

Geography and location are the primary drivers of office microbiome composition

Running title: Office microbes

John Chase^{1,2}, Jennifer Fouquier³, Mahnaz Zare⁴, Derek L Sonderegger⁵, Rob Knight^{6,7}, Scott T. Kelley³, Jeffrey Siegel^{4,8}, J Gregory Caporaso^{#,1,2}

¹Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA

²Center for Microbial Genetics and Genomics, Northern Arizona University, Flagstaff, AZ, USA

³Department of Biology, San Diego State University, San Diego, CA, USA

⁴Department of Civil Engineering, University of Toronto, Toronto, ON, Canada

⁵Department of Mathematics and Statistics, Northern Arizona University, Flagstaff, AZ, USA

⁶Department of Computer of Science, University of California San Diego, San Diego, CA, USA

⁷Department of Pediatrics, University of California San Diego, San Diego, CA, USA

⁸Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Address correspondence to: gregcaporaso@gmail.com

Abstract word count: 237

Text word count: 6477

Abstract

North Americans spend the majority of their time indoors where they are exposed to the microbiome of the built environment (BE) they inhabit. Despite the ubiquity of microbes in BEs, and their potential impacts on health and building materials, basic questions about the microbiology of these environments remain unanswered. We present a study on the impacts of geography, material type, human interaction, location in a room, seasonal variation, and indoor and microenvironmental parameters on bacterial communities in offices. Our data elucidates several important features of microbial communities in BEs. First, under normal office environmental conditions, bacterial communities do not differ based on surface material (e.g., ceiling tile or carpet), but do differ based on the location in a room (e.g., ceiling or floor), two features which are often conflated, but which we are able to separate here. We suspect that previous work showing differences in bacterial composition with surface material were likely detecting differences based on different usage patterns. Next, we find that offices have city-specific bacterial communities, such that we can accurately predict which city an office microbiome sample is derived from, but office-specific bacterial communities are less apparent. This differs from previous work which has suggested office-specific compositions of bacterial communities. We again suspect that the difference from prior work arises from different usage patterns. As has been previously shown, we observe that human skin contributes heavily to the composition of BE surfaces.

Importance

Our study highlights several points that should impact the design of future studies of the microbiology of the BEs. First, projects tracking changes to BE bacterial communities should

focus sampling effort on surveying different locations in offices and in different cities, but not necessarily different materials or different offices in the same city. Next, disturbance due to repeat sampling, though detectable, is small compared to other variables, opening up a range of longitudinal study designs in the built environment. Next, studies requiring more samples than can be sequenced on a single sequencing run (which is increasingly common) must control for run effects by including some of the same samples on all sequencing runs as technical replicates. Finally, detailed tracking of indoor and material environment covariates is likely not essential for BE microbiome studies, as the normal range of indoor environmental conditions is likely not large enough to impact bacterial communities.

Introduction

In the United States, humans spend over 90% of their time in built environments (BEs) (1, 2), such as homes, offices, hospitals, and cars. We know that microbes in the BE affect human health (3–5) and the rate of degradation of building materials (6, 7). However, until recently very little was known about the microorganisms that cohabit these environments with us. Over the past decade, molecular microbial diversity studies have shed new light on the spatial and temporal variation of microbial communities in the built environment (8–11). Recent work has revealed how microbial communities, or microbiomes, differ with different building systems (e.g., ventilation mechanisms) (5), how new buildings are colonized by microorganisms (11, 12), and how the human microbiome both impacts and is impacted by the microbiome of homes (13). Differences in microbiomes have been reported across different BE spaces (8, 10, 13), suggesting that our offices and homes have individualized microbiomes, and it has been suggested that microbiome composition differs based on the surface material where the community is found (13–15).

Prior work has not directly tested whether variation in BE microbiomes is mainly due to geographic location, the material that is being sampled, the location in the room that is being sampled, the specific inhabitants, or the environmental conditions that exist within a given indoor environment, all of which have been noted as potential sources of variation but which are difficult to separate. This study aims to expand our basic understanding of the microbiology of BEs by separating factors that are often conflated in BE studies, such as surface material type (e.g., ceiling tile) and location in an office (e.g., ceiling), to understand how factors independently contribute to the composition of BE microbiomes. Similarly, we aimed to understand which, if any, indoor or material microenvironment parameters, such as temperature or humidity, were associated with differences in office microbiomes. Finally, we wished to understand how the human microbiome of an office's inhabitants relates to the personalized office microbiome effect that has previously been suggested (14), and specifically whether this effect extends to surfaces that office inhabitants are not in direct contact with.

Determining the sources of variation in BE microbiomes requires that multiple offices are evaluated, and that multiple locations, material types, and human inhabitants of the offices are sampled in each. Further, if office microbiomes differ geographically, we must survey across a wide enough geographic range that climatic differences are apparent in the indoor environments. We therefore monitored nine offices, three each in Flagstaff, USA, San Diego, USA, and Toronto, CA and collected microbiome samples over the course of a year, along with dense indoor environment data such as temperature, occupancy, and humidity. If office microbiomes differ due to material and/or location in the office, we need to be able to separate these variables to determine their relative contribution. Accordingly, we installed carpet, ceiling tile, and drywall swatches on the floor, walls and ceiling. Microbiome samples were collected in four six-week intensive sampling periods in the summer, fall, winter and spring, and with some

material swatches being sampled more frequently than others. Finally, we collected human microbiome samples from the individuals who performed the sampling, and from 11 inhabitants of one of the offices.

Our design allowed us to evaluate and distinguish the impacts of building material, location in office, sampling frequency, city, office, time, and material and indoor environment covariates on the bacterial communities that established on each of the sample swatches. Our findings provide information on the factors associated with office microbiome composition under normal circumstances, and provide essential information for informed experimental design in future studies of the microbiology of BEs.

