Spring 2016

Microbial Resuspension in the Built Environment: The Role of Flooring Materials and Anthropogenic Traffic on Taxonomic Diversity

Gates Failing
Elizabethtown College, failingg@etown.edu

Microbial Resuspension in the Built Environment:

The Role of Flooring Materials and Anthropogenic Traffic on Taxonomic Diversity

Gates Failing

Elizabethtown College

Author Note

This thesis was submitted in partial fulfillment of the requirement for

Honors in the Discipline in Biology and the Elizabethtown College Honors Program.

May 6, 2016

Committee:

Primary Advisor - Dr. Debra Wohl

Secondary Advisors - Dr. David Bowne

Dr. Diane Bridge

Abstract

The interplay between humans and built environments is a new frontier for microbial ecology. Approximately 90% of human activities occur indoors, and current estimates for microorganism diversity in buildings are in the trillions. Previous studies have investigated microbial resuspension (i.e., the surface to air release of biotic and abiotic particulate matter) via temporal analysis of human occupancy patterns and spatial analysis of different flooring materials. However, prior research has not sufficiently addressed flooring structure, human-mediated resuspension in unconstrained environments, and phylogenetic analysis within the context of a single study. Our investigation examined the effect of surface composition and human traffic intensity on the taxonomic composition of airborne microbial communities. In a college academic building, 24 air samples were collected over carpeting during high (n=6) and low (n=6) human activity periods and over linoleum during high (n=6) and low (n=6) activity. DNA was extracted, amplified using prokaryotic 16S primers, sequenced using the Illumina MiSeq platform, and analyzed with the Quantitative Insights into Microbial Ecology statistical package. Prior to merging reads and quality filtering, 227 sequences were yielded across 14 samples and one control. Alpha (within sample) diversity indices of genera richness and evenness were reported along with beta diversity (between sample) comparisons of sequence counts and shared genera across sampling conditions. While low sequence yields precluded the determination of the explanatory power for the activity and flooring variables, the present study's limitations and new directions for investigating the composition of indoor microbial communities were discussed. With methodological revisions, we anticipate that future studies will help to elucidate the role of building design in modulating microbial resuspension dynamics induced by human traffic patterns between indoor and outdoor environments.

Introduction

Context

Bacteria and fungi are ubiquitous residents across nearly every environment on Earth. The discipline of microbial ecology seeks to apply ecological principles of diversity and species interactions to microorganisms (Konopka, 2009). One means of classifying microbial communities is by grouping them into microbiomes, or the sum total of microorganisms interacting within a community defined by a given parameter (Kembel et al., 2012). Microbiome research initially gained traction through the study of human microbial communities, such as those in mammalian gastrointestinal tracts, to determine the functionality of microbes in terms of disease susceptibility (Ursell et al., 2012). The rise of high-throughput characterization methods that are independent of culturing biases, such as 454 pyrosequencing and Illumina, has tremendously accelerated research and shifted the discipline towards a more genetic focus (Kembel et al., 2012; Ursell et al., 2012).

Genomic analysis has implemented the gene for the small ribosomal subunit in prokaryotes, also known as the 16S rRNA gene, as a metric for determining phylogenetic relatedness among microorganisms. Specifically, the V4 segment of the gene has conserved and variable regions that are apt for genetic comparisons (Yarza et al., 2014). Polymerase chain reaction (PCR) amplification of sample DNA followed by sequencing using “next-generation” technologies is far more cost effective and efficient than the previous standard, the Sanger chain-termination method. Such advances have had positive implications not just for microbial ecology, but for evolutionary biology, oncology research, and even the general public through the promises of personalized medicine through individual genome sequencing (Schuster, 2008).

Despite the focus on the human microbiome, other investigators have applied microbial ecology and genomic techniques to microbial communities of inanimate domains. One such area is the built environment microbiome, which focuses on microbial communities in indoor locations (Kembel et al., 2012). Numerous studies have identified relationships between microbial community distribution and the structure of the indoor environment as well as suggested implications for human health. Phylogenetic analysis of resuspended bioaerosols could provide insight into respiratory diseases, such as asthma (Hospodsky et al., 2012) and foster evidence-based building design decisions (Kembel et al., 2014). Pathogenic organisms, such as MRSA (methicillin-resistant *Staphylococcus aureus*), can spread via surface to air exchange (Rintala et al., 2008). A recent risk analysis assessed the probablitiy of pathogen resuspension from either air or human-mediated disturbance events (You & Wan, 2015). Furthermore, nonspecific externalities, such as sick building syndrome (SBS), can result from improper building ventilation (Chen & Hildemann, 2009; Kembel et al., 2012). Developmentally, bioaerosols may be able to impact colonization of the fetal intestine, which would have lifelong health ramifications (Rintala et al., 2008). Thus, resuspension is a broad public health concern given that 90% of human activities are indoors (Kelley & Gilbert, 2013).

Depending on the location and type of environmental sampling, some investigators concluded after analysis of indoor dust that the indoor microbiome is significantly comprised of human-associated taxa (Rintala et al., 2008), while other evidence suggests that outdoor taxa are the primary sources of diversity for indoor air (Meadow et al., 2014). In either case, these communities are subject to both spatial and temporal factors. For example, building use and seasonal change (Rintala et al., 2008) as well as dispersal mechanisms (Kembel et al., 2012) can impact species distribution. The term microbial biogeography, referring to spatial and temporal

variations in biodiversity, is often applied to microbiome studies. This perspective seeks to elucidate the relative contributions of selective pressures from the environment and stochastic historical events to community composition **(**Hughes Martiny et al., 2006; Kelley & Gilbert, 2013).

The physical structure of the built environment is intrinsically linked to the microbial biogeography of indoor air (Hughes Martiny et al., 2006; Kembel et al., 2014). Function (utilization), form (overall shape), and organization have been identified as key architectural factors that govern species diversity, linking construction and ecological processes. One area of interest is the circulation of microorganisms across various building domains. In particular, researchers have observed compositional homology among microbial communities of dust, flooring surfaces, and the air. *Proteobacteria* and *Firmicutes* are two phyla that are thought to be shared across these three locations, indicating that a particulate dispersal mechanism, such as resuspension, may be influential in the establishment of a core community as well as affecting change over time (Kembel et al., 2014).

Particulate Matter Resuspension

Indoor microbial dispersal is governed by both internal release, such as from building materials or occupants, and external inputs. The latter could include transfer via HVAC systems as well as human-mediated passive conveyance from an outdoor source. While outdoor air is a major compositional contributor, other abiotic sources of diversity, such as moisture availability and temperature, remain significant (Kembel et al., 2014). Conceptual research has generated a hypothetical model for inflows and outflows of particulate matter (Figure 1). The schematic of the two-compartment model addresses both the source and the destination of resuspended

particulate matter. Of primary interest are the inputs and outputs between the surface particulate matter and the rest of the environment. The major inflows are track-in, spillage (i.e., direct surface-to-surface conveyance), and deposition (i.e., air-mediated transfer). Resuspension is the only significant outflow from the surface to the atmosphere, besides purposeful cleaning activities (Schneider, 2008).

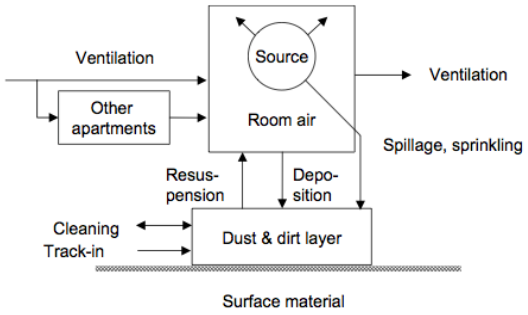

Figure 1. Schematic of inflows and outflows of particulate matter in a system.
Note. From "Dust and fibers as a cause of indoor environment problems" by T. Schneider, 2008, *Scandinavian Journal of Work, Environment & Health Supplement*, 4, p. 11.

General particulate resuspension is defined as "…a phenomenon in which particles, initially on a surface, join the passing fluid stream" (Mukai et al., 2009, p. 1022), and it is considered to be a relatively strong outflow source effector (Qian et al., 2012). Most simply, it can be mathematically modeled by r=R/L, whereas the L variable is particulate concentration on the surface layer, R is the resuspension flow, and r is the rate of release in a given time interval (Qian et al., 2008). Resuspension is governed by numerous factors, including air speed, size, air turbulence, vibration forces, air moisture, and interactions with the solid surface. For example, increasing air velocity can generally lead to increased rates of resuspension. Laboratory simulations have been attempted using artificial "seed" particles to determine the force of stimulation needed to reach a resuspension threshold (Mukai et al., 2009). Room airflow patterns can also dictate compositional dynamics, especially with regard to HVAC systems. One study

found that levels of resuspension were an estimated 60% to 80% lower when sampling was completed approximately 120 cm away from the activity source versus immediate source sampling. The quantity of particulate collection was inversely proportional to distance from the source of disturbance, suggesting that more global airflow patterns were affecting collection as a result of HVAC operation (Rosati et al., 2008).

Research on particulate resuspension has primarily sought the mechanistic explanations for anthropogenic release of abiotic particulate matter. The human foot is able to produce sufficient lift forces that outcompete the attractive forces acting on a particle; when this imbalance occurs, the particle leaves the depositional surface. Shearing forces and interactions with other particles lead to greater dispersal (Qian et al., 2008) along with increases in particle size, though the threshold varies in the literature. Kildeso et al. (1999) observed resuspension at diameters greater than 1 μm due to the reduction of Van der Waals inter-particle forces that prevent dispersal for smaller sizes. In contrast, other human traffic studies found resuspension was most elevated for particles between 10-25 μm and decreased as particle size decreased (Thatcher & Layton, 1995). Given that bacteria and other microorganisms suspended in the atmosphere occur anywhere from 0.2 μm to 30 μm, the resuspension mechanism would likely include a substantial biotic component (Chen & Hildemann, 2009).

