulation areas (NTAs) of New York City, representing three boroughs (Manhattan, Queens, and Brooklyn; Fig. 1). NTAs are geographic units designated by the Department of City Planning that are used to project population size and demographics at the small-area level (<http://www1.nyc.gov/site/planning/data-maps/open-data/dwn-nynta.page>). NTAs are units with a minimum population size of 15,000 and are used to collect and collate census/survey data; they do not necessarily reflect historical boundaries of NYC neighborhoods. In this study, sampling locations were preselected based on known neighborhood demographics, and ATM keypads were sampled randomly within areas proximate to a subway station within each NTA.

ATMs were sampled using sterile cotton swabs individually packaged in pairs (Covidien cotton-tipped applicators; Fisher Scientific catalog no. 22-037-924) and premoistened with 0.15 M NaCl solution–0.1% Tween 20. Each ATM keypad was sampled using 2 cotton swabs at a time, scrubbing all keys for a total of 10 s. A total of six control swabs were collected across different NTAs. For each control sample, a cotton swab was removed from sterile packaging, dipped in buffer solution, and held in ambient air for 10 s.

At each sampling location, the following metadata were also recorded: date, time, neighborhood (NTA), NYC borough, type of ATM keypad (metal or plastic), ATM location (indoor or outdoor), and site type (hospital, bank, convenience store, etc.). Global Positioning System (GPS) coordinates were not collected during sampling, in order to anonymize the location of all ATMs used in this study.

In order to analyze microbial community patterns in conjunction with NYC neighborhood demographics, additional metadata about each sampling area were collated from online sources. Demographic information for each NTA was obtained from the New York City Open Data portal (Fig. 1; public data set: <https://data.cityofnewyork.us/City-Government/Demographics-and-profiles-at-the-Neighborhood-Tabu/hyuz-tij8>), where all population information is derived from 5-year estimates (2008 to 2012) from the United States Census Bureau's American Community Survey (<https://www.census.gov/programs-surveys/acs/>). For this study, we selected a subset of metadata that was potentially relevant to the assessment of microbial community patterns on ATM keypads, such as population density, and indicators of genetic factors, lifestyle, or habits that may be influenced by cultural or age demographics of the population within each NTA. The following broad population statistics were obtained for each NTA: population recorded during the 2000 United States Census; population recorded during the 2010 United States Census; and population change (both number and percent change) between the 2000 United States Census and the 2010 United States Census. An additional set of population demographics related to race, ancestry, and age was also obtained for each NTA as follows: percentage of the population who reported belonging to a single race; top three race groups comprising the largest numbers and percentages of the population; top three age groups comprising the largest percentages of the population (quantified by age group and percentage); percentage of foreign-born residents; geographic origin of the primary foreign-born group (quantified by race); and predominant ancestry group (self-reported ethnic origin recorded for a small subset of the population).

DNA extraction, PCR amplification, and sequencing. We extracted DNA using a MoBio PowerLyzer PowerSoil extraction kit (catalog no. 12855-100) according to manufacturer instructions. One cotton swab from each sample site was clipped and placed into an individual well of a 96-well bead plate for DNA extractions (the second swab from each sample site was retained in frozen storage for future studies). For each ATM sample, the same environmental DNA extraction method was used to generate PCR amplicons for both the 16S rRNA and 18S rRNA genes.

The V4 fragment of the 16S rRNA gene was amplified using the 515F/806R primer set (54). All primers and protocols used for amplification and sequencing represent standardized workflows obtained from the Earth Microbiome Project (EMP) website (<http://www.earthmicrobiome.org/emp-standard-protocols/16s/>). Amplification was done in triplicate following the EMP protocol. After amplification, reactions were

quantified using a Pico Green (Invitrogen) assay kit, equal amounts of amplicons were pooled, and the final pool was cleaned using a QIAquick PCR purification kit (Qiagen). Amplicons were sequenced using a paired-end Illumina MiSeq platform, as previously described (54).

The V9 fragment of the 18S rRNA gene was amplified using Illumina primer constructs containing the universal primers 1391f-EukBr. Library synthesis and amplification using 2 μ l of input DNA were done in triplicate following the EMP protocol, available online (<http://www.earthmicrobiome.org/emp-standard-protocols/18s/>). After amplification, triplicate PCRs were pooled prior to purification and quantification and prepared for sequencing following the protocol described by Maritz et al. (submitted for publication). Purified libraries that showed high proportions of adapter dimers were size selected using a 2% agarose dye-free gel on a BluePippin instrument (Sage Science).