Results

Experimental design

To develop our understanding of how microbes establish in built environments over time and a range of building parameters, we sampled nine offices in three cities over a one-year period. The selected cities, San Diego, USA; Flagstaff, USA; and Toronto, CA, are climatically different from one another. Within each city, we chose three offices that were as similar to each other as possible (details on our selection and exclusion criteria are provided in *Materials and Methods*, and on the parameters of our offices in Supplementary Table 1). In each office we installed three sampling plates, with one plate each on the floor, ceiling and wall, as illustrated in Figure 1a. Each plate contained two or three swatches each of painted drywall, ceiling tile, and carpet, allowing us to dissociate location in the room from the material, as well as sensors that allowed us to monitor parameters of the environment including equilibrium relative humidity on the surfaces of the swatches, available light, occupancy, temperature, and relative humidity (Figure 1b). Samples were collected in four six-week sampling periods, one per season. Bacterial and

fungal genetic markers were sequenced from these samples: 16S rRNA gene sequencing was used to profile bacterial communities, and ITS-1 (the non-coding region between the 18S and 5.8S rRNA genes) sequencing was used to profile fungal communities. Our analysis focuses on bacterial data as we obtained less usable data for fungi, which we suspect may be a result of the low biomass of our samples. We present the results of our fungal sequencing in *Supplementary Results*.

Location in office, but not building material, drives community composition

The study was designed so that each material (drywall, carpet, and ceiling tile), was installed in each location (ceiling, floor, and wall), in every office in the study. This design allowed us to differentiate the roles that location within an office and surface material played in determining the richness (how many types of organisms are present) and composition (which taxa are present or absent, and their relative abundances) of bacterial communities on those surfaces.

Using unweighted and weighted UniFrac (16), qualitative and quantitative measures of microbial community dissimilarity, respectively, material was not observed to be a significant driver of bacterial community composition (Figure 2 and Figure 3). Similarly, using Faith's Phylogenetic Diversity (PD) (17), a measure of the phylogenetic richness of a community, material did not appear to be a driver of richness (Figure 1c). The location within an office that a sample was collected from, however, was associated with both community richness (Figure 1c, Figure 3) and community composition (Figure 2). Floor samples were significantly richer than either the ceiling or wall samples in all cities across all sampling periods, except for the first sampling period (two-sample Monte Carlo t-test with 999 permutations, p-values less than 0.05 for all comparisons). We suspect that the lack of difference in the first sampling period is due to the recent sterilization of the materials (see *Methods*). There were no statistical differences in community richness between wall and ceiling samples. This observation is consistent with the

higher deposition rates of larger particles and dust to upward facing horizontal surfaces like floors.

We focused our analysis of surface material and location effects on community composition on the Flagstaff samples taken during the fall and winter sampling periods, which were all sequenced in the same sequencing run to avoid sequencing run being a variable in this analysis. These samples showed significant differences in community composition by the location that was sampled in the office (Permanova with weighted UniFrac pseudo-F=7.80, p -value < 0.0001; unweighted UniFrac pseudo-F=3.54, p -value < 0.0001), but not by the surface material (Permanova with weighted UniFrac pseudo-F=1.17, p -value=0.22, unweighted UniFrac pseudo-F=1.01, p -value=0.89). As with community richness, the floor samples were more different from the wall and ceiling samples, which were virtually indistinguishable from one another. The specific taxa that differed across locations in the office are presented in Supplementary Dataset 1.

Interestingly, within a location, the pattern of differences among samples varied between floors and walls/ceilings. Using unweighted UniFrac, floor samples were more similar to other floor samples than wall/ceiling samples were to other wall/ceiling samples (Figure 2). In contrast, using weighted UniFrac, floor samples were less similar to one another. Thus, community differences among floor samples were driven primarily by differences in the relative abundance of the same OTUs, while the differences among wall/ceiling samples were driven more by the presence or absence of particular OTUs. The same pattern was statistically significant in all cities, but strongest in Flagstaff.

Sampling disrupts microbial communities detectably, but the effect is small

Each sampling plate in the study contained at least two rows of sampling swatches (Figure 1b). During each of the four sampling periods, the “frequently sampled” row of materials was

sampled every other day, while the “infrequently sampled” row was sampled every three weeks. Through this design, we could detect differences in bacterial community richness and composition between frequently and infrequently sampled materials (Figure 1c, Figure 3). Comparing pairs of frequently and infrequently sampled sites, the samples from infrequently sampled surfaces were more rich, as measured by PD, than the frequently sampled surfaces (Monte Carlo t-test 3.75, p-value = 0.0002, n=412) (Figure 1, Figure 3A). This sampling frequency effect was the strongest in floor samples (Figure 3A). While differences in bacterial community composition between frequently and infrequently sampled rows existed, they were not statistically significant (Weighted UniFrac, Permanova, p-value=0.109, pseudo-F=1.63; Unweighted UniFrac, Permanova, p-value= 0.175, pseudo-F=1.13) (Figure 3B-C). The specific taxa that are most different between frequently and infrequently samples sites are presented in Supplementary Dataset 2.

We note that although this sampling frequency effect is present, it is small compared to the effects of biological interest in our study. We clearly see in Figure 3B-C that the compositional differences that we observe as a result of sampling frequency are much smaller than differences observed based on other effects, such as season or location where the sampling plate is installed (i.e., the floor, ceiling, or wall), which are statistically significant.

Cities and offices have signature microbial communities, though the effect is less pronounced in offices

As in previous studies of office bacterial communities (8), we observe that Proteobacteria, Firmicutes, and Actinobacteria are three dominant phyla across all locations in all offices and all cities. To investigate the extent that offices have city-specific bacterial communities (i.e., offices within a given city have bacterial communities that look more like the communities in other offices from the same city than offices from other cities), we developed both support vector

machine (SVM) and Random Forests machine learning classifiers in attempt to classify the city that a microbiome sample was derived from, as described in (18). Though the two approaches gave similar results, the SVM classifiers performed slightly better based on F-1 scores (a weighted average of precision and recall, ranging between 0 and 1).