Resuspension Variables

Indoors, resuspension proceeds during general human activities, such as walking (Qian et al., 2008), dusting, and vacuuming (Schneider, 2008). For example, household walking led to a two-fold increase in resuspended particulate matter over basal levels (Thatcher & Layton, 1995). The amount of particulate matter released is less for moderate household movements over

intensive cleaning activities (Schneider, 2008), and human participant density is associated with higher rates of release for 5 μm particles (Ferro et al., 2004). Furthermore, the type of activity has been correlated with the average size of particles resuspended. Walking and cleaning have been shown to result in the release of particulate matter along the 2.5 to 10 μm size range. Excluding aerosol release from combustion activities, resuspension is regarded as a major mechanism for particulate matter efflux indoors (Hospodsky et al., 2012).

However, resuspension studies are limited by the effect of a confounding variable, anthropogenic deposition. Both general resuspension from environmental surfaces and human-mediated microbial dissemination combine to increase bioaerosol concentrations. Resuspension is inherently linked to human microbial shedding in a cycle of deposition and rerelease from a surface (Mukai et al., 2009). A personal cloud is hypothesized to facilitate dispersal on a per person basis, which would have ramifications for the heterogeneity of building airflow dynamics (Schneider, 2008). In occupied residential environments, one study by Qian and colleagues (2008) identified both resuspension and deposition as influential dispersal modes. Researchers found an average particulate matter increase of 2.5 times the basal level for 10 μm particles due to resuspension and deposition but could not distinguish among the competing variables. Finally, deposition is hypothesized to alter inter-particle interactions, including promoting adhesion, which could potentially impact the propensity for future resuspension.

In terms of temporal variation, Qian et al. (2008) also analyzed human resuspension of general particulate matter over a thirty-minute interval and found that the rate of release decreased as the sampling interval progressed. The change over time was attributed to particle-surface dynamics; species that have lower attractive interactions are displaced most easily during the initial sampling period. The rate of resuspension may also depend on the duration of the most

recent accumulation, with longer elapsed times from the original deposition event leading to resuspension resistance due to particle migration into the interior of the carpeting material.

Physically, resuspension can be influenced through surface variations. For example, different flooring materials impact the rate and content of particulate resuspension. One study found that resuspension was lowest with sheet metal, intermediate with linoleum flooring tiles, and highest with carpeting. The fibrous carpet was hypothesized to facilitate resuspension by promoting particle exposure to higher velocity airflow (Mukai et al., 2009). Carpets are considered to be dust sinks (Thatcher & Layton, 1995), and increasing the density of the upper carpet layer can retard resuspension by adhering particulate matter (Schneider, 2008). In general, older carpets have a lower rate of resuspension, reduced approximately 10 to 100 fold, when compared to that of newer carpets (Rosati et al., 2008).

One carpet study attempted to separate resuspension from deposition by covering carpeting with plastic sheeting to minimize resuspension and only permit deposition. During human traffic conditions, particulate release was 2.7 times greater than the background level without activity. In contrast, occupancy with a pure carpet that permitted resuspension and deposition resulted in a statistically significant level of release that was 4.7 times greater than the basal levels. While shedding can account for some variation in particulate matter release, resuspension remains an important mechanism for dispersal (Hospodsky et al., 2012).

Phylogenetic Resuspension Studies

Resuspension is applicable to microbial biogeography when there is a biologically active component to resuspended particulate matter, namely prokaryotes and fungi. In the previously discussed Hospodsky et al. (2012) study with plastic sheeting and carpeting, human traffic was

classified as both a source of general particulate matter mass (15-fold concentration increase of 10 μm particles vs. background) as well as bacterial genetic material (66-fold concentration increase of 10 μm particles vs. background). Using pyrosequencing, the researchers found evidence of compositional homology between air and surface microbial communities. Indoor resuspension and shedding appeared to facilitate dispersal of microbial taxa associated with humans (17% of total sequence abundance), such as *Propionibacterineae*, *Staphylococcus*, and *Streptococcus*, as well as external outdoor taxa, namely *Sphingomonadaceae*, *Rhodobacteraceae*, and *Streptophyla*. The human-associated taxa were hypothesized to originate from dermal squamous cells. Researchers have estimated that the percentage of human-associated taxa is highly variable, shifting from ~15%-75% for indoor dust analysis depending on the building type and usage patterns.

Other studies have used diverse methods to investigate the biotic composition of indoor air (Chen & Hildemann, 2009; Qian et al., 2012)**.** In an occupied classroom, Qian et al. (2012) found elevated atmospheric levels of bacterial DNA in the 3 to 5 μm fractions and fungal DNA in the 2 to 5 μm and greater than 9 μm range within air samples collected during activity periods (Qian et al., 2012). Interestingly, the largest general particulate matter increases during occupancy were for those greater than 9 μm, which stands in opposition to the size ranges established for bacterial hereditary material detailed above. The researchers hypothesized that the lack of size range overlap between general particulate matter and genomic fractions was the result of aggregate formation among bacteria/fungal species. Particle size may not be representative of the biotic or abiotic content of resuspended material, showing the deficits of non-phylogenetic methods. In the study, pyrosequencing ascertained that 18% of the collected

DNA was from microorganisms typically associated with humans, particularly taxa connected with the skin, nose, and hair.

Chen and Hildemann (2009) classified the components of particulate matter into protein (for general quantification), beta-glucan polymers (for fungi), and endotoxins (for gram-negative bacteria) as biomarker indicators for microbial diversity in ten residential homes**.** The researchers found some evidence that domesticated animals and extensive carpeting led to elevated particle release, especially >2.5 μm in size, which correlated to the beta-glucan size range. The study also suggested that levels of particulate matter could not be correlated to the cleaning history of the environment or the time elapsed since the flooring was placed, demonstrating that past building activities may be weaker effectors for resuspension.

A more recent study analyzed airborne microbial diversity in a 30x30 m controlled chamber where between one to eight participants either walked or sat during a two hour period. Walking conditions were conducted over carpeting or plastic sheeting. Samples were collected using 0.2 μm nitrocellulose membrane filters under a flow rate of 25 liters per minute (LPM). The researchers found evidence that higher occupancy (2 people walking or 8 people sitting) resulted in a 100% increase in OTU richness. OTUs, or operational taxonomic units, are a means of genetic classification based upon percent sequence homology at the 16S rRNA gene. Also, larger particles beyond 5 μm were more likely to be released in the carpeting condition versus the plastic sheeting treatment. Interestingly, based on Illuminia sequencing analysis, the researchers postulated that humans may be more likely to resuspend nonhuman-associated taxa, thus serving as a surrogate of sorts, rather than shed human-associated taxa. Overall, the researchers hypothesized that human activity does not substantially alter the taxonomic distribution of indoor air from that of outdoor air (Adams et al., 2015).

The results of Adams and colleagues (2015) highlight a more general trend in the literature. Phylogenetic analysis of an academic building found evidence that overall, outdoor air sources were the strongest predictor of microbial diversity (Meadow et al., 2014), and these results are consistent with other literature sources previous described (Adams et al., 2015; Hospodsky et al., 2012; Qian et al., 2012). In the study by Meadow and colleagues, the most prevalent taxa indoors were 88% homologous to taxa of the external environment. Therefore, ventilation is a primary means by which temporal variations in outdoor air, such as during seasonal fluctuations, can induce similar compositional changes within indoor air. Nevertheless, room occupants still significantly influenced bioaerosol diversity, and human-associated taxa were recorded, including the most prevalent genera *Corynebacterium*, *Staphylococcus*, and *Acinetobacter*. In contrast, Tringe et al. (2008) argued in a metagenomic study of two commercial shopping indoor environments for the existence of a core indoor microbiome distinct from outdoor air. In juxtaposing phylotypes across indoor environments, the authors found evidence that there were more composition similarities between the two indoor sampling locations in comparison to outdoor samples.

Current Study

Given the novelty of next-generation sequencing, phylogenetic studies of the built environment are not yet sufficiently comprehensive to explain how human activity intersects with the built environment to subsequently affect microbial diversity. Studies in the scientific literature have investigated resuspension of general particulate matter during human occupancy (Rintala et al., 2008) and with different flooring materials (Mukai et al., 2009). However, much of the current body of research has focused on particle size, rather than the identity of those

particles, such as classification of biotic vs. abiotic or fungal vs. bacterial components. The few studies that implemented phylogenetic analysis to characterize composition only sampled a single activity event (Hospodsky et al., 2012) and flooring type (Qian et al., 2012) or used a controlled chamber rather than a more ecologically valid environment (Adams et al., 2015). Therefore, prior research has not sufficiently integrated aspects of flooring design, *in situ* human-mediated resuspension, and genetic analysis to elucidate microbial community dynamics.

The 2011 Indoor Air Symposium on the microbiomes of built environments strongly recommended that future studies evaluate construction components, such as carpeting, and human traffic patterns (Corsi et al., 2012). Based on this directive, the present study sought to investigate some of the spatial and temporal variables influencing resuspension and determine their subsequent effect on microbial community composition. To this end, human-mediated microbial resuspension with different types of indoor flooring materials, carpeting and linoleum, was measured and analyzed within an indoor hallway. A weekday sampling period served as a high human activity treatment, and a weekend time was the low human activity treatment. The study aimed to answer the following research questions:

1. How is the phylogenetic profile of indoor air influenced by variations in flooring type and human activity?
2. What is the extent of within sample variation, or alpha diversity, for indoor air?
3. What is the extent of between sample variation, or beta diversity, for indoor air?

