HHS Public Access

Author manuscript

Sci Transl Med. Author manuscript; available in PMC 2017 November 29.

Published in final edited form as:

Sci Transl Med. 2017 May 24; 9(391): . doi:10.1126/scitranslmed.aah6500.

Colonization and Succession of Hospital-Associated Microbiota

Simon Lax[1,2], **Daniel Smith**, **Naseer Sangwan**, **Kim Handley**, **Peter Larsen**, **Miles Richardson**, **Stephanie Taylor**, **Emily Landon**, **John Alverdy**, **Jeffrey Siegel**, **Brent Stephens**, **Rob Knight**, and **Jack A Gilbert**[*]

[1]Department of Ecology and Evolution, University of Chicago, 1101 E 57th St., Chicago, IL 60637

[2]Institute for Genomic and Systems Biology, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439

Abstract

The microorganisms that inhabit hospitals may significantly influence patient recovery rates and outcomes (REFs). To develop a community level understating of how microorganisms colonize and move through the hospital environment, we mapped microbial dynamics between hospital surfaces, air and water to patients and staff over the course of one year as a new hospital became operational. Immediately following the introduction of staff and patients, the hospital microbiome became dominated by human skin-associated bacteria. Human skin samples had the lowest microbial diversity, while the greatest diversity was found on surfaces interacting with outdoor environments. The microbiota of patient room surfaces, especially bedrails, consistently resembled the skin microbial community of the current patient, with degree of similarity significantly correlated to higher humidity and lower temperatures. Microbial similarity between staff members showed a significant seasonal trend being greatest in late summer/early fall correlating with increased humidity.

The indoor environment has become our most intimate ecosystem, though its reduced microbial diversity relative to the outside world may be linked to an increased incidence of immunological diseases, asthma, and allergies[1–4]. A strong link has been observed between the microbial communities associated with human skin and those recovered from buildings[5–7]. However, the implication of this interaction on the incidence of hospital-acquired infections as a leading cause of patient death[8–10] has not been investigated. Culture-dependent analysis of HAIs and pathogen genotype tracking allows for the near-real time *post hoc* characterization of potential transmission routes[11]. However, outside of a small number of studies limited to ICUs and neonatal care rooms[12–14] no detailed longitudinal culture-independent analyses of the hospital microbiome have been performed[15].

Here, we present a yearlong microbial survey of the microbiota associated with the patients, staff, and surfaces of the newly constructed Center for Care and Discovery (University of Chicago). Sampling began 2 months prior to the hospital becoming operational on February 23rd 2013, and continued for nearly a year post-opening. We collected 6,523 microbial

Correspondence to gilbertjack@uchicago.edu.

samples from multiple sites (Table S1) in 10 patient care rooms and two nursing stations across two hospital levels. At least 5,000 high quality 16S rRNA V4 amplicons were generated per sample. One patient room on each level was sampled daily while all other environments where sampled weekly. Environmental conditions, including temperature, humidity, illuminance, CO_2 concentrations and infrared doorway beam breaks were continuously monitored[16–17].

A soon as the hospital became operational, the floor and nursing station surfaces saw a significant increase in the relative abundance of the human-skin associated genera *Corynebacterium*, *Staphylococcus*, and *Streptococcus*, and a decrease in *Acinetobacter* and *Pseudomonas*, which dominated pre-opening samples (Figure S1A,B,C; Figure S2). OTU-level Shannon diversity significantly increased in nursing station surfaces that commonly interact with human skin (Figure S1 D), but not in floor samples.

Skin samples were generally the least diverse of all sample types (although they had greater phylogenetic diversity than expected given their low Shannon index), while surfaces most likely to interact with the outdoors, such as air filters, shoes, and floors, were the most diverse (Figure 1A). Samples taken from the phone, chair armrest, countertop, computer mouse, and floor of the nurse stations, as well as the floor and air filters from the patient rooms, all had significantly lower beta diversity compared to samples taken from patient skin, staff noses, faucet handles, and unused latex gloves (Figure 1B). Staff hand microbial communities were significantly more similar to the microbiota of built hospital surfaces than were patient hands, likely as a result of greater staff mobility within the hospital. Supervised learning models could successfully differentiate nose and hand samples taken from staff members and patients (error ratios of 2.5 and 3.9, respectively) and hand samples could even be differentiated using genus-level data (error ratio = 3.3), with the genera *Micrococcus* (staff-associated) and *Prevotella* (patient-associated) having the highest feature importance scores. Pre-opening room and station floor samples had highly similar microbial communities, but were dissimilar to other surfaces; while post-opening floor samples had a greater degree of similarity to all surfaces (Figure 1B).

