The ecology of microscopic life in household dust

Albert Barberán1,†, Robert R. Dunn4,7,†, Brian J. Reich5, Krishna Pacifici6, Eric B. Laber3, Holly L. Menninger4, James M. Morton1, Jessica B. Henley1, Jonathan W. Leff3,2, Shelly L. Miller3 and Noah Fierer1,2

1Cooperative Institute for Research in Environmental Sciences, 2Department of Ecology and Evolutionary Biology, and 3Department of Mechanical Engineering, University of Colorado, Boulder, CO 80309, USA
4Department of Biological Sciences, 5Department of Statistics, and 6Department of Applied Ecology, North Carolina State University, Raleigh, NC 27695, USA
7Museum of Natural History, University of Copenhagen, Copenhagen, Denmark

We spend the majority of our lives indoors; yet, we currently lack a comprehensive understanding of how the microbial communities found in homes vary across broad geographical regions and what factors are most important in shaping the types of microorganisms found inside homes. Here, we investigated the fungal and bacterial communities found in settled dust collected from inside and outside approximately 1200 homes located across the continental US, homes that represent a broad range of home designs and span many climatic zones. Indoor and outdoor dust samples harboured distinct microbial communities, but these differences were larger for bacteria than for fungi with most indoor fungi originating outside the home. Indoor fungal communities and the distribution of potential allergens varied predictably across climate and geographical regions; where you live determines what fungi live with you inside your home. By contrast, bacterial communities in indoor dust were more strongly influenced by the number and types of occupants living in the homes. In particular, the female: male ratio and whether a house had pets had a significant influence on the types of bacteria found inside our homes highlighting that who you live with determines what bacteria are found inside your home.

1. Introduction

If there was ever any doubt, it has become clear in recent years that we are exposed to many thousands of bacterial and fungal species as we go about our daily lives [1–4]. As humans become ever-more urban and we now spend the majority of our lives indoors [5], the time we spend with the microbial taxa found inside homes is increasing. While some of these taxa have negative effects on human health, either as pathogens (e.g. antibiotic-resistant Staphylococcus aureus [6]) or by serving as triggers of allergies and allergic asthma [7–10], others may even be beneficial [11–14]. More generally, it has been argued that childhood exposure to reduced levels of microbial diversity in and around homes may partially explain the rise in the incidence of allergies and autoimmune disorders in many developed countries [13,15]. Clearly understanding the factors that influence which microbial taxa can be found within a particular home has great significance for human health and well-being.

Previous studies suggest that the composition of the microbes found inside homes can be strongly influenced by the presence, identity and activities of human occupants [16–20]. Non-human occupants, including dogs [3,21,22] and household insects [23], can also influence the types of microbes found inside homes. Likewise, there is evidence that the types of bacteria and fungi found inside homes can be affected by differences in ventilation, building design, the environmental characteristics found within buildings [20,22,24–28] or prior water damage from flooding [29]. If we consider a home to be a microbial ecosystem, the geographical location of a home should also be important in...
structuring the types of bacteria or fungi found inside the home, just as geography has been known for many decades to structure plant and animal communities across and within continents [30].

Geographical structure in the composition of microbe species in homes could arise from dispersal constraints—not all microbial taxa that can enter homes are evenly distributed across broad geographical regions, or they could arise from environmental factors (e.g. temperature, humidity) that select for certain taxa in certain regions. We expect that indoor fungi will exhibit particularly strong variability across geographical regions, given that most fungal taxa found indoors come from outside the home [27] and the fungi found in outdoor air can exhibit a high degree of geographical endemism [31,32]. Yet, the biogeography of indoor microbial life remains largely unstudied with the exception of a single study of 72 buildings, in which strong effects of climate were found on indoor fungal taxa [25].