All barcoded rRNA libraries were subsequently sequenced on an Illumina MiSeq platform (2-by-150 paired-end reads, with a 5 to 10% PhiX spike-in, based on the strategy outlined in reference 54). The 16S rRNA samples were sequenced on two separate MiSeq runs (as they represented two distinct sampling time points), while all 18S rRNA samples were pooled and sequenced together on a single MiSeq run (2-by-100 paired-end reads with a 6% PhiX spike-in).

Data filtering and processing. For both the 16S rRNA and 18S rRNA datasets, the majority of data filtering and processing steps were carried out using QIIME v1.8 (55). To process raw Illumina data, paired-end reads were merged using `join_paired_ends.py`, a minimum overlap of 10 bp, and a 15% error rate in the overlapping bases. Joined reads were subsequently demultiplexed using `split_libraries_fastq.py` with the `rev_comp_mapping_barcodes` flag; a minimum Phred quality score of 20, allowing 5 poor-quality bases before read truncation; and an 0.70 minimum fraction of consecutive high-quality base calls to include reads. Merged sequences that passed quality filtering thresholds were subsequently clustered into operational taxonomic units (OTUs) using the `pick_open_reference_otus.py` workflow (56) with 10% subsampling and *de novo* clustering of failure reads. Any resulting singleton OTUs were discarded (the minimum cluster size was set at 2 reads). A 97% clustering cutoff was used for 16S rRNA OTU picking, representing a standard approach in microbial ecology studies of bacteria/archaea. A more stringent 99% clustering cutoff was used for 18S rRNA OTU picking, as the 18S rRNA gene is typically more conserved and less variable in eukaryotic genomes (57) and because the use of a higher clustering cutoff is also in line with many comparable environmental sequencing studies of eukaryotes (58). For OTU picking, initial reference-based OTU clustering was carried out against Greengenes 13_8 (97% OTUs) for 16S rRNA data (59) and SILVA 119 (99% OTUs) for 18S rRNA data (60). Summaries of OTUs and quality-processed reads obtained across samples are provided in Fig. S6 (16S rRNA) and Fig. S7 (18S rRNA) in the supplemental material, and a more detailed record of demultiplexing and processing is provided in Table S1 in the supplemental material.

For 16S rRNA OTUs, taxonomy was assigned to representative sequences using QIIME's `uclust` consensus taxonomy assigner and Greengenes rRNA database version 13_8 (97% OTU representative sequences). For 18S rRNA OTUs, taxonomy was assigned in two steps (based on methods described in reference 61). First, BLAST was used to compare OTU representative sequences to a manually curated version of the SILVA 111 database, containing only eukaryotic sequences. In this curated database, taxonomic hierarchies were standardized and corrected, and some protist sequences were manually added to improve representation of some groups (Maritz et al., submitted). The taxonomy from the top BLAST hit was taken for any OTUs matching the curated database. Second, any OTUs without a BLAST hit were subsequently compared to the entire SILVA 111 database (99% OTU representative sequences) containing reference sequences from all three domains. The taxonomy from the top BLAST hit was again accepted for any OTUs matching the reference database, and failed sequences were denoted as "unassigned." In both taxonomy assignment steps, the minimum E value cutoff for BLAST searches was set at $1e-20$, and the top BLAST hit was taken for each representative OTU sequence; sequences were denoted as "unassigned" if they did not return any results meeting this E value cutoff criterion.

Representative OTU sequences from all ATM samples were subsequently aligned using the PYNAST aligner (62) and the same reference databases used for OTU picking. Chimera checking was also carried out using ChimeraSlayer (63) for 16S rRNA and 18S rRNA data (a database-dependent method) and additionally using the Blast Fragments method in QIIME for 18S rRNA data (a database-independent chimera checking method, given the sparse nature of 18S rRNA databases and the less variable nature of eukaryotic rRNA genes).