To answer these questions, DNA was extracted from air samples collected over two different flooring materials and during high or low levels of human occupancy. Using prokaryotic 16S primers for amplification, the Illumina MiSeq sequencing platform was implemented to provide insight into genetic relatedness to qualitatively and quantitatively

describe microorganisms of indoor air. The bioinformatics software Quantitative Insights Into Microbial Ecology (QIIME) was utilized for generating genera abundances to assess question 1. For question 2, genera richness and evenness were calculated for each air sample as an alpha diversity metric. Beta diversity, question 3, was assessed by comparing total sequence counts and identifying shared genera across sampling conditions.

It was hypothesized that the greatest quantity of sequences and highest biodiversity of microorganisms (richness and evenness) would be present in air samples collected during periods of high human activity over carpeting. Higher rates of activity are correlated to an increased rate of resuspension (Qian et al., 2008) and carpeting is known to facilitate particle release (Adams et al., 2015). In contrast, samples collected during low activity times over linoleum were hypothesized to be less taxonomically diverse in terms of number of sequences, richness, and evenness when compared to high activity samples. Lastly, it was predicted that a unique indoor microbiome would be present and contain taxonomic groups associated with both the human microbiome and the external environment. Overall, the present study sought to integrate the role of anthropogenic disturbance events and building structure on indoor microbial community diversity.

Materials and Methods

Study Location and Layout

All air samples were collected from a single spatial sampling location within Nicarry Hall, an academic building at Elizabethtown College in Elizabethtown, PA. High human traffic and the symmetrical nature of hallway structure within Nicarry Hall made it conducive for the present study. The sampled zone was a 2.5 m wide x 2.54 m tall hallway segment within the building, as shown in the first floor blueprint (Figure 2). The linoleum flooring treatment was

represented by the inherent flooring surface present in the building. The carpeting treatment was a used “StaticStopper” commercial mat measuring 5.9 m by 1.07 m x 0.07 m (Certified Carpet, Inc., Lancaster, PA). Analytical test filter funnels (Thermo Fischer Scientific, Inc., Waltham, MA) were placed 4.06 m apart and approx. 1 m above the floor on either end of the hallway zone via attachment to the wall (see schematic in Figure 3). The two funnels were each connected via flexible plastic tubing of 1.2 cm in diameter to a Leland Legacy air pump (SKC, Inc., Eighty Four, PA) as shown in Figure 4.

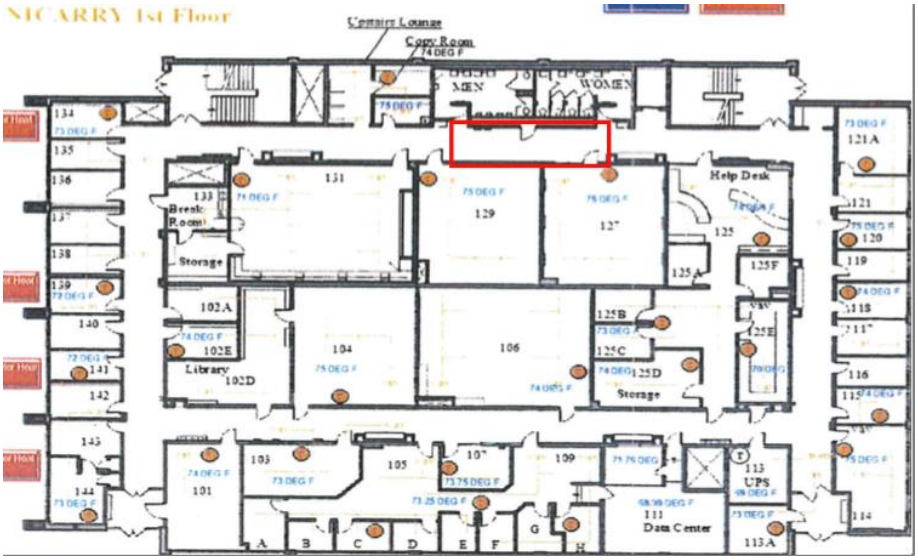

Figure 2. Blueprint for first floor of Nicarry Hall. Area outlined is the hallway sampling location.

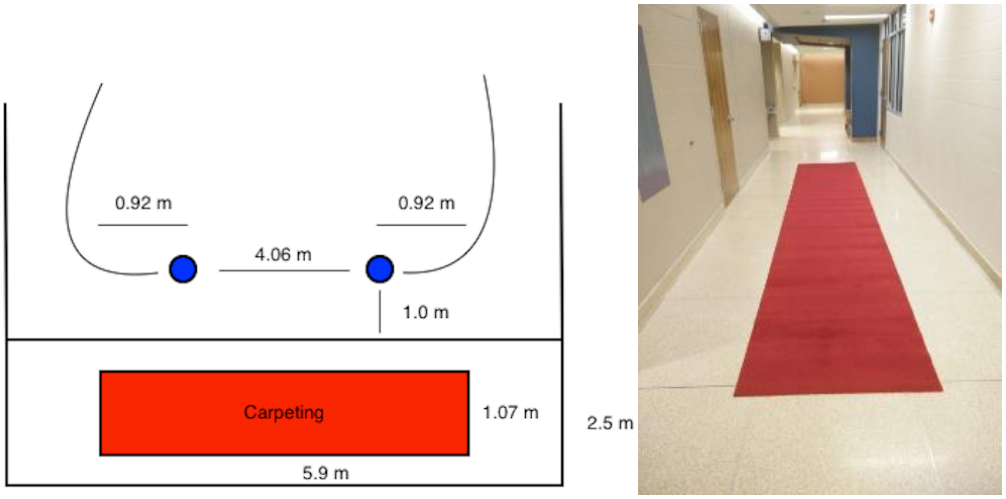

Figure 3. Hallway sampling schematic and image. Blue circles are the analytical test filter funnels connected via tubing to the air samplers placed above the ceiling tiles. 4.06 m is the distance between samplers, 1.0 m is the height above the floor, and 0.92 m is the distance from the sampler to the edge of the carpet. The carpet was removed for the linoleum sampling periods.

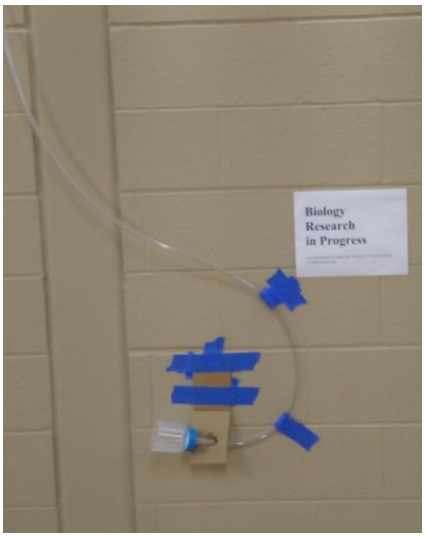

Figure 4. Air sampler wall attachment.

Sample Collection

A total of 24 samples were collected over a 4-week period from October 19, 2015 to November 16, 2015. Prior to sampling, the carpet was cleaned with detergent via an industrial washer, and the linoleum was cleaned by custodial staff. On October 30, 2015, the halfway point for sampling, the carpet and linoleum floor were vacuumed. All samples were collected under a

flow rate of 7 liters per minute (LPM) for 10 hours. Each sample consisted of two filter membranes that were subsequently pooled during PCR. At this flow rate and duration, approx. 4.2 m^3 of air passed through each funnel containing a separate nitrocellulose filter membrane with a 47 mm diameter and 0.45 μm pore size (Thermo Fischer Scientific, Inc., Waltham, MA). The high human activity samples were collected during weekdays (M-Th) between 8 AM and 6 PM. The low human activity samples were collected during weeknights (M-Th) between 9:30 PM and 7:30 AM. Human activity was approximated based upon average building occupancy data during a Fall 2015 week from a campus scheduling service (Table 1). There were no formally listed occupancy data for low activity sampling times.

Table 1. Estimated occupancy data for first floor academic classrooms in Nicarry Hall during a daytime sampling period

Day of the Week	**Estimated Number of Individuals**
Monday	208
Tuesday	368
Wednesday	203
Thursday	279

Out of the 24 samples, 6 were collected each during the high activity carpeting treatment, low activity carpeting treatment, high activity linoleum treatment, and low activity linoleum treatment. The distribution of samples across days of the week is shown in Table 2. Following a 10-hour runtime, both membranes were removed from their respective analytical test funnels and stored in 5 mL transport vials (Stockwell Scientific, Scottsdale, AZ) also containing approximately 1.8 g of 0.70 mm garnet beads in preparation for DNA extraction (MO BIO Laboratories, Inc., Carlsbad, CA). Membranes were folded inward and inserted via forceps so

that the exposure surface faced inward to maximize extraction potential. All membrane-containing vials were refrigerated at ~4°C.

Table 2. Air samples collected for all four treatments across the four possible weekdays/weeknights

Day of the Week	**HA Carpet (n=6)**	**LA Carpet (n=6)**	**HA Linoleum (n=6)**	**LA Linoleum (n=6)**
Monday	4	4	1	
Tuesday	2	2		
Wednesday			2	3
Thursday			3	3

HA= High Activity (Day); LA=Low Activity (Night)

DNA Extraction and Quantification

DNA was extracted in accordance with a modified version of the protocols used from both the UltraClean Microbial DNA Isolation Kit and the PowerWater DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad, CA). Reagents were utilized from the UltraClean Kit while volumes and centrifugation times were taken from the PowerWater Kit. First, 1 mL of MD1 was added to the 5 mL transport vial containing the filter membrane and the garnet beads. Tubes were attached horizontally to a Vortex Adaptor (MO BIO Laboratories, Inc., Carlsbad, CA) on a Vortex-Genie 2 (Scientific Industries, Inc., Bohemia, NY) to mechanically disrupt the membranes. Then, the vial was centrifuged via a Sorvall RC-5C Plus (Kendro Laboratory Products, Newtown, CT) at 4,000 x g for 1 minute at ~21°C. The supernatant was transferred to a 2 mL collection tube and centrifuged via an ALC micro Centrifugette 4212 (ALC International, Cologno Monzese, Italy) for 13,000 x g for 1 minute at room temperature. All subsequent centrifugations were performed using this instrument at the same temperature and rpm. The

supernatant was again transferred to a 2 mL collection tube and 200 μL of MD2 was added. The tube was vortexed to mix and incubated at ~4°C for 5 minutes.