If the types of bacteria and fungi inside homes track outdoor climatic conditions, then it could be because of the influence of outdoor climate on indoor climate (e.g. houses in Florida are warmer in the winter than those in Alaska [33]). Alternatively, such patterns could arise because of the influence of outdoor microbes on indoor microbes due to dispersal into homes through windows and doors or on occupants [24,34]. Just as pollutants inside homes often reflect pollution levels outside homes [35,36], indoor microbes might be influenced by outdoor conditions either because of frequent flow of outdoor microbes into homes [24,34] or because indoor conditions are incompletely buffered from outdoor conditions [33]. From previous work, we know that the types of bacteria and fungi found in outdoor air can vary as a function of surrounding soil conditions, land-use type, climate and net primary productivity [37], with these factors together yielding predictable regional differences in the overall composition of bacterial and fungal communities found in outdoor air [38]. All of these factors may influence the species found indoors directly, or they may do so via their effects on outdoor taxa. Alternatively since many homes, particularly those in developed countries, have environmental conditions distinct from those found outdoors with limited air exchange [33], the effects of regional conditions might be modest when compared with the effects of house design (including ventilation), operation and occupancy.

We currently lack a comprehensive understanding of how those microbial communities found in homes vary across broad geographical regions and what factors are most important in shaping the types of bacteria and fungi found inside our homes. This knowledge gap persists despite the well-recognized importance of these microbes to human health and well-being. Ideally, one wants to understand what factors influence which microbes live in homes and, in turn, how the composition of those microbes influences health and well-being. Here, we take the first of those two steps; we investigated the geography and local ecology of the fungi and bacteria found inside approximately 1200 homes located across the continental US, homes that represent a broad range of home designs and span many climatic zones. We used high-throughput sequencing coupled with multivariate statistical models to identify how home characteristics, the types of bacteria and fungi found outside the home, the numbers of occupants (both human and non-human), outdoor climate, degree of urbanization and geographical location shape the assemblages of fungi and bacteria found in dust collected from inside homes.

2. Material and methods
(a) Sample collection, molecular analyses and sample characterization

Details of the sample collection and analytical procedures are nearly identical to those described previously [38]. Briefly, indoor and outdoor dust samples were collected by volunteers participating in the Wild Life of Our Homes citizen science project (homes,yourwildlife.org). Participants were instructed to sample the upper door trim on an interior door in the main living area of the home and the upper door trim on the outside surface of an exterior door. These sampling locations were selected as they are unlikely to be cleaned frequently and serve as passive collectors of settled dust inside or outside home with little to no direct contact from the home occupants.

Swabs were prepared for sequencing using a direct PCR approach [39]. For bacteria, we sequenced the V4 hypervariable region of the 16S rRNA gene using the 515-F (GTGCCAGCMGCCGCGGTAA) and 806-R (GGACTACHVGGGTWTCTAAAT) primer pair. For fungi, we sequenced the first internal transcribed spacer (ITS1) region of the rRNA operon using the ITS1-F (CTT GGTATTAGAGGAGTAA) and ITS2 (GCTGCGTTGTTCGATGC) primer pair. The primers included the appropriate Illumina adapters with the reverse primers also having an error-correcting 12-bp barcode unique to each sample to permit multiplexing of samples [40]. Sequences were quality filtered (maxee value of 0.5) and clustered into greater than or equal to 97% similar phylotypes after removing singleton sequences using the UPARSE pipeline [41]. Phytype taxonomy was determined using the Ribosomal Database Project classifier [42] trained on the Greengenes 13.8 16S rRNA database [43] and the May 2014 version of the UNITE ITS database [44]. In order to remove potential amplicon sequencing biases, we first removed samples with less than 10,000 sequences and then we normalized the sequence counts using a cumulative-sum scaling [45]. For bacteria, the number of paired dust samples collected from both indoors and outdoors was 2284 (i.e. 1142 indoor samples and 1142 outdoor samples). For fungi, the number of paired dust samples was 2266 with an equal number of indoor and outdoor dust samples. The average number of sequences per sample was 79 450 and 99 145 for bacteria and fungi, respectively. The total number of phylotypes identified from both the indoor and outdoor samples was 125 066 and 72 284 for bacteria and fungi, respectively.