Before any microbial community diversity analyses and comparisons were carried out, the initial OTU tables resulting from open-reference OTU picking were filtered to remove unwanted and poor-quality sequences. First, nontarget OTUs were filtered from each rRNA data set on the basis of taxonomic assignments. Chloroplast, mitochondrial, and "Unassigned" sequences were removed from the 16S rRNA data set. Bacterial, archaeal, and *Archaeplastida* sequences were removed from the 18S rRNA data set. Next, any OTUs that failed to align to reference rRNA databases or were flagged as chimeric were removed. OTU tables were subsequently subjected to abundance-based filtering, removing low-abundance OTUs representing less than 0.0005% of total reads in the data set (using the `min_count_fraction` 0.000005 flag in QIIME for `filter_otus_from_otu_table.py`, following the methods described in reference 64). Finally, initial exploration of data and SourceTracker analysis (see below) suggested that a large proportion of environmental OTUs were derived from blank samples sequenced as controls. For maximal stringency, we additionally removed all OTUs that appeared in control samples, resulting in significant data reduction for both 16S and 18S rRNA (Table S1).

Microbial community analyses. For representative sequences of OTUs that satisfied all filtering criteria, phylogenetic trees were constructed from gap-filtered alignments using the FastTree algorithm (65) with default parameters. Alpha diversity metrics, including rarefaction curves and taxonomy sum-

maries, were calculated using the `alpha_rarefaction.py` script in QIIME. Beta diversity analyses, including weighted and unweighted Unifrac principal-coordinate analyses, were carried out using the `beta_diversity_through_plots.py` workflow script in QIIME. For both alpha and beta diversity analyses, OTU tables were rarefied on the basis of the sample with the lowest number of total sequence reads (the minimum rarefaction levels were 1,700 sequences for 16S rRNA data and 8,900 sequences for 18S rRNA data). Distance matrices from Unifrac PCoAs were imported into R studio and visualized using the `phyloseq` (66) and `ggplot2` packages. To further test for dissimilarity between sample groups in unweighted Unifrac beta diversity analyses, we carried out nonparametric PERMANOVA tests (using `compare_categories.py` in QIIME) and assessed intragroup and intergroup distances using nonparametric *t* tests (`make_distance_boxplots.py` in QIIME).

The 16S rRNA data set was additionally subjected to SourceTracker analysis (67), in order to identify the potential origin of microbial communities on ATM keypads. OTUs obtained from this study were compared to 12 public datasets compiled for a previous meta-analysis study (68). The chosen published datasets represent a range of potential source environments, including the human body (44, 45, 69, 70), surfaces in the built environment (10, 15, 16, 53, 71), soils (72), and indoor/outdoor air samples (13, 73). Demultiplexed 16S rRNA sequences from ATM keypads were subjected to closed-reference OTU picking (using the `pickTheseOTUs.sh` script as described in reference 68), and the resulting OTU table was merged with closed-reference tables from published studies (downloaded from GitHub: https://github.com/jfmeadow/BEMAFinalAnalysis/tree/master/individual_biom_tables). Closed-reference OTU picking was necessary for SourceTracker analysis, as this database-dependent approach avoids potential erroneous/chimeric OTUs and further enables the concurrent analysis of datasets obtained using different 16S rRNA primer sets, read lengths, and sequencing platforms. Among the ATM samples in this study, blank controls were also marked as potential sources of microbial communities in order to identify any potential contamination introduced during PCR and sequencing. Unfortunately, SourceTracker analysis was not possible for our eukaryotic data set because of the comparative lack of published 18S rRNA studies from urban and built environments, database limitations, and the prevalence of other markers used for other eukaryotic studies (internal transcribed spacer [ITS] rRNA, etc.).

Linear discriminant analysis (LDA) effect size (LEfSe) analysis (74) was carried out on both 16S and 18S rRNA datasets, to assess whether any microbial biomarkers could be identified across sample categories. LEfSe analyses were carried out using an online Galaxy server (<http://huttenhower.sph.harvard.edu/galaxy/>). Per-sample normalization was carried out on all input OTU tables, and LDA effect size was calculated using the strict “all-against-all” strategy for multiclass analysis and default parameters and thresholds.

Accession number(s). Raw Illumina sequence data for both 16S and 18S rRNA genes have been deposited in the NCBI SRA under BioProject [PRJNA330663](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA330663) (SRA accession no. [SRP079707](https://www.ncbi.nlm.nih.gov/sra/SRP079707)). Additional documentation of all QIIME workflows, parameters, and OTU picking outputs has been compiled and deposited in FigShare ([10.6084/m9.figshare.3498206](https://www.figshare.com/figure/10.6084/m9.figshare.3498206)).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/mSphere.00226-16>.