To determine the strength of microbial interaction between different hospital surfaces, we calculated the degree to which samples taken from two different surfaces on the same day and in the same room or building level, resembled each other using a method we term PC space correlation (Figure 2). Patient hands and bedrails had a strong interaction ($\rho = 0.5$); but all pairwise correlations within patient rooms were significant, suggesting a degree of microbial homogenization on the same day within each room. The strongest observed correlations were between the hand microbiota of hospital staff and their personal cell phones and pagers ($\rho = 0.52$ and 0.50 respectively), which has been observed previously[18–19]. Correlations within the nurse station environment were all significant but comparatively weak, probably due to the diversity of people using these environments daily. Correlations between environments were generally much weaker than those within environments, likely because of the reduced overlap in personnel occupancy. The station and patient floors had the greatest correlations between environments, likely due to homogenization by shoe traffic between locations. However, staff shoe and floor samples did not seem to interact significantly despite previous observations[18]. This may be due to regular

cleaning of the hospital floors or, more likely, the diversity of other individuals walking throughout the hospital who were not sampled in this study.

To investigate the degree to which the microbiota of individual patient rooms varied over the year, we calculated the variability of the core microbiome (Figure 3A). For all surface types, the average percent of 16S rRNA reads detected in every sample of a given room varied between 15 and 35%, being greatest for patient noses and lowest for patient axillae. The trends for all 5 surfaces were remarkably similar, suggesting that variability in the microbiota of room bedrails and floors could be attributable to variation in the skin microbiota of patients. We used a Bayesian source-tracking approach[20] to match microbial profiles of samples taken from the second day of a patients stay, to those taken on their first day (Figure 3B). For the same surface type (e.g. hand to hand), the model was highly predictive for matching the patient on day 2 to the correct patient on day 1. Models using the microbial profile of day 1 hands to predict day 2 bedrails, and day 1 bedrails to predict day 2 hands, were also highly accurate. Floor, nose, and axilla samples, by contrast, had much weaker predictive accuracy, though with the exception of nose or floor predicting axilla, were significantly better than random.

Within a patient room on the same day higher temperatures and higher illuminance was consistently associated with greater microbial dissimilarity between patient and surface microbial communities, while higher relative humidity and humidity ratio consistently correlated with greater microbial similarity (Figure 4A). Skin-associated microbial similarity between staff members working on the same floor showed a seasonal trend, with the greatest similarity in late summer/early fall, and the least similarity in the winter (Figure 4B). The greater the humidity, the greater the nose and hand-associated microbial similarity between staff; while staff hands became less similar with increasing temperatures (Figure 4C). Hand, nose, and floor samples taken on the same week were generally more similar than those taken on different weeks for both patient and staff environments (Figure S3).

The four dominant genera identified in the hospital comprised 303 unique strains (Oligotypes): with 116, *Staphylococcus*, 83 *Streptococcus*, 77 *Corynebacterium*, and 27 *Acinetobacter* strains (Figure S4). When we applied Bayesian source tracking to the two patient rooms that were sampled daily, using the microbial profiles on hand or axilla samples from each patient as a source, it was often possible to accurately predict the correct patient based on microbial profiles on the bedrails, and to a lesser extent on floors and faucet handles (Figure S5; Figure S6). This suggests that individual patients can harbor a unique strain-level skin microbial profile, though the fact that many oligotypes were found in a large number of samples (Figure S4) limits predictive accuracy.

The complex dynamic environment of a hospital is an extreme environment for microbial life. With frequent cleaning and constant turnover of occupants this environment is purpose built to reduce the microbial interaction between occupants. We have demonstrated that the building parameters shape the microbial sharing between occupants, but that these drivers are linked to seasonal humidity trends, driven by local climate.