Geographical coordinates were derived from location information (home address). Coordinates were then used to obtain georeferenced variables for each household. These variables included climatic and soil factors, population density, plant productivity and the presence of livestock (see details in [38]). Home characteristics and occupant (both human and pets) information were obtained through a questionnaire filled out by the participants. The list of house variables included age and size of the house, number of bedrooms, presence of basement, carpeting, ventilation index, number of days with the windows open, use of insecticides and mould products, number of inhabitants, female: male ratio, number of smokers and vegetarians, number of pets (dogs, cats, birds) and number of house plants.

(b) Data analyses
All analyses were carried out in the R environment (www.r-project.org). We calculated pair-wise dissimilarities between communities using the Bray–Curtis dissimilarity metric, with dissimilarity matrices visualized via non-metric multidimensional scaling.
(NMD5). We used analysis of similarities (ANOSIM) based on 1000 permutations to test for differences in microbial community composition between indoor and outdoor samples [46]. ANOSIM's R statistic ranges from 0 (no group separation) to 1 (complete separation). To estimate the explanatory power of environmental and house variables on bacterial or fungal community composition, we used permutational multivariate analysis of variance (PERMANOVA) based on 1000 permutations [47]. In order to identify potential taxa indicative of specific source environments (i.e. soil, oceans, human skin, human/livestock faeces, plants and insects), we used indicator value analyses, which combines abundance and frequency of occurrence [48], as described in [38]. We also used indicator value analyses to detect phylotypes associated with indoors. Only phylotypes with an indicator value greater than 0.3 and \( p < 0.01 \) were considered significantly associated with the indoor environment. We used the R packages vegan (http://vegan.r-forge.r-project.org/), labdsv (http://ecology.msu.montana.edu/labdsv/R/) and ecodist (http://cran.r-project.org/web/packages/ecodist/) for multivariate statistics.

We used machine learning, in particular boosted regression trees (BRTs), to predict the probability of the presence of pets (i.e. dogs and cats) based on the abundance of indoor bacterial phylotypes. BRTs combine regression trees (i.e. models with recursive binary splits) and boosting algorithms (i.e. large number of simple models merged to optimize prediction). In addition, BRTs can fit nonlinear relationships and allow for missing data and outliers [49]. We used cross-validation to select the optimal number of trees with tree complexity set to 10, learning rate to 0.001, and bag fraction to 0.5. Based on the results of cross-validation, the number of trees was 4450 and 3900 for the presence of dogs model and for the presence of cats model, respectively. Models were trained with 70% of the samples and the remaining 30% were used to test the predictive performance. We used the R package gbm (https://github.com/gbm-developers/gbm) for BRTs.

### 3. Results and discussion

#### (a) Comparing indoor to outdoor microbial communities

From the more than 1100 homes for which we had paired dust samples collected from both inside and outside the home, we could directly compare the diversity and composition of the indoor and outdoor microbial communities. Not surprisingly, indoor communities were, on average, distinct from those found outdoors (figure 1). While individual indoor and outdoor dust samples all harboured diverse microbial communities with hundreds of unique bacterial and fungal taxa, the diversity of bacteria and fungi was on average approximately 50% higher for the indoor samples (Mann–Whitney test \( p < 0.001 \) for both; figure 1a). These differences in diversity levels are likely driven by two interrelated phenomena. First, many of the outdoor communities were dominated by a handful of microbial taxa, while the indoor communities had greater Shannon diversity (Mann–Whitney test \( p < 0.001 \) for both; figure 1b). For example, fungi in the Cladosporium and Toxicocladusporium genera were often very abundant in individual outdoor samples (15.8% and 11.7% of ITS sequences, respectively) as were bacteria in the Actinomycetales and Sphingomonadales orders (11.1% and 8.5% of 16S rRNA sequences, respectively). Second, the bacterial and fungal communities found inside homes were more diverse, because they included both the microbes found outside of homes as well as those taxa that are restricted to the inside of homes or likely to have originated from indoor sources.