Using SVM, we were able to predict the city of origin of unlabeled samples (where we knew the city of origin, but withheld that information from the classifier) with 85% accuracy, only based on its microbiome composition (Figure 4; F-1 score 0.85). The SVM model correctly classified the city performed 2.67 times as well as random guessing (where the most common city was always chosen).

A potential confounding effect related to the city-specific microbial communities was the effect of office-specific bacterial communities. Because each city contained three offices, it is possible that our classifier actually predicts individual offices, because each office is only found in one city. If the bacteria found in each office were primarily derived from the inhabitants of the office, who are known to have highly personalized microbiomes (19, 20), we would expect to observe office-specific effects (8). To determine if individual offices were the source of the city-specific bacterial communities that we observed, we trained classifiers to predict the office from which a sample was taken. For all nine offices, our best classifiers achieved an F-1 score of 0.28, and a mean accuracy 2.9 times that achieved by random guessing. Evaluation of the confusion matrix (Figure 4b) shows that the majority of misclassifications happen when the classifier assigns an incorrect office in the correct city, further suggesting that offices within cities look more similar to each other than offices in different cities. Our design of including multiple offices in multiple cities therefore allows us to separate city-specific bacterial communities effects from office-specific bacterial community effects, and our findings suggest that communities differ across cities, but not necessarily across offices in those cities.

If we train the same classifiers only on offices from within individual cities, the classifiers could predict the office of origin more accurately, but still less accurately than predicting the city of origin (Flagstaff offices only:F-1=0.41; San Diego offices only:4-1=0.37; Toronto offices only:F-1=0.53; Figure 4c-e). This is especially interesting because, even within a city, offices were different from each other, for example in terms of size, usage patterns, and ventilation systems, suggesting that geography is more important than any of these features in driving the bacterial community composition of the offices.

In addition to the community composition differences that are shown by SVM, Flagstaff offices had richer communities (as determined by PD) than San Diego or Toronto (two-sample monte carlo t-test with 999 permutations, test-statistic = 7.029, p-value = 3.411e-12 and test-statistic = 9.156, p-value = 0.0, respectively), which were more similar to one another in community richness (two-sample monte carlo t-test with 999 permutations, test-statistic = 2.352, p-value = 0.019) (Figure 1). The specific taxa that are most different across cities are presented in Supplementary Dataset 3. Together these differences enable the classifiers to differentiate cities.

Office bacterial communities are moderately influenced by indirect contact with humans

In addition to our office surface samples, we collected human skin, nasal, oral, and fecal microbiome samples from 11 inhabitants of one of our Flagstaff offices, and from the individuals performing the sampling in all three cities. Using SourceTracker 2 (21, 22), we defined the human microbiome samples as potential “sources”, and the office surface samples as “sinks”. This allows us to determine which human subjects, and which body sites from those subjects, might serve as sources for the microbes found on the office surfaces.

Across all nine offices, human skin bacterial communities were the largest identifiable source of the office bacterial community samples, with an estimated 25-30% of the office surface microbiome being derived from human skin (Figure 7A). The human nasal microbiome also appeared to be a small but consistent source of office surface microbial communities. The largest source of microbial communities in these offices was from non-human sources (the *unknown* proportion in Figure 7A).

We next defined the source samples as skin from the individuals who collected the samples in each city and 11 inhabitants of Flagstaff office 1 (some of whom only worked in this office for part of the year). Our goal was to determine whether the personalized skin microbiomes (20) of the individuals working in an office were drivers of the office bacterial communities, or whether the office bacterial communities looked more generically like human skin (Figure 7B). The office surface bacterial communities from office 1 in Flagstaff do not appear to be more derived from the inhabitants of that office than the other offices in the study. Similarly, the offices do not appear to more influenced by the individual who sampled in that city (who wore gloves during sampling), relative to the individuals sampling in other cities, or the Flagstaff 1 office inhabitants. Thus it appears that the personalized microbiomes of the office inhabitants or samplers was not transferred to our office surfaces.

Other work has shown that the personalized human microbiome is transferred from human subjects to their offices through direct contact (13, 14). In our study we specifically requested that office inhabitants did not touch the materials, and we required our samplers to wear gloves while collecting samples. We therefore suspect that our inability to detect a personalized human microbiome signal in our office microbiome data is a result of our sampling swatches not being in direct contact with the office inhabitants, though the surfaces still look

more generically like human skin suggesting a role for indirect transfer of skin microbes to office surfaces.

Sequencing run can be an important confounding factor in long-term temporal studies of bacterial communities

Our bacterial samples were sequenced in three sequencing runs, to facilitate the sequencing of the large number of samples collected in this study and to provide us with a way to begin analysing data before all samples were collected. We were specifically interested in using the samples collected early in the study to inform decisions that would be made during later sampling periods, such as how frequently we should collect samples to track changes in the office bacterial communities. As the cost of sequencing continues to decline, we expect that this model will become increasingly frequent. An issue with this approach, however, is that it conflates time with sequencing run, potentially introducing a batch effect. Samples collected during the summer sampling period were sequenced on our first sequencing run, samples collected during the fall and winter periods were sequenced on our second run, and samples collected during the spring period were sequenced on our third sequencing run (Figure 5a-b).

To detect and quantify the run effect, the same eight samples (our technical replicates) were sequenced on three runs: the run of summer samples (Run 1), the run of spring samples (Run 3), and an additional partial run used to understand inter-run variation (Run 4).