The tube was then centrifuged for 1 minute. The supernatant was transferred to a new 2 mL collection tube and 650 μL of MD3 was added with subsequent vortexing for 10 seconds. Then, 650 μL of the total tube volume was delivered to a spin filter tube and centrifuged for 1 minute. The flow-through was removed, and the remaining volume in the initial collecting tube was added to the spin filter tube, which was then centrifuged for 1 minute. The spin filter (containing DNA to be isolated) was transferred to a new 2 mL collection tube, and 650 μL of MD4 was added to the spin filter. The tube was centrifuged for 1 minute, and the flow through was removed. The spin filter tube was centrifuged again for 2 minutes, and then, the spin filter was placed into a new collection tube. Finally, 50 μL of deionized H_2O was added to the spin filter within the tube, which was centrifuged for 1 minute. The spin filter was discarded, and isolated DNA in the collection tube was stored at -22°C. Two extraction controls using sterile filter membranes were also performed using the above protocol. Following extraction, some selected sample DNA concentrations were quantified using a Qubit 3.0 Fluorometer following the manufacturer's protocol (Thermo Fischer Scientific, Inc., Waltham, MA).

DNA Amplification

For polymerase chain reaction (PCR) amplification, isolated DNA was amplified via 16S rRNA gene Illumina tag PCR in accordance with the protocols from the Earth Microbiome Project. All PCR reactions were performed in duplicate for each of the two filter membranes acquired per each sampling condition (four total reactions per 10-hour sampling interval). One unique barcoded reverse primer was used for all four PCR reactions for each sampling condition.

Primers were of the 2015 HHMI set from Juniata College. The PCR master mix for each reaction was composed of 0.13 μL of Ex Taq polymerase (final concentration: 0.65 U/rxn), 2.5 μL of Ex Taq Buffer (final concentration: 1X), and 2 μL of dNTP mixture (final concentration: 0.2 mM) from TAKARA BIO, Inc., Shiga, Japan as well as 1 μL of non-barcoded forward primer, 515f, at a final concentration of 0.2 μM (Integrated DNA Technologies, Inc., Coralville, IA). Each 25.0 μL PCR reaction tube contained 5.63 μL of the master mix, 18.37 μL of template DNA, and 1 μL of barcoded reverse primer, 806r, at a final concentration of 0.2 μM (Integrated DNA Technologies, Inc., Coralville, IA). Three PCR control reactions were performed under identical conditions except that 18.37 μL of PCR-grade deionized H_2O was substituted for DNA. Thermocycling was conducted in a PTC-200 Peltier Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules, CA). The cycling conditions were 94°C for 3 minutes to denature DNA, 94°C for 45 seconds, 50°C for 60 seconds, and 72°C for 90 seconds. The latter three temperature cycles were repeated 35 times. Then, the final extension was conducted at 72°C for 10 minutes, and samples were held at 4°C. Amplicons were then stored at -22°C.

DNA Sequencing

Next-generation sequencing of the 28 unique barcoded PCR amplicons (24 air samples, one extraction control, and three PCR controls) was performed at Juniata College in Huntington, PA using the Illumina MiSeq platform (Illumina, Inc., San Diego, CA). Initial sequence library assembly was performed there as well.

Sequence Analysis

Computation analysis was performed using the Quantitative Insights into Microbial Ecology (QIIME) statistical package. Access to a compute cluster (Nor-Tech, Burnsville, MN) was provided by Juniata College in Huntington, PA. Sequencing data were processed in accordance with the Metagenomics Workshop Instructional Booklet (Lamendella et al., 2015). To start, multiplexed sequences were paired to merge forward and reverse reads for the target V4 region of the 16S rRNA gene. During quality filtering, reads beyond 253 base pairs or with an average expected error of 0.5% or greater (base assignment error for every 1 out of 200 bases) were discarded.

Then, operational taxonomic units (OTUs) were selected in accordance with the open reference model. Briefly, reads were first compared to the Greengenes 16S rRNA reference database. Sequences that did not match the library were clustered internally among other unmatched sequences in a *de novo* format. After clustering, chimeras and other DNA artifacts were removed from the sequences. Data were then converted into the BIOM table format, taxonomy was assigned, and a phylogenetic tree was constructed. Using QIIME 1.9.0, the OTU table was filtered by OTU abundance to remove rare taxa present at 0.005% or less. Then, taxa were summarized at levels ranging from phylum to genus.

In terms of statistical analysis, one-tailed Student's T-tests were calculated to ascertain significant differences between total sequence counts across sampling conditions using Excel (Microsoft, Redmond, WA). Genera richness was calculated as the number of genera present in each sample. Shannon's equitability for genera evenness, or relative abundance, was calculated using the equation in Figure 5. A one-way analysis of variance (ANOVA) was performed for

selected sampling conditions versus mean richness and evenness calculations using SPSS Statistics 22 (IBM, Armonk, NY).

$$\text{Sample Evenness} = \frac{-\Sigma P*\ln(P)}{\ln(R)}$$

P=genus proportion within sample R=genera richness for that sample

Figure 5: Equation for calculation of genera evenness. Ln is the natural logarithm.

Results

Quantitative and qualitative results were obtained from the analysis of indoor microbial communities. Initially, Qubit DNA concentration values obtained following DNA extraction but prior to PCR were assessed for a select number of samples and are displayed in Table 3.

Table 3. Selected DNA concentrations prior to PCR for two sampling conditions

Sample (date and condition)	**Concentration in ng/μL**
10/20/2015 HA carpet	0.107 (L); 0.038 (R)
10/20/2015 LA carpet	0.125 (L); 0.04 (R)

Note: L/R refers to left or right positions of the analytical test filter funnel membranes.

Across 14 samples and one control, 227 sequences were obtained with a mean length of 151 bases. The 227 value refers to sequences that had undergone some initial quality filtering but had not yet been paired. Of the original 28 PCR reactions (24 samples, three PCR controls, and one extraction control), 13 yielded no sequenceable DNA. The distribution of sequences across samples is shown in Table 4. From this sequence per sample data, sequence counts across sampling conditions were generated, and a Student's t-test was conducted to ascertain if the differences between the number of sequences generated in each condition were significant

(Figure 6 and 7). In comparing the number of sequences for high activity (163) versus low activity (54) samples, a statistically significant difference was obtained (P=0.0315). However, there was no significant difference between the number of sequences obtained for carpet (96) verses linoleum (121) as P=0.4471.

Table 4. Distribution of sequences across varying activity levels and flooring types

Sampling Condition	Activity Level	Flooring	# Seqs
10/22, Thurs.	High	Linoleum	44
11/16, Mon.	High	Carpet	34
10/26, Mon.	High	Carpet	25
11/4, Wed.	Low	Linoleum	18
11/12, Thurs.	High	Linoleum	16
10/20, Tue.	High	Carpet	13
10/28, Wed.	High	Linoleum	11
11/9, Mon.	High	Linoleum	11
10/20, Tue.	Low	Carpet	11
Extraction control		N/A	10
10/27, Tue.	High	Carpet	9
11/11, Wed.	Low	Linoleum	8
10/28, Wed.	Low	Linoleum	7
10/22, Thurs.	Low	Linoleum	6
10/27, Tue.	Low	Carpet	4

Note: Counts for post-initial quality filtering, unpaired sequence data.

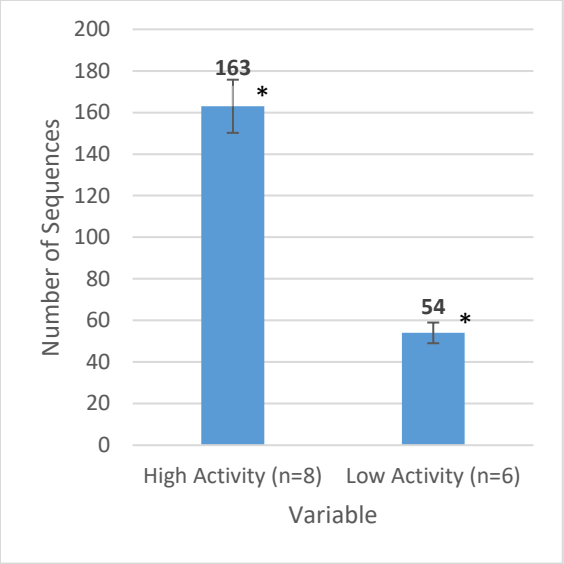

Figure 6. Sequence counts for all samples based upon activity level. Asterisk indicates a significant difference ($P<0.05$).

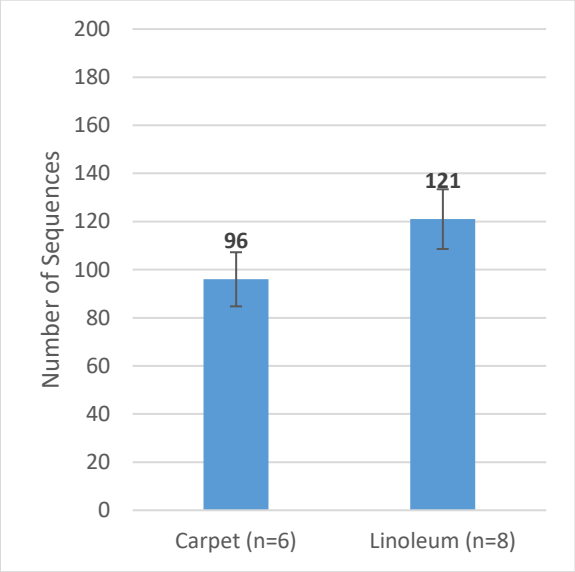

Figure 7. Sequence counts for all samples based upon flooring type. No significant difference between flooring types was found.