In addition to the diversity differences, indoor and outdoor samples harboured distinct microbial communities (figure 1c). These differences were larger for bacteria (ANOSIM \( R = 0.34, \ p < 0.001 \)) than for fungi (ANOSIM \( R = 0.29, \ p < 0.001 \)), but in both cases these differences between indoor and outdoor communities could be attributed, in part, to certain taxa that were relatively more abundant and prevalent indoors than outdoors. For fungi, only 1.1% of taxa (763 out of 72 284 phylotypes) were relatively more abundant inside homes than outside homes. These findings are directly in line with previous work demonstrating that most fungi found inside buildings come from outside the home [16,34], and they suggest that there are unlikely to be many sources of fungi inside homes. However, examining the small subset of fungal phylotypes that were significantly more abundant inside homes than outside homes (electronic supplementary material, table S1) provides some insight into the possible sources of fungi within homes. These included well-known household moulds such as Aspergillus, Penicillium, Alternaria and Fusarium [50], and wood-degrading fungi such as Stereum, Trametes, Phlebia and Ganoderma that are likely wood colonizers [51]. We also find that fungi associated with human skin (including Candida and Trichosporon, [52]), and gastronomically relevant fungi (such as Saccharomycetes, Pneumocystis and Agaricus, [53]) were rare, but significantly more abundant inside homes than outside homes (electronic supplementary material, table S1).

As with the fungal communities, the bacterial communities found inside the home were also distinct from those found outside the home. Only 1.6% of the bacterial phylotypes (1994 out of 125 066) were significantly more abundant inside homes than outside homes but many of these taxa were quite frequent and, when encountered, relatively abundant (electronic supplementary material, table S1). Most notably, bacterial taxa that likely originate from the occupants of homes were relatively more abundant inside homes than outdoors. For example, bacteria likely associated with human skin (e.g. Staphylococcus, Streptococcus, Corynebacterium and Propionibacterium [52]), vaginas (e.g. Lactobacillus, Bifidobacterium and Lactococcus [54]) and faeces (e.g. Bacteroides, Faecalibacterium and Ruminococcus [55]) were far more abundant inside homes (electronic supplementary material, table S1). To quantify the relative importance of different environmental sources of bacteria to the indoor and outdoor dust samples, we compared the relative abundances of taxa indicative of different source environments using the approach described previously [38]. These analyses confirm that skin and faeces are more important sources of bacteria found in indoor dust compared to outdoor dust (figure 2). Furthermore, these analyses show that bacteria from non-human occupants, including those taxa commonly associated with insects, including household insects (e.g. Wolbachia, Buchnera, Rickettsia and Bartonella, [56]) were found to be relatively more abundant inside homes (figure 2).

Together these results highlight that we can use the relative abundances of bacterial and fungal taxa found inside homes to identify potential sources of these microbes within our homes, just as analyses of bacteria found in outdoor air can be used to identify how the sources of outdoor bacteria change over time [57] or across different geographical locations [37]. More generally, these results indicate that those bacteria and fungi living with us in our homes are
not identical to those found outdoors and attempts to link microbes to human health outcomes (e.g. [12,13]) cannot rely exclusively on data collected from outside homes.