METHODS

Sample Collection

Samples were collected by trained technicians at the Center for Care and Discovery at the medical center of the University of Chicago in compliance with IRB12-1508. With the exception of air samples, which were collected via filters placed in patient room air vents, all samples were collected by rubbing sterile swabs pre-moistened with 0.15M saline solution on the site of interest. After collection, samples were immediately frozen at −20° pending shipment to Argonne National Laboratory on dry ice. Environmental factors and occupancy proxies were continuously collected as previously reported[16–17].

Amplicon Sequencing

All samples were processed using a modified version of the manufacturer's protocol of the Extract-N-Amp kit (Sigma-Aldrich). Swabbed tips were placed into 2ml 96-well Deep Well plates (Axygen). 200μl of Extract-N-Amp Extraction solution was added, vortexed for 5 seconds, and incubated at 90°C for 10 minutes. Samples were centrifuged a 2,500 × g for 1 minute. 200μl of Extract-N-Amp Dilution solution was added to each sample to obtain a 1:1 ratio of extraction to dilution solution. Genomic DNA was amplified using the Earth Microbiome Project barcoded primer set, adapted for Illumina HiSeq2000 and MiSeq by adding nine extra bases in the adapter region of the forward amplification primer that support paired-end sequencing. The V4 region of the 16S rRNA gene (515F-806R) was amplified with region-specific primers that included the Illumina flowcell adapter sequences. The reverse amplification primer also contained a twelve base barcode sequence that supports pooling of up to 2,167 different samples in each lane[21]. Each 20μl PCR reaction contains 5μl of MoBio PCR Water (Certified DNA-Free), 10μl of Extract-N-Amp Ready Mix, 1μl of Forward Primer (5uM concentration, 200pM final), 1μl Golay Barcode Tagged Reverse Primer (5μM concentration, 200pM final), and 4μl of template DNA. The conditions for PCR were as follows: 94°C for 3 minutes to denature the DNA, with 35 cycles at 94°C for 45s, 50°C for 60s, and 72°C for 90s; with a final extension of 10 minutes at 72°C to ensure complete amplification. PCR amplifications were completed in triplicate, and then pooled. Following pooling, amplicons were quantified using PicoGreen (Invitrogen) and a plate reader. Once quantified, different volumes of each of the products were pooled into a single tube so that each amplicon was represented equally. This pool was then cleaned using the UltraClean® PCR Clean-Up Kit (MoBIO) and quantified using Qubit (Invitrogen). After quantification, the molarity of the pool was determined and diluted to 2nM, denatured, and then diluted to a final concentration of 4pM with a 30% PhiX spike for loading on the Illumina HiSeq2000 sequencer. Amplicons were then sequenced in two 151bp×12bp HiSeq2000 runs using custom sequencing primers and procedures described in the supplementary methods of reference [21].

Quality Control and Sequence Clustering

Paired end reads were quality trimmed and processed for operational taxonomic unite (OTU) clustering using the open reference method implemented in the QIIME pipeline[22]. The sequence identity cutoff set at 97% and taxonomy was assigned to the high quality (<1% incorrect bases) candidate OTUs using the *parallel_assign_taxonomy_rdp.py* script of the

QIIME software. Multiple sequence alignment and phylogenetic reconstruction was performed using PyNast and FastTree. OTUs containing less than 5 reads were discarded and the OTU table was rarefied to an even depth of 5000 reads.

Oligotyping

We used the Oligotyping pipeline[23] to identify sub-OTU level variation in 4 highly abundant genera: *Acinetobacter*, *Cornebacterium*, *Streptococcus* and *Staphylococcus*. Usearch was used to align reads back to OTUs based on a 97% identity cut-off and mapped reads were quality trimmed using the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). The minimum substantive abundance threshold for an Oligotype (-M) was set to 500 reads and the minimum number of samples (-s) and percent abundance cutoff (-a) were set to 1800 and 5%, respectively.