Once forward and reverse reads were merged to yield combined reads of ~250 bases, sequence quality was assessed. A Phred quality score and an average expected error value were assigned to each nucleotide position across reads. The Phred score distribution, also known as average Q, is shown in Figure 8. It indicates the mean accuracy of the sequencing at every single

nucleotide in all reads. All nucleotide identity assignments were above a score 30, though the score for nucleotides between position 100-150 was elevated above 40. Since average Q is logarithmic, a score of 30 means 99.9% accuracy, and a score of 40 means 99.99% accuracy. The average expected error was also calculated and is shown in Figure 9. Average expected error is the cumulative error of base identification across a read. Except for one outlier beyond 250 bases, all assigned nucleotides had an error below 0.4%, or an error rate of 1 per every 250 bases (Figure 9). Given this sequence quality data, reads with an average expected error beyond 0.5% and possessing more than 250 bases were subsequently discarded, which left 84.26% of the 216 sequences obtained post-merge pairs for downstream analyses. After this step, 12 samples and the control extraction remained containing sequences of sufficient quality.

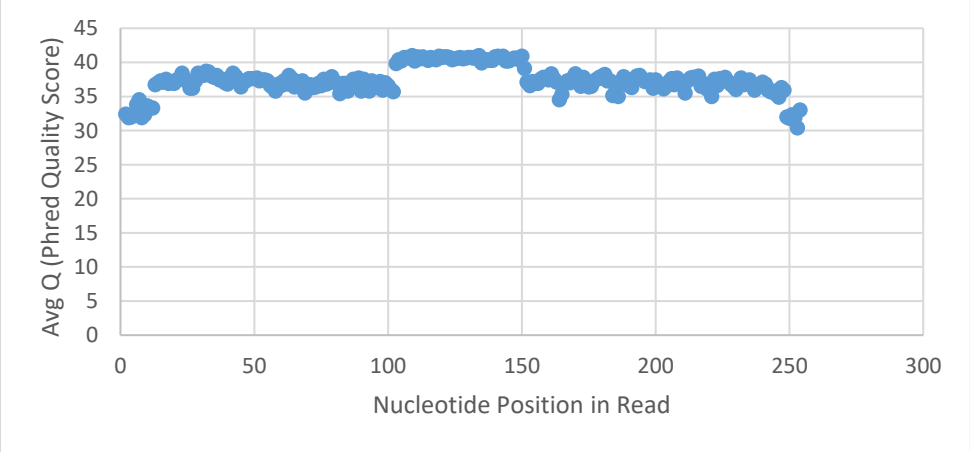

Figure 8. Average Phred quality score at each nucleotide position across all sequences.

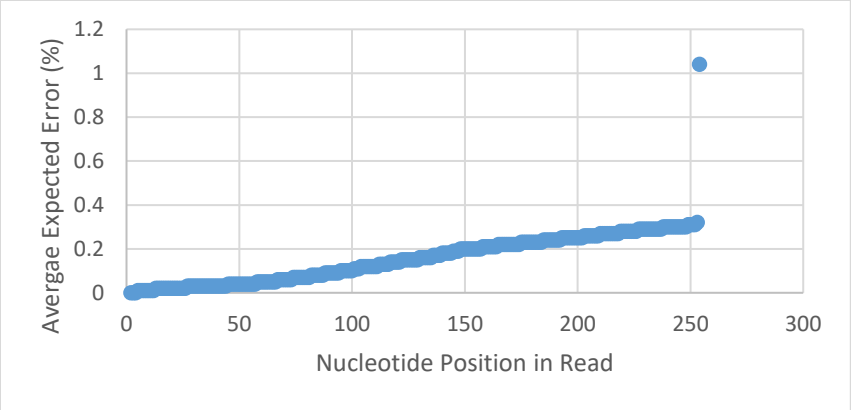

Figure 9. Average expected error at each nucleotide position across all sequences. Note that error compounds as nucleotide position advances.

OTU picking for the remaining sequences yielded taxonomic group distributions across various phylogenetic resolutions. Summaries were generated using the summarize_taxa.py command. Figure 10 depicts the distribution of genera across the samples containing sequences and the extraction control. Figure 11 is a quantitative interpretation of Figure 10, indicating the number of genera and unique genera present across the sampling conditions. Genera that were shared across two or more sampling conditions are shown in Table 5.

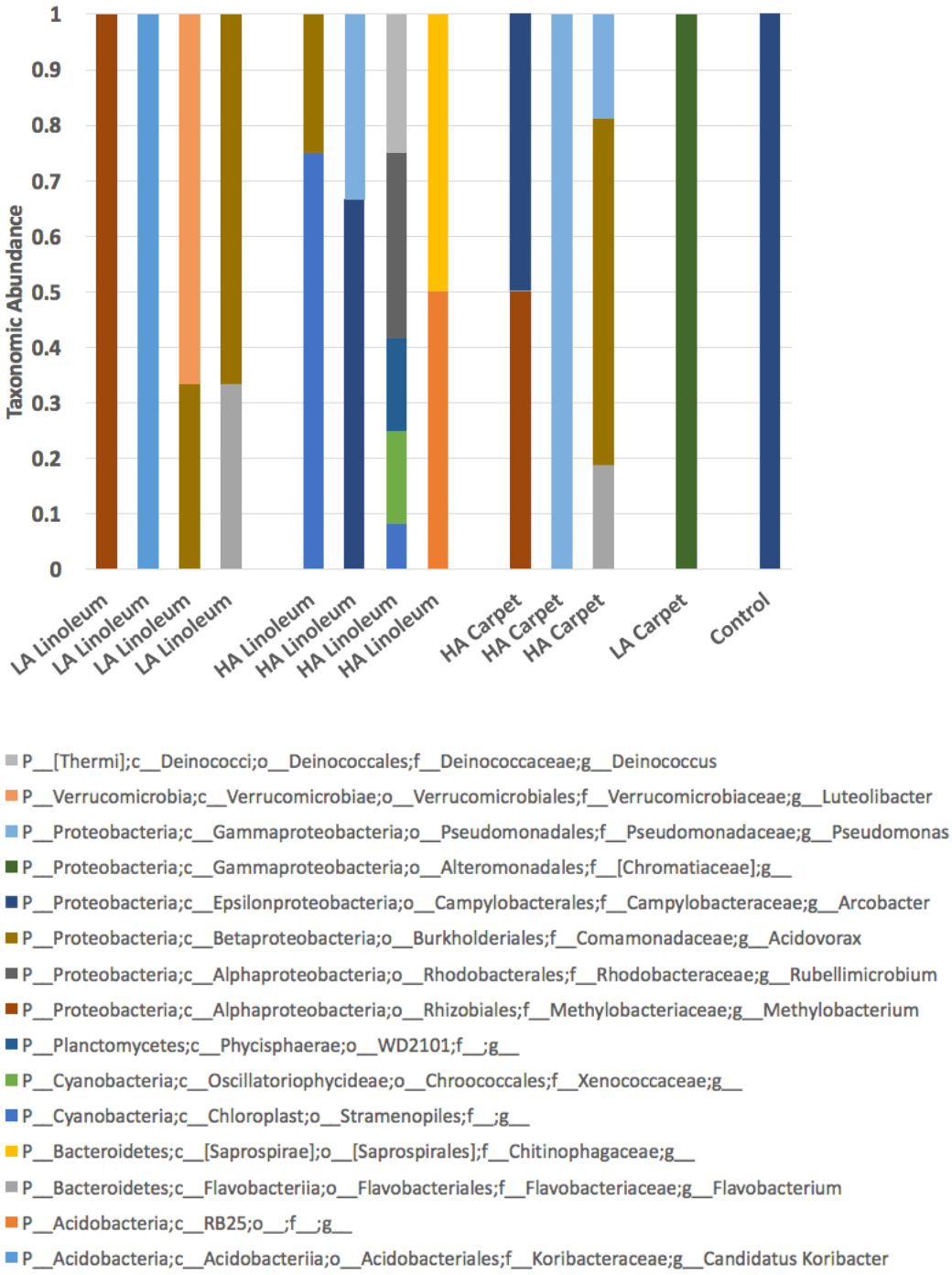

Figure 10. Fractional abundances for genera across sampling conditions and control. HA=high activity, LA=low activity. Note that in the legend, some genera were not specified as sufficient taxonomic data were not available to classify at that particular phylogenetic rank.

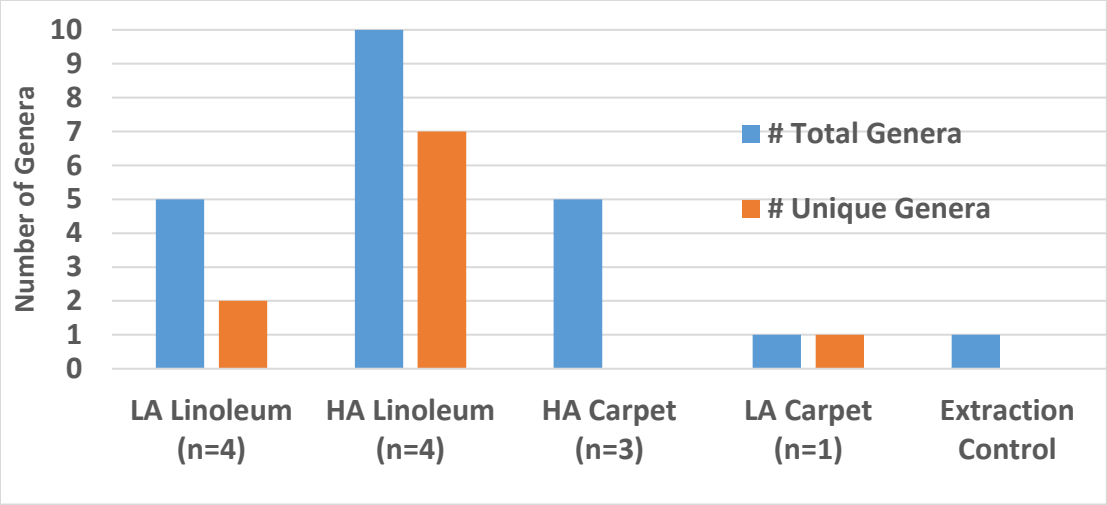

Figure 11. Total genera and unique genera across sampling conditions.