Theoretically, these eight samples should have been identical in both composition and richness across the three runs. While we expected differences between sequencing runs, the observed variation between runs was larger than we expected. The technical replicates sequenced in run 3 (spring samples) in particular were very different in composition (using unweighted UniFrac) and richness (PD) from runs 1 and 4, but all of these runs were significantly different from one another based on these metrics (Figure 5D). Using weighted UniFrac, none of the runs were

significantly different from the others. The larger differences using unweighted UniFrac than weighted UniFrac suggest that the run effect we observe is primarily due to differences in low abundance taxa, and not shifts in the dominant taxa between runs.

We attempted several approaches for removing this run effect. First, we ensured that the the sequence length was the same across all runs by trimming longer reads (251 bases) to the same length as our shortest reads (151 bases). All data presented in this study is based on these length-normalized reads. We additionally tried various filtering strategies, including filtering low abundance OTUs, and filtering the n OTUs that were the most differentially abundant across the technical replicates (as identified by ANCOM (23)), where n was varied between 0 and 1000. While the variation across runs was reduced by these strategies, we did not observe significant differences in the results until we filtered out over 1000 of the differentially abundant OTUs, and even at this level the run effect was only minimally reduced.

Despite the differences in replicates across runs, there is evidence of seasonal variation in the bacterial communities. Figure 5C illustrates differences in the bacterial community composition of two samples each of the same 10 sampling swatches, in the summer and fall. These samples were sequenced on the same run, so there is no confounding run effect. In weighted and unweighted UniFrac principal coordinates analysis (PCoA), the samples collected during the same season cluster with one another, suggesting that these are more similar to each other than samples collected from the same swatches but in different seasons. Similarly, the fall samples had higher community richness than the summer samples when sequenced on the same run (Figure 5). Taken together, Figures 5c-d suggest that the observed seasonal differences are likely representative of the underlying biology though conflated with a run effect. We note that we are only able to determine this because our experimental design involved

repeated sequencing of samples. This is essential for studies that require splitting sequencing across multiple runs.

Community richness of office surfaces is consistent over time, and impacted by sampling

Our summer sampling period, which took place immediately after UV sterilization of the plates, showed that bacterial communities were less rich than the fall or winter sampling periods that followed (Figure 8). The spring sampling period subsequently appears to be less rich on average than the summer, fall or winter. While these data suggest that community richness increases following installation of plates, plateaus, and begins to decrease, we note that the summer, fall/winter, and spring samples were sequenced on sequencing runs 1, 2 and 3, respectively, and that these sequencing runs differed in their mean richness (Figure 5d), so as described above this is likely a combination of real biological effects and run effects.

Figure 8 additionally illustrates the diverging of the richness of the floor samples and the wall/ceiling samples, with the distributions becoming increasingly more bimodal with time. The upper mode in these distributions is primarily composed of floor samples, where the lower mode is composed of wall/ceiling samples.

Community compositions were not found to be associated with any indoor or material environmental covariates

Throughout the study, indoor environment as well as surface microenvironment (i.e., sensors specifically collecting data from our sampling swatches) data were collected, including surface and air temperatures, equilibrium relative humidity, relative humidity, room occupancy and visible light illumination. Despite the extent of this environmental data, and considerable effort to identify associations between microbial composition and these data, we failed to identify

significant associations between building science and microbiome data. For example, there were no meaningful correlations between levels of community richness, composition, or abundance of specific taxa of interest and equilibrium relative humidity of the material (the humidity at which moisture is no longer being absorbed by or evaporated from the material). The single exception to this was a weak but significant positive correlation between fungal community richness and equilibrium relative humidity ($r: 0.315$, $p\text{-value}: 3.24e-12$, Figure 6). A significant correlation was also observed between bacterial community richness and equilibrium relative humidity, but this was weaker and negative, so we suspect that it is a false positive. We provide detail on additional tests we performed to test for associations between the office microbiome and environmental parameters in *Supplementary Results*, though none of these tests yielded significant associations.

Discussion

Our experimental design allowed us to isolate variables that were confounded in previous studies such as surface material, human usage patterns, and location in an office (15, 20, 24). It is important to note that the goals of those studies did not necessarily require separating these variables, so our findings are not in disagreement with theirs. Our ability to better explore these parameters in isolation has led to a better understanding of the microbiology of the BE, and enables us to make recommendations that can improve future studies.

Our data suggest that the personalized office microbiomes that have previously been reported are more likely to be the result of direct human contact with the surfaces being sampled, rather than the result of climatic or other differences in the offices themselves. Many of the studies that report that office microbial communities are personalized (i.e., that individual offices have signature microbial communities) have sampled materials that were

already present in the offices in areas of high human traffic (9, 10, 13), rather than from materials installed for the specific purpose of microbial tracking. Our sampling materials were specifically placed in low traffic areas and contained signs requesting that individuals not touch the materials. While we did observe an office-specific microbial community effect, the effect size was smaller than what has been previously reported, and smaller than the city-specific bacterial community effect that we observed (Figure 4b). This suggests that “personalized office microbiomes” are likely largely a result of the “personalized human microbiomes” (19) of the office inhabitants, based on the microbes that they leave behind on the surfaces through direct interaction. The source of the city-specific BE microbial communities that we observe will be important to understand better, as it could for example result in city-specific “cage effects” in murine microbiome studies. We observe that around 25 to 30% of the office surface microbiomes are human-derived (Figure 7), primarily from human skin, suggesting that indirect contact does impact office microbiomes (e.g., through the the office inhabitants’ “personal microbial clouds” (25)), but this is far less than what would be expected on surfaces that the inhabitants are in direct contact with (14, 20).