PC Space Correlation

We calculated the weighted UniFrac distance[24] between each pair of samples and then found the principal coordinates (eigenvectors) of the distance matrix. To reduce the complexity of the data and minimize noise we focused only on the minimum set of eigenvectors whose eigenvalues summed to fifty percent of the variance, which for our distance matrix were the first ten. We calculated the PC correlation along these 10 eigenvectors (n) such that our PC correlation (ρ) was an average of the Pearson correlation along eigenvectors i weighted by their associated eigenvalues (example calculation in Figure S7):

$$\rho = \sum_{i=1}^{n} \rho_i \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}$$

Correlation along each eigenvector was checked for significance using the *cor.test* function in R, and all non-significant correlations ($p > 0.05$) were reduced to zero before averaging.

Correlations were determined between pairs of samples taken from within the same environment (patient room, nurse station, hospital staff) on the same day. For between environment comparisons, such as room floor to station floor, samples could only be linked by the same building level (level 9 or level 10) and day, resulting in pairwise comparisons between all possible combinations of samples (nurse station floor to all 5 patient room floor samples taken that day on level 9, for example). Glove and water samples were excluded from these analyses due to the small and unique subset of ordination space they occupied in the PCoA of all samples (Figure S8).

Supervised Learning

Random forest supervised learning models were used to determine the diagnostic power of microbial community profiles in predicting whether hand and nose samples were taken from a hospital patient or staff member. The models were run using the *supervised_learning.py* command in QIIME, with 1,000 trees per model and 10-fold cross validation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Lax, Simon, Nagler, Cathryn R., Gilbert, Jack A. Our interface with the built environment: immunity and the indoor microbiota. Trends in Immunology. 2015; 36(3):121–123. [PubMed: 25754179]
2. Fujimura, Kei E., et al. House dust exposure mediates gut microbiome Lactobacillus enrichment and airway immune defense against allergens and virus infection. Proceedings of the National Academy of Sciences. 2014; 111(2):805–810.
3. Stefka, Andrew T., et al. Commensal bacteria protect against food allergen sensitization. Proceedings of the National Academy of Sciences. 2014; 111(36):13145–13150.
4. Blaser, Martin J., Falkow, Stanley. What are the consequences of the disappearing human microbiota? Nature Reviews Microbiology. 2009; 7(12):887–894. [PubMed: 19898491]
5. Adams, Rachel I., et al. Microbiota of the indoor environment: a meta-analysis. Microbiome. 2015; 3(1):1–18. [PubMed: 25621171]
6. Wood, Mariah, et al. Athletic equipment microbiota are shaped by interactions with human skin. Microbiome. 2015; 3(1):1–8. [PubMed: 25621171]
7. Lax, Simon, et al. Longitudinal analysis of microbial interaction between humans and the indoor environment. Science. 2014; 345(6200):1048–1052. [PubMed: 25170151]
8. Department of Health and Human Services. Summary of Notifiable Diseases – United States, 2007. Centers for Disease Control and Prevention; 2009.
9. Anderson, Robert N., Smith, Betty L. Deaths: leading causes for 2002. National Vital Statistics Reports. 2005; 53(17):67–70.
10. Groseclose SL, et al. Summary of notifiable diseases – United States, 2002. MMWR Morb Mortal Wkly Rep. 2004; 51(1)
11. Quick, Joshua, et al. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. Genome Biology. 2015; 16:114. [PubMed: 26025440]
12. Brooks, Brandon, et al. Microbes in the neonatal intensive care unit resemble those found in the gut of premature infants. Microbiome. 2014; 2(1):1. [PubMed: 24468033]
13. Hewitt, Krissi M., et al. Bacterial diversity in two neonatal intensive care units (NICUs). PloS One. 2013; 8(1):e54703. [PubMed: 23372757]
14. Poza, Margarita, et al. Exploring bacterial diversity in hospital environments by GS-FLX Titanium pyrosequencing. PloS one. 2012; 7(8):e44105–e44105. [PubMed: 22952889]
15. Lax, Simon, Gilbert, Jack A. Hospital-associated microbiota and implications for nosocomial infections. Trends in Molecular Medicine. 2015
16. Ramos, Tiffanie, et al. Spatial and temporal variations in indoor environmental conditions, human occupancy, and operational characteristics in a new hospital building. PloS One. 2015; 10(3):e0118207. [PubMed: 25729898]
17. Dedesko, Sandra, et al. Methods to assess human occupancy and occupant activity in hospital patient rooms. Building and Environment. 2015; 90:136–145.
18. Lax, Simon, et al. Forensic analysis of the microbiome of phones and shoes. Microbiome. 2015; 3(1):21. [PubMed: 25969737]
19. Meadow, James F., Altrichter, Adam E., Green, Jessica L. Mobile phones carry the personal microbiome of their owners. PeerJ. 2014; 2:e447. [PubMed: 25024916]
20. Knights, Dan, et al. Bayesian community-wide culture-independent microbial source tracking. Nature Methods. 2011; 8(9):761–763. [PubMed: 21765408]
21. Caporaso, J Gregory, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME journal. 2012; 6(8):1621–1624. [PubMed: 22402401]
22. Caporaso, J Gregory, et al. QIIME allows analysis of high-throughput community sequencing data. Nature Methods. 2010; 7(5):335–336. [PubMed: 20383131]