(b) Strong influence of the outdoor environment on indoor fungi

Given that 65% of fungal phylotypes found indoors were also found outdoors, it is not surprising that the overall composition of the indoor and outdoor fungal communities were highly correlated across the sampled homes (Mantel’s $r = 0.43$, $p < 0.001$). The best predictor of indoor fungal communities is outdoor fungal community composition. Likewise, those factors that best predict the composition of fungi found in outdoor dust, particularly regional climatic conditions [38], are also those factors that best predict the composition of indoor fungal communities (electronic supplementary material, table S2) with environmental variables (particularly mean annual temperature and mean annual precipitation) explaining 14% of the variability in the composition of the fungal communities found in the indoor dust samples (by contrast, only 5% of the variability was explained by house characteristics; electronic supplementary material, table S2). The correlations with climate variables indicate that homes in different geographical regions of the US harbour distinct fungal communities indoors. More generally,
homes located closer in geographical proximity tended to
harbour more similar fungal communities (Mantel’s $r = 0.35$, $p < 0.001$). Broad regional patterns are evident in figure 3a, where we see that homes in the eastern USA have distinct communities from those found in more western regions. These patterns are partially driven by homes in the eastern USA having higher relative abundances of Basidiomycota (particularly taxa in the class Agaricomycetes) while western homes have higher relative abundances of Ascomycota (specifically from the classes Dothideomycetes and Leotiomycetes). Taken together, these results highlight that the fungal taxa found inside homes are largely determined by the geographical location of the home (as already observed at the local scale; [34]), a finding that may explain strong regional differences in sensitivities to known fungal allergens [10]. Potential fungal allergens usually described in the literature [58] were abundant in indoor samples (on average they represented 31% of the total number of fungal sequences; figure 4). Geographical variability in the relative abundances of potential fungal allergens was driven by the individual distributions of the three most abundant allergenic genera (i.e. Alternaria, Aspergillus and Phoma; electronic supplementary material, figure S1), and these genera were concentrated along the regions of the Great Plains and the Great Lakes (due to Alternaria), the Northeast region (due to Aspergillus) and the state of Arizona (due to Phoma; figure 4 and electronic supplementary material, figure S1). We are careful to note that the species in these genera differ in their levels of allergenicity [8,58] and that linking the abundances of these putative fungal allergens found in household dust to rates of allergic diseases is non-trivial but worthy of additional study.

Owing to the strong influence of those fungi coming from outside the house on indoor fungal communities, we re-ran the analyses just focusing on the 1.1% of fungal phylotypes that were consistently more abundant inside homes than outside homes. By focusing on just this subset, we hoped to better understand how conditions within homes may alter those fungi that are preferentially found within homes. The fungi that were more common in homes were influenced by a relatively small number of indoor variables with the presence of a basement, the age of the home and the presence of dogs having weak, but significant, effects on the relative abundances of these fungal taxa (electronic supplementary material, table S2). Since these effects are far weaker than the effects of outdoor variables or geographical locations (electronic supplementary material, table S2), we conclude that the geographic allocation of the home, not the house or occupant characteristics measured here, is the best predictor of what types of fungi will be found in a given home.

(c) Home occupants structure indoor bacterial communities

Just as for fungi, the composition of the bacterial communities found in indoor dust were significantly correlated with the composition of the bacterial communities found in outdoor...
However, there were larger differences between indoor and outdoor bacterial communities with many of the bacterial taxa found indoors likely coming from occupants (as mentioned above). While an important portion of the variation among homes in fungal community composition was associated with geographical location and climatic variables, this was not the case for the bacterial communities (only 6% of the variability was explained by environmental variables; electronic supplementary material, table S2). Likewise, geographical location was not a good predictor of the types of bacterial communities found in a given home and, although geographical distance was significantly correlated with bacterial community dissimilarity, the relationship was far weaker than for fungi (Mantel’s $r = 0.09$, $p < 0.001$). Unlike the maps showing clear geographical gradients in fungal communities inside homes across the USA (figure 3a)—the corresponding maps for bacterial communities showed no clear regional patterns (figure 3b). Homes located far apart, for example one in California and one in North Carolina, do not necessarily harbour more distinct bacterial communities than homes located in the same state.