Our nested experimental design has allowed us to draw several conclusions that should impact the design of future BE microbiome studies. First, sampling different locations (such as the floor and ceiling in multiple locations) in a built environment is likely to result in observing greater variability among microbial communities than sampling different surface materials (e.g., carpet versus tile flooring). Thus limited sampling effort is likely better expended on sampling different locations rather than different materials. Next, it is likely more useful to sample offices in buildings in different climates than to sample multiple offices in the same city, as there seem to be consistent differences in the compositions of the offices we studied by city. Finally, sampling of bacterial communities using dry swabs (as were used here) is sufficient for

obtaining consistent bacterial community 16S rRNA profiles using Illumina sequencing.

Sampling of these communities does “disturb” them, but the effect size of those disturbances will likely be small relative to the biological effects of interest. Our approach of dry swabbing did not work as well for sampling of fungal communities (we received very low PCR yield, see *Supplementary Results*), though it has been successfully applied in previous work that used the same swabs (26). We expect this differential success is a result of differences in biomass (public restroom floor tiles were previously sampled, swabbing a larger area with much heavier human traffic). Additional work should be performed to understand how collection of fungal samples can be performed on low biomass BE samples.

Despite considerable effort we did not detect any difference in microbial communities associated with material or indoor environment covariates such as temperature or equilibrium relative humidity. While this doesn't prove that these parameters do not impact microbial communities, it does suggest that the variation across indoor environmental conditions, which are restricted to a narrow range for comfort of the inhabitants, may not be enough to drive changes in the microbial communities. Another compatible and viable hypothesis is that rather than observing succession of microbial communities over the course of the year, we instead have observed passive accumulation of biologically inactive microbes. This is consistent with the relatively small amount of change in these communities over the year, as illustrated in Figure 8. BE surface microbial communities may behave similarly to those found in the soils of the Atacama Desert (27), waiting for liquid water to become active. Our findings suggest that detailed monitoring of material and indoor environment covariates is not necessary in studies of the composition of microbiology of the BE, except perhaps when the parameters are likely to be microbially relevant (such as the addition of liquid water through real or simulated flooding), or

fall far outside the normal range. This information may however be important for understanding bacterial and/or fungal load on BE surfaces (26, 28).

Our experimental design included incremental sampling over the period of one year. Due to the large number of samples that were collected, and to allow us to analyse data as it accumulated to make decisions about future sampling, it was not possible to include all of our samples on a single Illumina MiSeq run. As we illustrate in Figure 5, this has resulted in a sequencing run artifact in our data. As microbiome sequencing becomes more routine, for example as we transition toward human microbiome monitoring in clinical settings to discover early signals of dysbioses, this type of design is likely to become more common and sequencing run effects will need to be understood, controlled for, and ideally eradicated. We note that in this study we are working with low biomass samples and relatively small biological effect sizes, so the possibility of a run effect interfering with biological effects of interest is likely amplified. However, our approach for detecting and understanding this run effect was useful here, and we recommend that this be routinely adopted. Specifically, we suggest that in studies necessitating multiple sequencing runs that researchers include either some of their samples (as we did here) or control samples (such as artificial communities of known composition) in all sequencing runs as technical replicates. An accumulation of publicly available sequence data replicated across sequencing runs will help us to understand, and possibly develop methods to control for, sequencing run effects.

This work expands our basic understanding of the factors impacting the microbiology of built environments. The human skin microbiome has a considerable impact on the composition of microbiomes on built environment surfaces, even when humans are not in direct contact with those surfaces. Further, within the range of human comfort, differences in the indoor environment do not appear to impact microbial community composition. These findings provide

insight into what drives the composition of BE microbiomes, and taken together we suspect that in the absence of extreme conditions (for example, flooding), microbes may be passively accumulating on surfaces rather than undergoing an active succession process. We additionally show that features previously suspected to be important in driving microbial composition on BE surfaces, such as the surface material, do not seem to impact community composition under typical circumstances. As we continue to expand our understanding of the microbiology of the BE, possibly including routine monitoring of microbial communities as indicators of changes toward communities that may impact human health, the results presented here will help with making critical decisions about important dependent and independent variables in future research efforts.

Materials and Methods

City and office selection

To develop our understanding of how microbes establish in built environments over time and a range of building parameters, we sampled nine offices in three cities over the period of one year. We selected three cities that were climatically different from one another, as determined by their Köppen climate classifications: San Diego, USA, which has an arid Mediterranean climate (Köppen: Bsh or Csa/Csb); Flagstaff, USA, which has a semi-arid climate (Köppen: Dsb/Csb); and Toronto, CA, which has a humid continental climate (Köppen: Dfa/Dfb).

In each of these three cities we selected three offices, with the goal of making these offices as similar as possible to one another. Offices were always shared spaces, meaning they had more occupancy than just a single individual, but had controlled access. All offices had similar patterns of usage, in that they tend to have the highest occupancy on weekdays during business hours. We selected offices that had consistent ventilation rates, and were

approximately the same size. Finally, we excluded older buildings that contained laboratory space. Supplementary Table 1 summarizes parameters of these offices.

Sampling Plate Construction and Installation

In each office we installed three sampling plates, one on the floor, one on the ceiling, and one on the wall, as illustrated in Figure 1 (which contains a photograph showing the three installed sampling plates in a single office). The plates were constructed from 600 mm x 600 mm x 6 mm sheets of birch plywood. Swatches of each sampling material (painted drywall, ceiling tile, and carpet tile) were mounted to the back of the plate and exposed through holes in the plates. Each plate contained a minimum of two swatches of each material that were sampled for microbial community composition. One of each of these swatches was sampled frequently (as often as every two days during our sampling periods) and the other was sampled infrequently (every three weeks during our sampling periods) to evaluate the impact of sampling on the microbial communities. Our wall-mounted plates (i.e., those not on the floor or ceiling) contained an additional three swatches of each material that were monitored for equilibrium relative humidity on the surfaces. These swatches were not sampled for microbial composition as the equilibrium relative humidity sensor covered most of the available space. The relative locations of each material was consistent across all plates.