23. Eren, A Murat, et al. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. Methods in Ecology and Evolution. 2013; 4(12):1111–1119.

24. Lozupone, Catherine, Knight, Rob. UniFrac: a new phylogenetic method for comparing microbial communities. Applied and Environmental Microbiology. 2005; 71(12):8228–8235. [PubMed: 16332807]

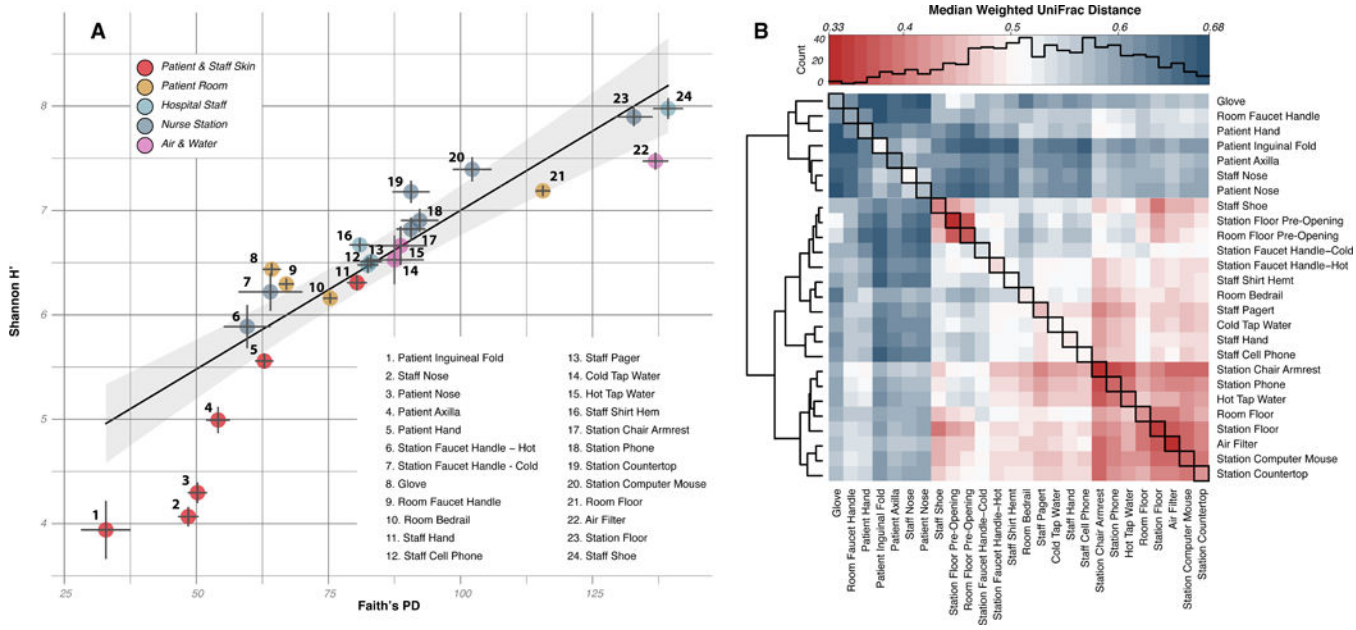

Figure 1. Alpha and Beta Diversity of Hospital Sample Types

(**A**) Average alpha diversity of sample types based on Faith's phylogenetic diversity (x-axis) and the Shannon diversity index (y-axis). Bars indicate standard error of the mean. (**B**) Heat map of beta diversity relationships between sample types based on the median weighted UniFrac distance between pairwise comparisons. Sample groups are clustered based on similarity in beta diversity patterns and median distances within individual sample types are highlighted in black along the diagonal.