Table 5. Genera shared across two or more sampling conditions.

Genus (Phylum)	**LA Linoleum**	**HA Linoleum**	**LA Carpeting**	**HA Carpeting**	**Control**
Acidovorax (Proteobacteria)	X	X		X	
Arcobacter (Proteobacteria)		X		X	X
Flavobacterium (Bacteriodetes)	X			X	
Methylobacterium (Proteobacteria)	X			X	
Pseudomonas (Proteobacteria)		X		X	

Note: X marks genus presence in any of the sampling conditions.

Genera distribution data were also used to calculate genera richness and evenness across sampling conditions. Genera richness is displayed in Figure 12. A one-way ANOVA was conducted for genera richness comparing the three sampling conditions containing at least one replicate: low activity linoleum, high activity linoleum, and high activity carpeting. Means are

shown in Table 6, and the results of the ANOVA are in Table 7. The ANOVA yielded F=1.296 and P=0.325, indicating no statistically significant difference for genera richness across the samples.

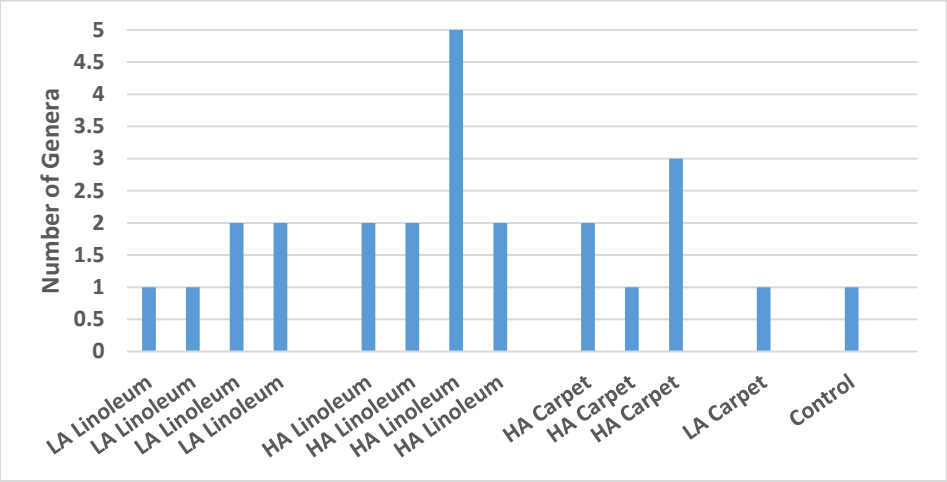

Figure 12. Genera richness across each sampling condition and control.

Table 6. Genera richness means and standard deviations for low activity linoleum, high activity linoleum, and high activity carpet.

	N	Mean	Std. Deviation
LA Linoleum	4	1.5	0.58
HA Linoleum	4	2.8	1.5
HA Carpet	3	2.0	1.0
Total	11	2.1	1.1

Table 7. Results from ANOVA test for the three conditions of low activity linoleum, high activity linoleum, and high activity carpet versus genera richness. P value was not statistically significant.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3.159	2	1.580	1.296	0.325
Within Groups	9.750	8	1.219		
Total	12.909	10			

Genera evenness was determined for each sample. Shannon's equitability or evenness is given on a 0 to 1 scale, with 1 being perfectly even and 0 being perfectly uneven. Evenness was calculated using the equation shown in Figure 5. Evenness values for each sample are displayed in Figure 13. A one-way ANOVA was conducted for genera evenness comparing the three sampling conditions containing at least one replicate: low activity linoleum, high activity linoleum, and high activity carpeting. Means are shown in Table 8, and the results of the ANOVA are in Table 9. The ANOVA yielded F=0.119 and P=0.89, indicating no statistically significant difference for genera evenness across the samples.

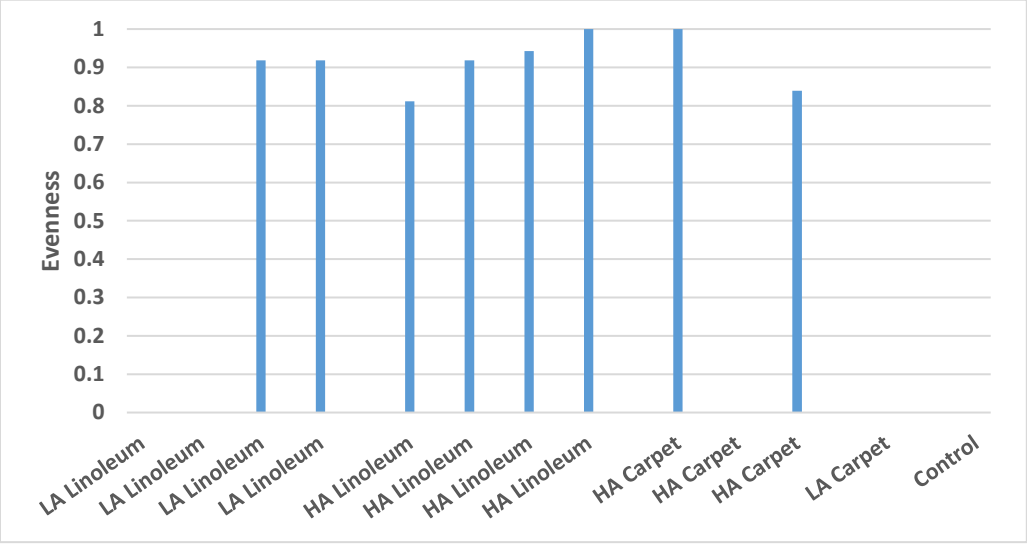

Figure 13. Genera evenness across sampling conditions. Samples without a bar indicate a sample where evenness was mathematically undefined given that richness was 1.

Table 8. Genera evenness means and standard deviations for low activity linoleum, high activity linoleum, and high activity carpet.

	N	Mean	Std. Deviation
LA Linoleum	3	0.89	0.046
HA Linoleum	4	0.92	0.079
HA Carpet	2	0.92	0.11
Total	9	0.91	0.068

Table 9. Results from ANOVA test for the three conditions of low activity linoleum, high activity linoleum, and high activity carpet versus genera evenness. P value acquired was not statistically significant.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.001	2	.001	.119	0.890
Within Groups	.036	6	.006		
Total	.037	8			

Discussion

Air samples were collected under varying levels of activity and flooring types in a college academic building. Extracted DNA from membrane filters was amplified and sequenced via the Illumina method. Sequence data were processed using QIIME and analyzed both quantitatively qualitatively.

General Sequence Analysis

The first initial data point to be considered are the sample DNA concentrations from filter membrane extracts prior to PCR. These values served as a nonspecific indicator of the biotic content of indoor air. As shown in Table 3, DNA concentrations were assayed using QuBit for the left and right filter membranes in two different sampling conditions on a range of 0.038-0.125 ng/μL. Furthermore, DNA levels were undetectable in some instances (data not shown). Given that there are only two data points, no substantial conclusions can be drawn. In comparison to other phylogenetic studies, such as those for blood samples, DNA extraction concentrations were typically in the 50-200 ng/μL range (Psifidi et al., 2015), which are several orders of magnitude greater than what was observed in the present study. However, there is evidence that indoor microbial DNA concentrations are typically far lower than those from other sampling locations. One protocol review using much larger filters and higher flow rates yielded

concentrations on a range of 0.073-0.632 ng/μL, which are comparable with the present study (Jiang et al., 2015). Therefore, there is some weak evidence that obtained DNA concentrations may be consistent with the literature.

Total sequence counts and distribution across sampling conditions served as a pre-QIIME analysis metric for the biotic content of indoor air. Prior to merging the forward and reverse reads, 227 initial sequences were yielded for one extraction control and 14 of the 24 experimental samples. In comparison, a phylogenetic study of indoor air within a hospital environment reported 179,146 initial sequences (Kembel et al., 2012), and a study of a university classroom using next-generation sequencing yielded 10,675 sequences after some quality filtering/chimera removal (Hospodsky et al., 2012). Thus, the number of sequences yielded in the present study was substantially lower than the numbers found in other studies.

As shown in Table 4, samples that did yield sequences were distributed across all four treatment combinations. The maximum number of sequences in any one sample was 44 (high activity linoleum) while the minimum was 4 (low activity carpet). Interestingly, the extraction control yielded 10 sequences, a substantial percentage of the total sequence number. Had sequence counts been several orders of magnitude larger, the taxonomic data yielded from the extraction control could have been subtracted out from the rest of the experimental samples. Additionally, 10 samples of the original 24 did not yield any detectable sequences, suggesting methodological difficulties associated with sample collection or DNA extraction.

While the 227 total count is not reflective of the final number of sequences to be used for taxonomic analysis (the analyzed sequence counts were actually lower), it can be used to illustrate comparative differences in microbial diversity for indoor air. As shown in Figure 6, the number of sequences yielded from high activity samples was significantly elevated in

comparison to those in low activity samples. This is consistent with expectations, given the evidence that human occupancy is a primary contributor of taxa via resuspension and particulate shedding in indoor environments (Adams et al., 2015; Hospodsky et al., 2012). During lower activity levels (nighttime sampling periods), fewer disturbance events would reduce the biotic content of indoor air. Thus, there is some limited empirical support for the original hypothesis predicting a greater number of sequences for high activity samples.