Environmental monitoring

Each plate contained an Onset HOBO[®] UX90-005 occupancy sensor and an Onset HOBO[®] U12-012 temperature, relative humidity and light sensor. The occupancy sensors monitored occupancy by sensing movement up to approximately five meters away from the plate. The wall plates additionally contained a Decagon EM-50 data logger with three VP3 sensors that measured equilibrium relative humidity and near-surface temperatures (one on each material).

In one office in each city, an additional two VP3 sensors were used to measure the same conditions on the drywall samples on the floor and ceiling. All measurements were recorded every five minutes except for the occupancy sensors which recorded the time of each change of state (e.g., going from occupied to unoccupied). Data were downloaded from the data loggers once a month using the HOBOWare 3.5.0 and the Decagon ECH₂O Utility software packages.

Plate Sterilization

The plates were constructed in Toronto and then shipped to San Diego State University to be sterilized in order to maintain consistent sterilization techniques. We hoped to get all of the surfaces as close to the same starting point as possible to provide a baseline to compare subsequent samples. Any DNA present in the first post-sterilization sampling event could be controlled for as DNA present prior to the experiment start. Ultraviolet (UV) sterilization was selected because of its ability to denature DNA via cross-linking nucleotides and creating thymine dimers, while not damaging the surface material. The sampling portion of the plate containing the 6 or 9 office materials was covered with Saran™ Wrap and sealed around the edges with packing tape. The back of the plate was also sealed with Saran™ Wrap. Both sides of the plate were sterilized under UV light for 10 minutes. This sterilization procedure was repeated for all 27 plates. Plates were swabbed before sealing, and again (at time zero) when the Saran™ Wrap was first removed at installation time in each of the nine offices, to determine whether UV sterilization worked as expected.

Swabbing procedure

The plates were shipped sealed in Saran™ Wrap as described above. Immediately after unwrapping the plates, each surface was swabbed. This was done to either ensure that surfaces were sterile, and if not sterile to establish the starting community of each surface. Once

the plates were installed the regular sampling schedule was initiated. The sampling was done with BD CultureSwab™ sterile swabs. Each swab tube was labeled with a unique sample ID and corresponding barcode. The swab was removed from the tube immediately prior to swabbing. The cotton swab was swiped left to right across the surface moving downward (or toward the individual who was performing the swabbing) for approximately 3 seconds. The swab was turned 180 degrees and the process was repeated starting from the bottom and moving upwards. Once the swabbing was complete, the swab was returned to the sterile tube and the next swab was removed to swab the next surface. Once all of the surfaces were swabbed, the tubes were placed in a -20°C freezer for storage.

Sampling

Sampling was performed in four, six-week periods, one period per season (summer, fall, winter and spring). The first day of sampling in San Diego was May 29, 2013; in Flagstaff, June 27, 2013; and Toronto July 3, 2013. Samples were taken every other day in the first sampling period, and every Monday, Wednesday and Friday in the subsequent periods to simplify collection. The last three sampling periods took place simultaneously in all cities (this was not possible for the first sampling period, as we had one team member travel to each city to ensure consistent experimental setup, and our first sampling period had to begin immediately after removal of Saran™ Wrap). Period two took place from 9/23/2013-11/4/2013, period three from 1/6/2014-2/17/2014, and period four from 4/7/2014-5/19/2014.

Sequencing and Data Analysis

Samples were collected in four six-week sampling periods, one per season the samples were labeled based on the cual-id labeling protocol(29). 16S rRNA gene sequencing was used to

profile bacterial communities. Human bacterial microbiome samples were collected and processed through collaboration with American Gut (30).

The V4 hypervariable region of the 16S rRNA gene was amplified using barcoded PCR with 515F (GTGCCAGCMGCCGCGGTAA) and 806R (GGACTACHVGGGTWTCTAAT) primers following the Earth Microbiome Project protocol (31). All sequencing was performed at Argonne National Laboratories on an Illumina MiSeq. Raw fastq files containing sequence data for both the 3' and 5' reads and the barcodes were provided by the sequencing facility via secure file transfer protocol (SFTP). The bacterial raw sequence data contained 41,335,672 DNA reads for the three runs. The sequence length for the first two sequencing runs was 251 bases, and 151 bases for the third sequencing run. All reads were trimmed to exactly 151 bases before analysis so they could be directly compared.

All bioinformatics analysis was performed on the 5' reads. The reads were demultiplexed and assigned to sample ids using QIIME 1.9.1 (32), and quality filtering was performed as described in (33). After demultiplexing and quality filtering, 33,799,179 reads remained.

The sequences were assigned to operational taxonomic units (OTUs) using QIIME's uclust-based (34), open-reference OTU picking protocol (35) with the Greengenes 13_8 reference sequence set (36) at 97% similarity. The centroid of each OTU was chosen as the representative sequence for the OTU. OTU representative sequences were aligned with PyNAST (37) and phylogenetic trees were constructed with FastTree (38) for phylogenetic diversity calculations. Taxonomy was assigned to sequences using QIIME's uclust consensus taxonomy assigner (39). The resulting bacterial OTU table contained 2309 samples with a median of 7849 sequences per sample.

Beta diversity calculations were performed using QIIME's implementations of weighted and unweighted UniFrac (16) with exactly 1000 randomly selected sequences per sample (40).

Samples with less than 1000 sequences were not included in the calculations. Community richness was calculated using Phylogenetic Diversity (17) and compared across categories using a nonparametric t-test with 999 permutations. Comparisons of significantly different OTUs across sample groups were performed with scikit-bio (scikit-bio.org) using the ANCOM (23). For all permutation based tests, the nested structure of the experimental design was respected so that after the label shuffling, all further nested observations continued to be grouped together.