Table S1, PDF file, 0.1 MB.

Figure S1, PDF file, 0.1 MB.

Figure S2, PDF file, 0.2 MB.

Figure S3, PDF file, 0.1 MB.

Figure S4, PDF file, 0.4 MB.

Figure S5, PDF file, 1.7 MB.

Figure S6, PDF file, 0.2 MB.

Figure S7, PDF file, 0.2 MB.

ACKNOWLEDGMENTS

We thank Steven Sullivan for providing comments and edits on a draft of the manuscript. We thank the personnel of the NYU Center for Genomics and Systems Biology GenCore for their sequencing services.

J.M.C. and M.G.D.-B. designed the study. A.L. collected swab samples from ATM keypads, performed DNA extractions, and carried out PCR and sequencing of 16S rRNA genes. J.M.M. carried out PCR and sequencing of 18S rRNA. H.M.B., J.M.M., and H.S. performed computational analyses of 16S and 18S rRNA data. J.M.C. and M.G.D.-B. provided support during sequencing and data analysis. H.M.B., J.M.M., and J.M.C. wrote the manuscript.

J.M.M. is supported by the MacCracken Program in the Graduate School of Arts and Science at New York University. This study was partially funded by a New York University Grand Challenge project, “Microbes, Sewage, Health and Disease: Mapping the New York City Metagenome” (J.M.C.). We thank the Sloan Foundation for their

support to pioneer studies, including ours, on the microbiology of the built environment. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