However, for the total sequence count comparison between flooring types (Figure 7), there was no significant difference apparent for carpeting versus linoleum sample sequence quantities. This is contrary to the original hypothesis, which predicted elevated sequences counts for carpeting versus linoleum. Prior research showed that carpet facilitated particulate resuspension in comparison to more regular, solid surfaces (Adams et al., 2015; Mukai et al., 2009) and also demonstrated that particle release at larger size ranges was elevated for human traffic in a carpeting condition versus a plastic sheeting condition (Hospodsky et al., 2012). Nevertheless, sequence counts were much lower than expected and likely grossly underestimated microbial diversity, which could account for the surprising results. Given the low sequence yields, it would be imprudent to draw any substantial conclusions from the data set.

Sequence quantity is just one metric to assess the reliability of Illumina sequencing data. Accuracy of nucleotide identification is another means to evaluate sequencing reads prior to phylogenetic analysis. As shown in Figures 8 and 9, the average Phred quality score and the average expected error illustrated that though sequence counts were low overall, they met the quality benchmarks established by the literature. For average Q, a score of 30 or greater is considered to be of sufficient quality for inclusion in a data set (Jalali et al., 2015). As shown in Figure 8, average Q was consistently above the threshold, indicating that the vast majority of the

sequences were high quality. The only slight depression below 30 occurred near the 250 base pair point, which is logical given that 253 bp is the maximum length threshold for the read. Interestingly, quality was elevated beyond 40 for nucleotides at positions between 100 and 150 bp within a read. When forward and reverse reads (approx. 150 bases each) of the V4 region were merged, there is some overlap at the midpoint region. Hence, the nucleotide repetition helps to corroborate base assignment accuracy, elevating the Phred score for that interval.

Average expected error is another quality metric indicating the percentage chance that a nucleotide's identity was incorrectly assigned (Figure 9). In general, sequences with an expected error beyond 0.5% are excluded from a data set (Puente-Sanchez et al., 2015). Again, given that the vast majority of the nucleotides in the reads had an error below 0.5%, the sequence data were of high quality despite a paucity of overall sequence counts.

Phylogenetic Analysis

Progression through the QIIME analysis pipeline in accordance with the protocol by Lamendella et al. (2015) generated summaries of taxonomic abundances across samples for quality filtered, chimera-free sequences. Across 12 samples and one control, 15 distinct genera (or families/orders if genera identification could not be resolved) were reported (Figure 10). Qualitatively, there appears to be some evidence that the high activity linoleum samples were more diverse (contained more total taxa) than low activity linoleum samples. This finding was corroborated by quantitative enumeration of total genera and unique genera present across sampling conditions (Figure 11). High activity linoleum had 10 total genera and 7 unique genera, which in both cases exceeded the values for the other three conditions and the control. No

comparisons between high and low activity carpet were possible due to the absence of any replicates with measureable sequences beyond a single sample for low activity carpeting.

Just as total sequence counts were used as a beta diversity metric to evaluate the validity of the original hypotheses, genera abundances can also be considered in light of those predictions. These data provide some evidence, at least for samples collected over linoleum, supporting the original hypothesis that higher activity samples would be more diverse than lower activity samples (Figure 10), corroborating sequence count data of Figure 6. However, there was not sufficient evidence to suggest that samples collected over carpeting were more or less taxonomically diverse than those collected over linoleum given the statistical power limitations for low activity carpet. Once again, methodological shortcomings precluded the generation of any scientifically valid generalizations.

Computation beta diversity could not be calculated through QIIME due to sampling limitations. However, qualitative comparisons were conducted to assess which taxonomic groups were shared across samples. As shown in Table 5, only five genera were shared across two or more conditions, and no genera were shared across all four conditions. *Acidovorax* was shared across three conditions, and the genus includes several species that act as plant pathogens (Giordano et al., 2012). *Arcobacter* was shared across two conditions and the control. This genus has been found in human fecal samples, and a subset of species has been classified as enteropathogens (Collado & Figueras, 2011). *Flavobacterium* and *Methylobacterium* were shared across the same two conditions, LA linoleum and HA carpeting. *Flavobacterium* has been linked to diverse outdoor locations, including soil and water (Bernardet et al., 1996), while *Methylobacterium* is a prolific airborne organism in indoor and outdoor environments (Brenner & Krieg, 2006). *Pseudomonas*, shared across both high activity conditions, is a highly

heterogeneous genus classified in aquatic as well as terrestrial locations. Some species are pathogenic to plants and animals (Moore et al., 2006).

However, there is empirical evidence that many of the genera positively identified in samples, such as *Acidovorax*, *Flavobacterium*, *Methylobacterium,* and *Pseudomonas,* are common contaminants of DNA extraction kits (Salter et al., 2014). Though *Arcobacter* was not listed in the study by Salter and colleagues (2014), it was found in the control extraction and may likely be a contaminant as well. While these data were unable to implicate any genera in forming a core microbiome, meaning that the original hypothesis was not supported, it does present a constrained view of some potential microbial homology among indoor air samples.

Ecological Diversity Analysis

Independent of the QIIME pipeline, the alpha diversity metrics of genera richness and evenness were calculated from the genera abundances presented in Figure 10. In terms of the genera richness in Figure 12, no clear trends were readily discernable. Consequently, an ANOVA for a portion of the data set (low activity linoleum, high activity linoleum, and high activity carpet) was performed. The mean richness for high activity linoleum was slightly elevated in comparison to the other treatments (Table 6), but the ANOVA found no significant differences among richness data for the three selected sampling conditions of low activity linoleum, high activity linoleum, and high activity carpet (Table 7).

Genera evenness, graphically displayed in Figure 13, appeared similar across sampling conditions upon initial qualitative analysis. Given that “1” represents complete genera evenness, there is evidence that all indoor sampling conditions were relatively even, as evidenced by the overall mean evenness of 0.91 for low activity linoleum, high activity linoleum, and high activity

carpet (Table 8). Thus, indoor environments as a whole may have elevated general microbial biodiversity, at least in terms of evenness. As with richness, the ANOVA was not significant for differences in genera evenness across the three sampling conditions (Table 9). Taken together in the context of the present study's limitations, the richness and evenness data suggest that the diversity of indoor air in a college academic building was not affected by variations in flooring type and human activity levels. This evidence runs counter to literature studies, as other investigators have recorded a 100% increase in OTU richness with elevated human occupancy in an indoor environment (Adams et al., 2015). Other computation alpha diversity metrics, such as PD_whole_tree and chao1, were attempted but could not be completed due to insufficient sequence counts across samples.

Limitations

Many limitations for the present study have already been discussed, namely in terms of few total sequence counts and the low fraction of samples possessing detectable sequences. The reasons for this deficit are numerous. One postulated explanation is that the DNA extraction protocol was faulty and consequently a substantial percentage of DNA in each sample was not isolated. Since this was an adapted protocol, rather than one identical to a manufacturer's specifications, the likelihood of errors was increased. Nevertheless, the extraction concentrations were consistent with a literature source (Jiang et al., 2015), and there is evidence that pre-PCR DNA concentrations can often be undetectable (Dr. Ann Klein, personal communication, April 9, 2015). Therefore, other mechanisms could potentially be implicated to explain the low sequence yields across samples.

Another possible outcome, which would still involve a deficit in the extraction protocol, is that isolated DNA was of poor quality. In theory, poor quality DNA would still be quantified by the Qubit assay, which is selective for dsDNA only. However, the Qubit metric provided no information as to the quality of the V4 region to be amplified within the DNA isolates. Thus, perhaps there was minimal PCR amplification within isolated DNA due to deficits in reaction kinetics or the presence of small molecule PCR inhibitors (Dr. Ann Klein, personal communication, November 16, 2015). An additional potential error mechanism may have been with the PCR cycling conditions. As shown in Jiang et al. (2015), there is evidence that PCR conducted with isolated airborne genetic material requires greater than 34 cycles for minimal amplification and that maximum amplification could be achieved with 48-56 cycles. Hospodsky et al. (2012) used 45 cycles for their amplification of air sample DNA. In contrast, the present experiment used 35 cycles, which was suggested in the Earth Microbiome Project protocols. However, that recommended cycling value may not have been sufficient for DNA originating from air samples. Future manipulations of the present study are required to more closely elucidate the effect of PCR cycle elongation on DNA amplification.

The conditions used for sampling collection, namely sampling duration and flow rate, may have contributed to experimental error. A flow rate of 7 LPM and a sampling period of 10 hours were utilized to achieve a total air volume of 4.2 m^3 per sample. This is above the approximately 3 m^3 volume minimum for indoor air studies (Dr. James Meadow, personal communication, January 16, 2015). However, implementation of a more powerful air sampler would have increased the flow rate and therefore the deposition of biotic and abiotic materials onto the filter membranes. Other indoor air studies have used a 10 LPM flow rate over 22 hours (Hospodsky et al., 2012) or 28.3 LPM for an unspecified interval (Qian et al., 2012). Therefore,

the flow rate and sampling interval will need to be optimized in future studies. Also, a less porous filter could have been implemented to be more selective than the 0.45 μm one used in the present study to increase cell retention on the membrane filter. Adams and colleagues (2015) utilized a 0.2 μm filter membrane for sampling. They also placed air samplers at 1.5 m from the ground and positioned them upside down. This differs from the present study that placed filters at a height of 1 m and positioned them approximately horizontally. Future studies will have to examine the affect of altered filter placement on bioaerosol collection.