The SVM analysis to predict city based on sample OTU features was performed using a linear kernel by randomly splitting into two halves (training and test sets), tuning the SVM using 5-fold cross-validation on the training set and then observing the prediction error rates on the test set. SVM and Random Forest analyses were run using the implementations of these methods from scikit-learn (41).

Differentially abundant OTUs across plate locations, frequently/infrequently rows, and cities were identified with ANCOM (23) and provided as Supplementary Datasets 1, 2, and 3, respectively.

All analyses were run using Jupyter Notebooks (jupyter.org). The notebooks used in this project are all available at <https://github.com/gregcaporaso/office-microbes>.

Acknowledgements

We would like to thank Bo Stevens, Alex Embusch, Pascal Reyes, Sandra Dedesko, and Dylan Mctavish for assistance with office microbiome sampling, and Daniel McDonald, Gail Ackermann, and the American Gut Consortium for assistance with human microbiome sampling and sequencing. Human subjects research was performed in accordance with the University of Colorado's Institutional Review Board protocol #12-0582 and the University of California, San Diego's Human Research Protection Program protocol #141853.

Funding information

This work was funded by a grant from the Alfred P. Sloan Foundation to JGC, RK, SK and JS.

References

1. **EPA**. 1989. Report to Congress on indoor air quality: Volume 2 **EPA/400/1-89/001C**.
2. **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.
3. **Husman T**. 1996. Health effects of indoor-air microorganisms. *Scand J Work Environ Health* **5**–13.
4. **Lee L, Tin S, Kelley ST**. 2007. Culture-independent analysis of bacterial diversity in a child-care facility. *BMC Microbiol* **7**:27.
5. **Kembel SW, Jones E, Kline J, Northcutt D, Stenson J, Womack AM, Bohannan BJM, Brown GZ, Green JL**. 2012. Architectural design influences the diversity and structure of the built environment microbiome. *ISME J* **6**:1469–1479.
6. **Ortega-Calvo JJ, Hernandez-Marine M, Sáiz-Jiménez C**. 1991. Biodeterioration of building materials by cyanobacteria and algae. *Int Biodeterior Biodegradation* **28**:165–185.

7. **Viitanen H, Vinha J, Salminen K, Ojanen T, Peuhkuri R, Paajanen L, Lähdesmäki K.** 2010. Moisture and bio-deterioration risk of building materials and structures. *J Building Phys* **33**:201–224.
8. **Hewitt KM, Gerba CP, Maxwell SL, Kelley ST.** 2012. Office space bacterial abundance and diversity in three metropolitan areas. *PLoS One* **7**:e37849.
9. **Hewitt KM, Mannino FL, Gonzalez A, Chase JH, Caporaso JG, Knight R, Kelley ST.** 2013. Bacterial diversity in two Neonatal Intensive Care Units (NICUs). *PLoS One* **8**:e54703.
10. **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.
11. **Mole B.** Patients leave a microbial mark on hospitals. *Nature News*.
12. **Shogan BD, Smith DP, Packman AI, Kelley ST, Landon EM, Bhangar S, Vora GJ, Jones RM, Keegan K, Stephens B, Ramos T, Kirkup BC Jr, Levin H, Rosenthal M, Foxman B, Chang EB, Siegel J, Cobey S, An G, Alverdy JC, Olsiewski PJ, Martin MO, Marrs R, Hernandez M, Christley S, Morowitz M, Weber S, Gilbert J.** 2013. The Hospital Microbiome Project: Meeting report for the 2nd Hospital Microbiome Project, Chicago, USA, January 15(th), 2013. *Stand Genomic Sci* **8**:571–579.
13. **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.
14. **Lax S, Hampton-Marcell JT, Gibbons SM, Colares GB, Smith D, Eisen JA, Gilbert JA.** 2015. Forensic analysis of the microbiome of phones and shoes. *Microbiome* **3**:21.
 15. **Meadow JF, Altrichter AE, Kembel SW, Moriyama M, O'Connor TK, Womack AM, Brown GZ, Green JL, Bohannan BJM.** 2014. Bacterial communities on classroom surfaces vary with human contact. *Microbiome* **2**:7.
 16. **Lozupone CA, Hamady M, Kelley ST, Knight R.** 2007. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Appl Environ Microbiol* **73**:1576–1585.
 17. **Faith DP.** 1992. Conservation evaluation and phylogenetic diversity. *Biol Conserv* **61**:1–10.
 18. **Knights D, Costello EK, Knight R.** 2011. Supervised classification of human microbiota. *FEMS Microbiol Rev* **35**:343–359.
 19. **Califf K, Gonzalez A, Knight R, Caporaso JG.** 2014. The human microbiome: getting personal. *Microbe Wash DC* **9**:410–415.
 20. **Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R.** 2010. Forensic identification using skin bacterial communities. *Proc Natl Acad Sci U S A* **107**:6477–6481.
 21. **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.
22. SourceTracker2. Biota Technology.
23. **Mandal S, Van Treuren W, White RA, Eggesbø M, Knight R, Peddada SD.** 2015. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis* **26**:27663.
24. **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.
25. **Meadow JF, Altrichter AE, Bateman AC, Stenson J, Brown GZ, Green JL, Bohannan BJM.** 2015. Humans differ in their personal microbial cloud. *PeerJ* **3**:e1258.
26. **Fouquier J, Schwartz T, Kelley ST.** 2015. Rapid assemblage of diverse environmental fungal communities on public restroom floors. *Indoor Air*.
27. **Neilson JW, Quade J, Ortiz M, Nelson WM, Legatzki A, Tian F, LaComb M, Betancourt JL, Wing RA, Soderlund CA, Maier RM.** 2012. Life at the hyperarid margin: novel bacterial diversity in arid soils of the Atacama Desert, Chile. *Extremophiles* **16**:553–566.
28. **Liu CM, Aziz M, Kachur S, Hsueh P-R, Huang Y-T, Keim P, Price LB.** 2012. BactQuant: an enhanced broad-coverage bacterial quantitative real-time PCR assay. *BMC Microbiol* **12**:56.