- Creer S, Deiner K, Frey S, Porazinska D, Taberlet P, Thomas WK, Potter C, Bik HM. 2016. The ecologist's field guide to sequence-based identification of biodiversity. *Methods Ecol Evol* **7**:1008–1018. <http://dx.doi.org/10.1111/2041-210X.12574>.
- Fuhrman JA, Cram JA, Needham DM. 2015. Marine microbial community dynamics and their ecological interpretation. *Nat Rev Microbiol* **13**:133–146. <http://dx.doi.org/10.1038/nrmicro3417>.
- Zinger L, Gobet A, Pommier T. 2012. Two decades of describing the unseen majority of aquatic microbial diversity. *Mol Ecol* **21**:1878–1896. <http://dx.doi.org/10.1111/j.1365-294X.2011.05362.x>.
- Barberán A, Bates ST, Casamayor EO, Fierer N. 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J* **6**:343–351. <http://dx.doi.org/10.1038/ismej.2011.119>.
- Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH. 2013. Going back to the roots: the microbial ecology of the rhizosphere. *Nat Rev Microbiol* **11**:789–799. <http://dx.doi.org/10.1038/nrmicro3109>.
- Boetius A, Anesio AM, Deming JW, Mikucki JA, Rapp JZ. 2015. Microbial ecology of the cryosphere: sea ice and glacial habitats. *Nat Rev Microbiol* **13**:677–690. <http://dx.doi.org/10.1038/nrmicro3522>.
- Konstantinidis KT. 2014. Do airborne microbes matter for atmospheric chemistry and cloud formation? *Environ Microbiol* **16**:1482–1484. <http://dx.doi.org/10.1111/1462-2920.12396>.
- Klepeis NE, Nelson WC, Ott WR, Robinson JP, Tsang AM, Switzer P, Behar JV, Hern SC, Engelmann WH. 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *J Expo Anal Environ Epidemiol* **11**:231–252. <http://dx.doi.org/10.1038/sj.jea.7500165>.
- Barberán A, Dunn RR, Reich BJ, Pacifici K, Laber EB, Menninger HL, Morton JM, Henley JB, Leff JW, Miller SL, Fierer N. 2015. The ecology of microscopic life in household dust. *Proc R Soc B Biol Sci* **282**:20151139. <http://dx.doi.org/10.1098/rspb.2015.1139>.
- Dunn RR, Fierer N, Henley JB, Leff JW, Menninger HL. 2013. Home life: factors structuring the bacterial diversity found within and between homes. *PLoS One* **8**:e64133. <http://dx.doi.org/10.1371/journal.pone.0064133>.
- Lax S, Gilbert JA. 2015. Hospital-associated microbiota and implications for nosocomial infections. *Trends Mol Med* **21**:427–432. <http://dx.doi.org/10.1016/j.molmed.2015.03.005>.
- Bokulich NA, Mills DA, Underwood MA. 2013. Surface microbes in the neonatal intensive care unit: changes with routine cleaning and over time. *J Clin Microbiol* **51**:2617–2624. <http://dx.doi.org/10.1128/JCM.00898-13>.
- Meadow JF, Altrichter AE, Kembel SW, Moriyama M, O'Connor TK, Womack AM, Brown GZ, Green JL, Bohannon BJ. 2014. Bacterial communities on classroom surfaces vary with human contact. *Microbiome* **2**:7. <http://dx.doi.org/10.1186/2049-2618-2-7>.
- 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* **1**:e00022-16. <http://dx.doi.org/10.1128/mSystems.00022-16>.
- Hewitt KM, Gerba CP, Maxwell SL, Kelley ST. 2012. Office space bacterial abundance and diversity in three metropolitan areas. *PLoS One* **7**:e37849. <http://dx.doi.org/10.1371/journal.pone.0037849>.
- Flores GE, Bates ST, Knights D, Lauber CL, Stombaugh J, Knight R, Fierer N. 2011. Microbial biogeography of public restroom surfaces. *PLoS One* **6**:e28132. <http://dx.doi.org/10.1371/journal.pone.0028132>.
- Buse HY, Lu J, Struewing IT, Ashbolt NJ. 2013. Eukaryotic diversity in premise drinking water using 18S rDNA sequencing: implications for health risks. *Environ Sci Pollut Res Int* **20**:6351–6366. <http://dx.doi.org/10.1007/s11356-013-1646-5>.
- Baron JL, Vikram A, Duda S, Stout JE, Bibby K. 2014. Shift in the microbial ecology of a hospital hot water system following the introduction of an on-site monochloramine disinfection system. *PLoS One* **9**:e102679. <http://dx.doi.org/10.1371/journal.pone.0102679>.