Conclusion

Air samples over two different flooring types and activity levels were collected in a college academic building in Fall 2015. Following DNA isolation, 16S rRNA gene amplification, and Illumina sequencing, 227 unpaired sequences were generated initially for 14 samples and one extraction control. Though limited sequence counts precluded extensive analysis, several metrics were utilized to assess alpha (within sample) and beta (between sample) diversity. In terms of alpha diversity, statistical analysis of species richness and evenness found no significant differences across low activity linoleum, high activity linoleum, and high activity carpet conditions. For beta diversity, samples collected under high activity levels had a statistically greater number of sequences than those collected under low activity. No significant difference was apparent for sequences yielded from carpeting versus linoleum samples. Five total genera were shared across at least two or more sampling conditions. Phylogenetically, samples collected over high activity periods and linoleum flooring had the greatest number of total genera and unique genera in comparison to the other three conditions and controls.

These results are both supported and refuted by numerous literature studies. More importantly, the statistical power limitations of the present study, both in terms of total number of sequences and total number of samples containing sequences, prevented the postulation of any conclusions from the data set. Future investigations will seek to reevaluate the DNA extraction and PCR protocol to improve pre-sequencing DNA quality and concentration. Revisions to sampling flow rate, duration, and means of collection are also necessary to sample indoor environments at a more comprehensive depth. Therefore, despite methodological and analytical challenges, this study provides a limited perspective on how the variables of anthropogenic traffic and flooring type may be affecting microbial diversity in the indoor built environment.

Acknowledgements

I would like to acknowledge financial support from the Sigma-Xi Grants-In-Aid Research Foundation [#G201510151560786], Beta Beta Beta Research Scholarship Foundation Fund, Elizabethtown College Department of Biology, Elizabethtown College Honors Program Academic Grants, and Certified Carpet of Lancaster. For logistical and analytical support, I would like to thank the Wohl Lab of the Elizabethtown College Department of Biology, Elizabethtown College Facilities Management, the GCAT-SEEK Consortium at Juniata College, and Dr. Ann Klein and Dr. James Meadow at the University of Oregon.

References

Adams, R. I., Bhangar, S., Pasut, W., Arens, E. A., Taylor, J. W., Lindow, S. E., Nazaroff, W. W., & Bruns, T. D. (2015). Chamber bioaerosol study: Outdoor air and human occupants as sources of indoor airborne microbes. *PLOS ONE, 10*(5), 1-18.

Bernardet, J. F., Segers, P., Vancanneyt, M., Berthe, F., Kersters, K., & Vandamme, P. (1996). Cutting a gordian knot: Emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* norn. nov. (Basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *International Journal of Systematic Bacteriology, 46*(1), 128-148.

Brenner, D. J., & Krieg, N. R. (2006). *Bergey's Manual® of Systematic Bacteriology: Volume Two: The Proteobacteria, Part 3* (2nd ed.). East Lansing: Springer, p. 568.

Chen, Q., & Hildemann, L. M. (2009). Size-resolved concentrations of particulate matter and bioaerosols inside versus outside of homes. *Aerosol Science and Technology, 43*, 699-713.

Collado, L., & Figueras, M. J. (2011). Taxonomy, epidemiology, and clinical relevance of the genus *Arcobacter*. *Clinical Microbiology Reviews, 24*(1), 174-192.

Corsi, R. L., Kinney, K. A., & Levin, H. (2012). Microbiomes of built environments: 2011 symposium highlights and workgroup recommendations. *Indoor Air, 22*, 171-172.

Ferro, A. R., Kopperud, R. J., & Hildemann, L. M. (2004). Source strengths for indoor human activities that resuspend particulate matter. *Environmental Science & Technology, 38*(6), 1759-1764.

Giordano, P. R., Mitkowski, N. A., & Vargas, J. M. (2012). Identification, characterization, and distribution of *Acidovorax avenae* subsp. *avenae* associated with creeping bentgrass etiolation and decline. *Plant Diseases, 96*(12), 1736-1742.

Hospodsky, D., Qian, J., Nazaroff, W. W., Yamamoto, N., Bibby, K., Rismani-Yazdi, H., & Peccia, J. (2012). Human occupancy as a source of indoor airborne bacteria. *PLOS ONE, 7*(4), 1-10.

Hughes Martiny, J. B., Bohannan, B. J., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M, Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Ovreas, L., Reysenbach, A., Smith, V. H., & Staley, J. T. (2006). Microbial biogeography: Putting microorganisms on the map. *Nature Reviews: Microbiology, 4*, 102-112.

Jalali, S., Kohli, S., Latka, C., Bhatai, S., Vellarikal, S. K., Sivasubbu, S., Scaria, V., & Ramachandran, S. (2015). Screening currency notes for microbial pathogens and antibiotic resistance genes using a shotgun metagenomic approach. *PLOS ONE, 10*(6), 1-15.

Jiang, W., Liang, P., Wang, B., Fang, J., Lang, J., Tian, G., Jiang, J., & Zhu, T. F. (2015). Optimized DNA extraction and metagenomic sequencing of airborne microbial communities. *Nature Protocols, 10*(5), 768-779.

Kelley, S. T., & Gilbert, J. A. (2013). Studying the microbiology of the indoor environment. *Genome Biology, 14*(202), 1-9.

Kembel, S. W., Jones, E., Kline, J., Northcutt, D., Stenson, J., Womack, A. M., Bohannan, B. J., Brown, G. Z., & Green, J. L. (2012). Architectural design influences the diversity and structure of the built environment microbiome. *The ISME Journal, 6*, 1469-1479.

Kembel, S. W., Meadow, J. F., O'Connor, T. K., Mhuireach, G., Northcutt, D., Kline, J., Moriyama, M., Brown, G. Z., Bohannan, B. J., & Green, J. L. (2014). Architectural design drives the biogeography of indoor bacterial communities. *PLOS ONE, 9*(1), 1-10.

Kildeso, J., Vinzents, P., & Schneider, T. (1999). A simple method for measuring the potential resuspension of dust from carpets in the indoor environment. *Textile Research Journal, 69*(3), 169-175.

Konopka, A. (2009). What is microbial ecology? *The ISME Journal, 3*, 1223-1230.

Lamendella, R., Ulrich, N., Hager, N., & Tokarev, V. (2015). *Metagenomics Workshop Instructional Booklet.* GCAT-SEEK Metagenomics Workshop. Huntingdon: Juniata College.

Meadow, J. F., Altrichter, A. E., Kembel, S. W., Kline, J., Mhuireach, G., Moriyama, M., Northcutt, D., O'Connor, T. K., Womack, A. M., Brown, G. Z., Green, J. L., & Bohannan, B. J. (2014). Indoor airborne bacterial communities are influenced by ventilation, occupancy, and outdoor air source. *Indoor Air, 24*, 41-48.

Moore, E. R., Tindall, B. J., Martins Dos Santos, V. A., Pieper, D. H., Ramos, J., & Palleroni, N. J. (2006). Nonmedical: Pseudomonas. *Prokaryotes, 6*, 646-703.

Mukai, C., Siegel, J. A., & Novoselac, A. (2009). Impact of airflow characteristics on particle resuspension from indoor surfaces. *Aerosol Science and Technology, 43*, 1022-1032.

Psifidi, A., Dovas, C. I., Bramis, G., Lazou, T., Russel, C. L., Arsenos, G., & Banos, G. (2015). Comparison of eleven methods for genomic DNA extraction suitable for large-scale whole-genome genotyping and long-term DNA banking using blood samples. *PLOS ONE, 10*(1), 1-18.

Puente-Sanchez, F., Aguirre, J., & Parro, V. (2015). A novel conceptual approach to read-filtering in high-throughput amplicon sequencing studies. *Nucleic Acids Research*, 1-11.

Qian, J., Ferro, A. R., & Fowler, K. R. (2008). Estimating the resuspension rate and residence time of indoor particles. *Air & Waste Management Association, 58*, 502-516.

Qian, J., Hospodsky, D., Yamamoto, N., Nazaroff, W., & Peccia, J. (2012). Size-resolved emission rates of airborne bacteria and fungi in an occupied classroom. *Indoor Air, 22*, 339-351.

Rintala, H., Pitkaranta, M., Toivola, M., Paulin, L., & Nevalainen, A. (2008). Diversity and seasonal dynamics of bacteria community in indoor environment. *BMC Microbiology*, *8*(56), 1-13.

Rosati, J. A., Thornburg, J., & Rodes, C. (2008). Resuspension of particulate matter from carpet due to human activity. *Aerosol Science and Technology, 42*, 472-482.

Salter, S. J., Cox, M. J., Turek, E. M., Callus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J., & Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology, 12*(87), 1-12.

Schneider, T. (2008). Dust and fibers as a cause of indoor environmental problems. *Scandinavian Journal of Work, Environment & Health Supplement*, *4*, 10-17.

Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nature Methods, 5*(1), 16-18.

Thatcher, T. L., & Layton, D. W. (1995). Deposition, resuspension, and penetration of particles within a residence. *Atmospheric Environment, 29*(13), 1487-1497.

Tringe, S. G., Zhang, T., Liu, X., Yu, Y., Lee, W. H., Yap, J., Yao, F., Suan, S. T., Ing, S. K., Haynes, M., Rohwer, F., Wei, C. L., Tan, P., Bristow, J., Rubin, E. M., & Ruan, Y. (2008). The airborne metagenome in an indoor urban environment. *PLOS ONE*, *3*(4), 1-10.

Ursell, L. K., Metcalf, J. L., Parfrey, L. W., & Knight, R. (2012). Defining the human microbiome. *Nutrition Reviews, 70*(Suppl. 1), S38-S44.

Yarza, P., Yilmaz, P., Pruesse, E., Glockner, F. O., Ludwig, W., Schleifer, K., Whitman, W. B., Euzeby, J., Amann, R., & Rossello-Mora, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology, 12*, 635-645.

You, S., & Wan, M. P. (2015). A risk assessment scheme of infection transmission indoors incorporating the impact of resuspension. *Risk Analysis, 35*(8), 1488-1502.