29. **Chase JH, Bolyen E, Rideout JR, Caporaso JG.** 2016. qual-id: Globally Unique, Correctable, and Human-Friendly Sample Identifiers for Comparative Omics Studies. *mSystems* **1**:e00010–15.
30. **Knight R.** 2014. American Gut.
31. **Caporaso JG1, 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.
32. **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.
33. **Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA, Caporaso JG.** 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* **10**:57–59.
34. **Edgar RC.** 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**:2460–2461.
35. **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 H-W, Knight R, Caporaso JG.** 2014. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* **2**:e545.
36. **McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P.** 2011. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* **6**:610–618.
37. **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.
38. **Price MN, Dehal PS, Arkin AP.** 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* **5**:e9490.
39. **Bokulich NA, Rideout JR, Kopylova E, Bolyen E, Patnode J, Ellett Z, McDonald D, Wolfe B, Maurice CF, Dutton RJ, Turnbaugh PJ, Knight R, Caporaso JG.** 2015. A standardized, extensible framework for optimizing classification improves marker-gene taxonomic assignments. e1502. *PeerJ PrePrints*.
40. **Lozupone C, Knight R.** 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**:8228–8235.
41. **Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher**

M, Perrot M, Duchesnay É. 2011. Scikit-learn: Machine Learning in Python. *J Mach Learn Res* **12**:2825–2830.

Figure captions

Figure 1: Experimental design. (a) Configuration of sampling site in Flagstaff office 1. This configuration was similar to those set up in all offices. Signs on the wall adjacent to wall sampling plate describe the project, as request that the materials not be touched. (b) Diagram of single sampling plate illustrating nine sampling swatches (circles) of three different materials, one row for tracking equilibrium relative humidity of the materials (Row #1), one row for infrequent sampling (Row #2), and one row for frequent sampling (Row #3). (c) Samples were collected from rows 2 and 3 of all sampling plates from three offices in each of our three cities in four intensive sampling periods over the course of one year. Coloring of sampling swatches in this figure illustrates the change in bacterial Phylogenetic Diversity over the year.

Figure 2: Bacterial community dissimilarity as measured by (a) weighted UniFrac, a quantitative measure, and (b) unweighted UniFrac, a qualitative measure for the Flagstaff samples (similar patterns are observed across all cities). Darker colors indicate that groups of samples are more dissimilar and lighter colors indicate that groups of samples are more similar. The labels indicate the location in the office followed by the material type. For example, *Ceiling/Carpet* indicates the samples of carpet that are installed on the ceiling sampling plate, and investigation of the first column of (a) shows that the carpet samples on the ceiling are more similar to the carpet samples on the wall than they are to the carpet samples on the floor.

Figure 3: Disturbance due to repeat sampling, though detectable, is small compared to other variables. (a) Phylogenetic diversity (PD) of frequently sampled swatches (“Row 2”, see Figure 1b) was consistently lower than PD of the frequently sampled swatches, suggesting that sampling decreases the PD of the sites. (b) Weighted UniFrac shows that the samples with the same sampling frequency are more similar to each other than samples with different sampling frequencies, suggesting an effect of repeat sampling. However comparison of this difference to effects of biological interest show that sampling frequency has a larger effect on the bacterial communities than the material does (which we show in Figure 2 is effectively null), but a small effect than the location of the plate in the office or the sequencing run, which we show in Figures 2 and 5 to impact our observed bacterial communities. (c) Similar results are observed with unweighted UniFrac.

Figure 4: Confusion matrices illustrating the performance of city and office SVM classifiers. (a) True (actual) city and predicted city for our SVM classifier when trained and tested on office microbiomes by city. Dark colors on the diagonal indicate the predicted label is very frequently correct. (b) True and predicted office for our SVM classifier when trained and tested on office microbiomes from all cities labeled by individual office. Note that the diagonal is

not as dark as in (a), illustrating that the predicted office labels are not correct as often as the predicted city labels. When an incorrect office is predicted, it is often from the correct city, as indicated by the darker colors surrounding the diagonal. (c-e) True and predicted office for our SVM classifiers when trained and tested on office microbiomes from cities individually.

Figure 5: Investigation of sequencing run effect on observed bacterial community composition. (a) Weighted and unweighted UniFrac principal coordinates analysis (PCoA) and Phylogenetic Diversity (PD) by sequencing run for sequencing runs 1, 2 and 3. (b) Weighted and unweighted UniFrac PCoA and PD by season. (c) Weighted and unweighted UniFrac PCoA and PD for samples of the same sites from summer and fall when sequenced in a single sequencing run. (d) Weighted and unweighted UniFrac PCoA and PD of technical replicate samples sequenced on sequencing runs 1, 3, and 4.

Figure 6: Community richness as a function of equilibrium relative humidity of sites for bacterial (16S) and fungal (ITS) communities. We observe weak correlations between community richness and equilibrium relative humidity (ERH) of sites. While these correlations are statistically significant, we do not find these relationships to be convincing for reasons discussed in *Results* and *Supplementary Results*.

Figure 7: Human-based source tracking of office microbiome samples. (a) Percent contribution of microbiomes of different sites of the human body to the office bacterial communities. *Unknown* indicates contribution from a source other than the body sites tested here. (b) Percent contribution of microbiomes of different human subjects to the office bacterial communities. *Unknown* indicates contribution from a source other than the individuals tested here.

Figure 8: Longitudinal analysis of bacterial Phylogenetic Diversity (PD) over one year in (a) Flagstaff, (b) San Diego, and (c) Toronto. Each “violin” represents a two week period from the beginning, middle, and end of our four sampling periods.