- Stanish LF, Hull NM, Robertson CE, Harris JK, Stevens MJ, Spear JR, Pace NR. 2016. Factors influencing bacterial diversity and community composition in municipal drinking waters in the Ohio River Basin, USA. *PLoS One* **11**:e0157966. <http://dx.doi.org/10.1371/journal.pone.0157966>.
- Dybwad M, Skogan G, Blatny JM. 2014. Temporal variability of the bioaerosol background at a subway station: concentration level, size distribution, and diversity of airborne bacteria. *Appl Environ Microbiol* **80**:257–270. <http://dx.doi.org/10.1128/AEM.02849-13>.
- Cao C, Jiang W, Wang B, Fang J, Lang J, Tian G, Jiang J, Zhu TF. 2014. Inhalable microorganisms in Beijing's PM2.5 and PM10 pollutants during a severe smog event. *Environ Sci Technol* **48**:1499–1507. <http://dx.doi.org/10.1021/es4048472>.
- Leung MH, Wilkins D, Li EK, Kong FK, Lee PK. 2014. Indoor-air microbiome in an urban subway network: diversity and dynamics. *Appl Environ Microbiol* **80**:6760–6770. <http://dx.doi.org/10.1128/AEM.02244-14>.
- Robertson CE, Baumgartner LK, Harris JK, Peterson KL, Stevens MJ, Frank DN, Pace NR. 2013. Culture-independent analysis of aerosol microbiology in a metropolitan subway system. *Appl Environ Microbiol* **79**:3485–3493. <http://dx.doi.org/10.1128/AEM.00331-13>.
- Yooseph S, Andrews-Pfannkoch C, Tenney A, McQuaid J, Williamson S, Thiagarajan M, Bami D, Zeigler-Allen L, Hoffman J, Goll JB, Fadrosch D, Glass J, Adams MD, Friedman R, Venter JC. 2013. A metagenomic framework for the study of airborne microbial communities. *PLoS One* **8**:e81862. <http://dx.doi.org/10.1371/journal.pone.0081862>.
- Firth C, Bhat M, Firth MA, Williams SH, Frye MJ, Simmonds P, Conte JM, Ng J, Garcia J, Bhuvu NP, Lee B, Che X, Quan PL, Lipkin WI. 2014. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal *Rattus norvegicus* in New York City. *mBio* **5**:e01933-14. <http://dx.doi.org/10.1128/mBio.01933-14>.
- Reese AT, Savage A, Youngsteadt E, McGuire KL, Kolling A, Watkins O, Frank SD, Dunn RR. 22 September 2015. Urban stress is associated with variation in microbial species composition—but not richness—in Manhattan. *ISME J* <http://dx.doi.org/10.1038/ismej.2015.152>.
- McGuire KL, Payne SG, Palmer MI, Gillikin CM, Keefe D, Kim SJ, Gedalovich SM, Discenza J, Rangamannar R, Koshner JA, Massmann AL, Orazi G, Essene A, Leff JW, Fierer N. 2013. Digging the New York City skyline: soil fungal communities in green roofs and city parks. *PLoS One* **8**:e58020. <http://dx.doi.org/10.1371/journal.pone.0058020>.
- Hsu T, Joice R, Vallarino J, Abu-Ali G, Hartmann EM, Shafquat A, DuLong C, Baranowski C, Gevers D, Green JL. 2016. Urban transit system microbial communities differ by surface type and interaction with humans and the environment. *mSystems* **1**:e00018-16. <http://dx.doi.org/10.1128/mSystems.00018-16>.
- Afshinnekoo 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. 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>.
- Martin LJ, Adams RI, Bateman A, Bik HM, Hawks J, Hird SM, Hughes D, Kembel SW, Kinney K, Kolokotronis S-O, Levy G, McClain CR, Meadow JF, Medina RF, Mhuireach G, Moreau CS, Munshi-South J, Nichols LM, Palmer C, Popova L, Schal C, Täubel M, Trautwein M, Ugalde JA. 2015. Evolution of the indoor biome. *Trends Ecol Evol* **30**:223–232. <http://dx.doi.org/10.1016/j.tree.2015.02.001>.
- Nworie O, Mercy M, Chukwudi A, Oko I, Chukwudum SO, Agah VM, Ekuma UO. 2012. Antibigram of bacteria isolated from automated teller machines within Abakaliki metropolis. *Am J Infect Dis* **8**:168–174. <http://dx.doi.org/10.3844/ajidsp.2012.168.174>.
- Reynolds KA, Watt PM, Boone SA, Gerba CP. 2005. Occurrence of bacteria and biochemical markers on public surfaces. *Int J Environ Health Res* **15**:225–234. <http://dx.doi.org/10.1080/09603120500115298>.
- Roberts MC, Soge OO, No D, Helgeson SE, Meschke JS. 2011. Characterization of methicillin-resistant *Staphylococcus aureus* isolated from public surfaces on a university campus, student homes and local com-