Our analyses focusing on the 1.6% of bacterial phylotypes that were consistently more abundant inside homes than outside homes suggest that the occupants of the homes play a significant role in determining the types of bacterial communities found indoors (electronic supplementary material, table S2). In particular, whether a home had dogs was the most important predictor of the relative abundances of those bacteria preferentially found within homes (electronic supplementary material, table S2). This effect of dogs on indoor bacterial communities has been noted previously [3,20,22] with some work suggesting that these bacterial differences are associated with human health outcomes.
Figure 5. Differences in the proportion of bacterial indoor genera (Mann–Whitney test $p < 0.01$ after false discovery rate correction) between homes with or without dogs (a), and between homes with or without cats (b). (c) Differences in the proportion of bacterial indoor genera (Spearman’s rank correlation $p < 0.01$ after false discovery rate correction) between homes with more females than males (female : male ratio $> 0.5$) and homes with more males than females (female : male $< 0.5$). (Online version in colour.)

[21]. To the best of our knowledge, this is the first study to demonstrate that cats can also influence the types of bacteria found in household dust or on household surfaces.

We identified 56 and 24 bacterial genera that were significantly more abundant in homes with dogs or cats, respectively, a subset of which are shown in figure 5, including members of the Prevotella, Porphyromonas, Moraxella and Bacteroides genera. These bacterial taxa have previously been shown to be common in the mouths or faeces of dogs and cats [59,60], suggesting that the pet effect on bacterial communities is in part caused by these pets directly shedding these bacterial taxa from their bodies into our home environment. We then used a machine learning approach (BRTs) focusing on indoor ‘indicator’ taxa to determine whether we could use the relative abundances of these bacterial phylotypes to predict which homes have dogs or cats in them (electronic supplementary material, figure S2). When we tested the predictive performance of these models, we found that we could predict, just from the indoor bacterial phylotypes alone, whether a home has a dog or cat in it with 92% and 83% accuracy, respectively, highlighting the predictable influence that our pets have on the bacterial communities found in household dust.

We also found that the number of inhabitants and the female : male ratio of human occupants in the home predicted, albeit weakly, the relative abundances of those bacteria preferentially found within homes (electronic supplementary material, table S2). While the female : male effect was subtle and only 6% of homes had no women living in them (we had no homes with only women), we could still identify those bacterial taxa that are more common in homes with more women versus homes with less women. Interestingly, we found that Corynebacterium and Dermabacter (both skin-associated taxa, [52]) and Roseburia (faecal-associated, [55]) were relatively more abundant in homes with fewer women (figure 5c). This pattern is probably driven by differences between the skin biology (and perhaps to body size and hygiene practices) of men and women. In particular, it has been shown that men shed more bacteria into their surrounding environment than women [61,62], and also that Corynebacterium is relatively more abundant on men’s skin than on women’s skin [63]. In turn, Lactobacillus was relatively more abundant in homes with women than in homes without women (figure 5c). While there are a wide range of potential sources of Lactobacillus, we hypothesize that members of this particular genus are vaginally associated given that this taxon contains important members of the vaginal microbiome [54]. Members of the Lactobacillus genus have been reported to be protective against allergies and asthma [13,14]. Although additional sampling of residences is required to confirm whether these patterns are generally observed across homes, our results do demonstrate that we may be able to use information on the bacteria found in household dust to identify the sex of occupants—a finding that could be useful for forensic applications.

4. Conclusion

We show that the types of bacteria and fungi found in the settled dust collected from inside homes are, to some degree, predictable. Most of the fungi found in homes come from outside the homes with the indoor dust fungal communities varying predictably across climate and geographical regions. By contrast, the composition of the bacterial communities found in indoor dust was not predictable from the house location and was more strongly influenced by the number and types of occupants living in the homes. In particular, whether a house had dogs or cats had a significant influence on the types of bacteria we live within our homes. If you want to change the types of fungi you are exposed to in your home, then it is best to move to a different home (preferably one far away). If you want to change your bacterial exposures, then you just have to change who you live with in your home.

Ethics. Enrolled participants were provided a written informed consent form approved by the North Carolina State University’s Human
References