- munity. *J Appl Microbiol* **110**:1531–1537. <http://dx.doi.org/10.1111/j.1365-2672.2011.05017.x>.
34. **Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW.** 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* **12**:87. <http://dx.doi.org/10.1186/s12915-014-0087-z>.
 35. **Bates ST, Clemente JC, Flores GE, Walters WA, Parfrey LW, Knight R, Fierer N.** 2013. Global biogeography of highly diverse protistan communities in soil. *ISME J* **7**:652–659. <http://dx.doi.org/10.1038/ismej.2012.147>.
 36. **Perntaler J.** 2005. Predation on prokaryotes in the water column and its ecological implications. *Nat Rev Microbiol* **3**:537–546. <http://dx.doi.org/10.1038/nrmicro1180>.
 37. **Parfrey LW, Walters WA, Knight R.** 2011. Microbial eukaryotes in the human microbiome: ecology, evolution, and future directions. *Front Microbiol* **2**:153. <http://dx.doi.org/10.3389/fmicb.2011.00153>.
 38. **Adl M, Gupta VS.** 2006. Protists in soil ecology and forest nutrient cycling. *Can J Forest Res* **36**:1805. <http://dx.doi.org/10.1139/x06-056>.
 39. **Farmer MA.** 1993. Ultrastructure of *Ditrichomonas honigbergii* n. sp. (Parabasalia) and its relationship to amitochondrial protists. *J Eukaryot Microbiol* **40**:619–626. <http://dx.doi.org/10.1111/j.1550-7408.1993.tb06119.x>.
 40. **Edgcomb V, Viscogliosi E, Simpson AG, Delgado-Viscogliosi P, Roger AJ, Sogin ML.** 1998. New insights into the phylogeny of trichomonads inferred from small subunit rRNA sequences. *Protist* **149**:359–366. [http://dx.doi.org/10.1016/S1434-4610\(98\)70042-2](http://dx.doi.org/10.1016/S1434-4610(98)70042-2).
 41. **Mostegl MM, Wetscher A, Richter B, Nedorost N, Dinhopf N, Weisenböck H.** 2012. Detection of *Tritrichomonas foetus* and *Pentatrichomonas hominis* in intestinal tissue specimens of cats by chromogenic in situ hybridization. *Vet Parasitol* **183**:209–214. <http://dx.doi.org/10.1016/j.vetpar.2011.07.050>.
 42. **Maritz JM, Land KM, Carlton JM, Hirt RP.** 2014. What is the importance of zoonotic trichomonads for human health? *Trends Parasitol* **30**:333–341. <http://dx.doi.org/10.1016/j.pt.2014.05.005>.
 43. **Su C, Khan A, Zhou P, Majumdar D, Ajzenberg D, Dardé M-L, Zhu X-Q, Ajika JW, Rosenthal BM, Dubey JP, Sibley LD.** 2012. Globally diverse *Toxoplasma gondii* isolates comprise six major clades originating from a small number of distinct ancestral lineages. *Proc Natl Acad Sci U S A* **109**:5844–5849. <http://dx.doi.org/10.1073/pnas.1203190109>.
 44. **Fierer N, Hamady M, Lauber CL, Knight R.** 2008. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci U S A* **105**:17994–17999. <http://dx.doi.org/10.1073/pnas.0807920105>.
 45. **Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R.** 2009. Bacterial community variation in human body habitats across space and time. *Science* **326**:1694–1697. <http://dx.doi.org/10.1126/science.1177486>.
 46. **Macher JM, Huang F-Y, Flores M.** 1991. A two-year study of microbiological indoor air quality in a new apartment. *Arch Environ Health* **46**:25–29. <http://dx.doi.org/10.1080/00039896.1991.9937425>.
 47. **Fischer G, Dott W.** 2003. Relevance of airborne fungi and their secondary metabolites for environmental, occupational and indoor hygiene. *Arch Microbiol* **179**:75–82. <http://dx.doi.org/10.1007/s00203-002-0495-2>.
 48. **Pel JJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, Andersen MR, Bendtsen JD, Benen JA, van den Berg M, Breestraat S, Caddick MX, Contreras R, Cornell M, Coutinho PM, Danchin EG.** 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat Biotechnol* **25**:221–231. <http://dx.doi.org/10.1038/nbt1282>.
 49. **Hayden NJ, Maude RB, Proctor FJ.** 1994. Studies on the biology of black mould (*Aspergillus niger*) on temperate and tropical onions. 1. A comparison of sources of the disease in temperate and tropical field crops. *Plant Pathol* **43**:562–569. <http://dx.doi.org/10.1111/j.1365-3059.1994.tb01591.x>.
 50. **Sampaio JP, Bauer R, Begerow D, Oberwinkler F.** 1999. *Occultifur externus* sp. nov., a new species of simple-pored auricularioid heterobasidiomycete from plant litter in Portugal. *Mycologia* **91**:1094–1101. <http://dx.doi.org/10.2307/3761639>.
 51. **Leong SL, Pettersson OV, Rice T, Hocking AD, Schnürer J.** 2011. The extreme xerophilic mould *Xeromyces bisporus*—growth and competition at various water activities. *Int J Food Microbiol* **145**:57–63. <http://dx.doi.org/10.1016/j.ijfoodmicro.2010.11.025>.
 52. **Lax S, Smith DP, Hampton-Marcell J, Owens SM, Handley KM, Scott NM, Gibbons SM, Larsen P, Shogan BD, Weiss S, Metcalf JL, Ursell LK, Vázquez-Baeza Y, Van Treuren W, Hasan NA, Gibson MK, Colwell R, Dantas G, Knight R, Gilbert JA.** 2014. Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science* **345**:1048–1052. <http://dx.doi.org/10.1126/science.1254529>.
 53. **Flores GE, Bates ST, Caporaso JG, Lauber CL, Leff JW, Knight R, Fierer N.** 2013. Diversity, distribution and sources of bacteria in residential kitchens. *Environ Microbiol* **15**:588–596. <http://dx.doi.org/10.1111/1462-2920.12036>.
 54. **Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R.** 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6**:1621–1624. <http://dx.doi.org/10.1038/ismej.2012.8>.
 55. **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>.
 56. **Rideout JR, He Y, Navas-Molina JA, Walters WA, Ursell LK, Gibbons SM, Chase J, McDonald D, Gonzalez A, Robbins-Pianka A, Clemente JC, Gilbert JA, Huse SM, Zhou HW, Knight R, Caporaso JG.** 2014. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* **2**:e545. <http://dx.doi.org/10.7717/peerj.545>.
 57. **Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C.** 2014. Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS One* **9**:e87624. <http://dx.doi.org/10.1371/journal.pone.0087624>.
 58. **Caron DA, Countway PD, Savai P, Gast RJ, Schnetzer A, Moorthi SD, Dennett MR, Moran DM, Jones AC.** 2009. Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Appl Environ Microbiol* **75**:5797–5808. <http://dx.doi.org/10.1128/AEM.00298-09>.
 59. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL.** 2006. Greengenes, a chimeric-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>.
 60. **Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO.** 2007. Silva: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**:7188–7196. <http://dx.doi.org/10.1093/nar/gkm864>.
 61. **Parfrey LW, Walters WA, Lauber CL, Clemente JC, Berg-Lyons D, Teiling K, Kodira C, Mohiuddin M, Brunelle J, Driscoll M, Fierer N, Gilbert JA, Knight R.** 2014. Communities of microbial eukaryotes in the mammalian gut within the context of environmental eukaryotic diversity. *Front Microbiol* **5**:298. <http://dx.doi.org/10.3389/fmicb.2014.00298>.
 62. **Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R.** 2010. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**:266–267. <http://dx.doi.org/10.1093/bioinformatics/btp636>.
 63. **Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E, Methé B, DeSantis TZ, Human Microbiome Consortium, 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. <http://dx.doi.org/10.1101/gr.112730.110>.
 64. **Korajkic A, Parfrey LW, McMinn BR, Baeza YV, VanTeuren W, Knight R, Shanks OC.** 2015. Changes in bacterial and eukaryotic communities during sewage decomposition in Mississippi river water. *Water Res* **69**:30–39. <http://dx.doi.org/10.1016/j.watres.2014.11.003>.
 65. **Price MN, Dehal PS, Arkin AP.** 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* **26**:1641–1650. <http://dx.doi.org/10.1093/molbev/msp077>.
 66. **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.0061217>.
 67. **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>.
68. **Adams RI, Bateman AC, Bik HM, Meadow JF.** 2015. Microbiota of the indoor environment: a meta-analysis. *Microbiome* **3**:49. <http://dx.doi.org/10.1186/s40168-015-0108-3>.
69. **Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S, Harris HM, Coakley M, Lakshminarayanan B, O'Sullivan O, Fitzgerald GF, Deane J, O'Connor M, Harnedy N, O'Connor K, O'Mahony D, van Sinderen D, Wallace M, Brennan L, Stanton C.** 2012. Gut microbiota composition correlates with diet and health in the elderly. *Nature* **488**:178–184. <http://dx.doi.org/10.1038/nature11319>.
70. **De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, Collini S, Pieraccini G, Lionetti P.** 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* **107**:14691–14696. <http://dx.doi.org/10.1073/pnas.1005963107>.
71. **Adams RI, Miletto M, Lindow SE, Taylor JW, Bruns TD.** 2014. Airborne bacterial communities in residences: similarities and differences with fungi. *PLoS One* **9**:e91283. <http://dx.doi.org/10.1371/journal.pone.0091283>.
72. **Lauber CL, Hamady M, Knight R, Fierer N.** 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl Environ Microbiol* **75**:5111–5120. <http://dx.doi.org/10.1128/AEM.00335-09>.
73. **Bowers RM, Lauber CL, Wiedinmyer C, Hamady M, Hallar AG, Fall R, Knight R, Fierer N.** 2009. Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei. *Appl Environ Microbiol* **75**:5121–5130. <http://dx.doi.org/10.1128/AEM.00447-09>.
74. **Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C.** 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* **12**:R60. <http://dx.doi.org/10.1186/gb-2011-12-6-r60>.