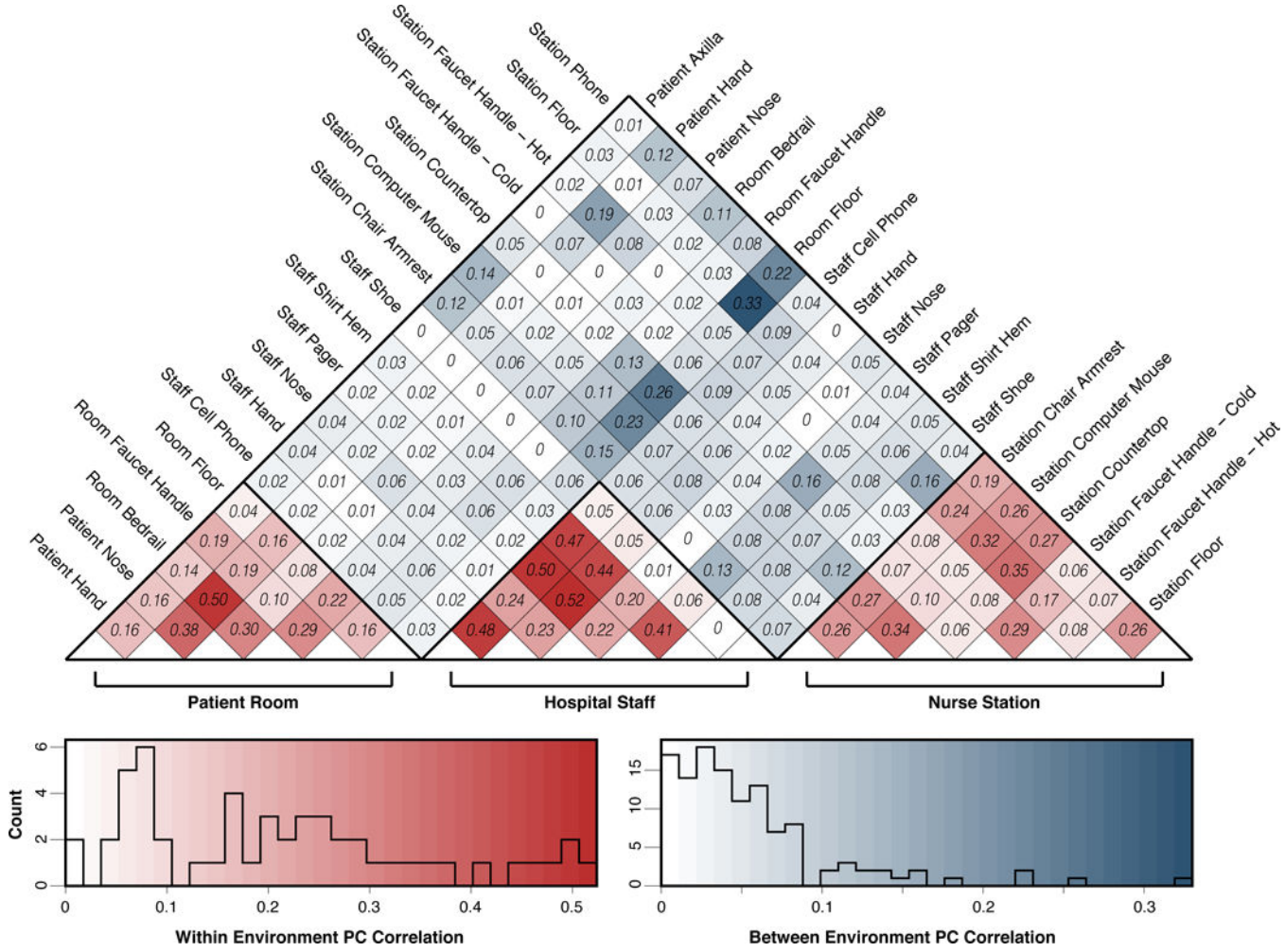

Figure 2. Heat map of PC Space Correlations between Sample Types

Within- and between-environment comparisons are differentiated by color scheme.

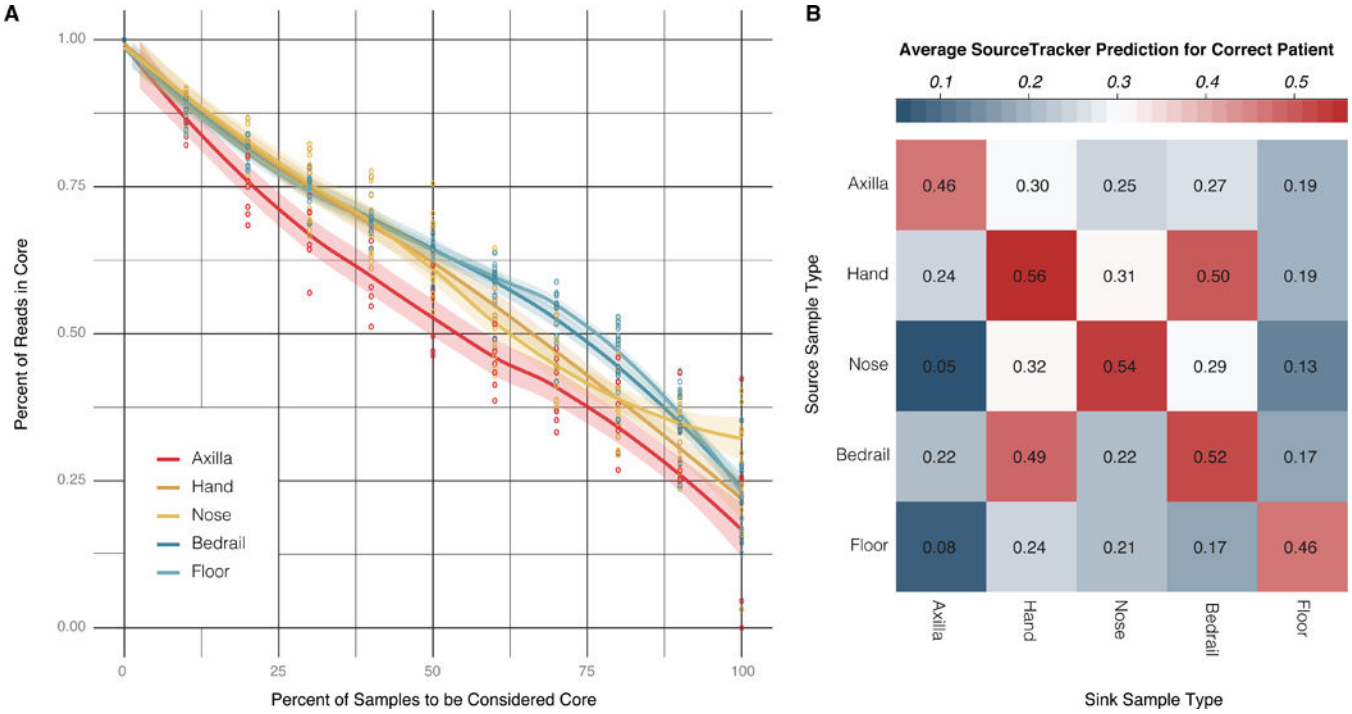

Figure 3. Variability in Patient Room Microbiota

(**A**) Scatter plot of the percent of 16S reads in the "core microbiome" for the 8 rooms sampled weekly, with the core definition on the x-axis and the percent of reads in the core on the y-axis. Points represent the 8 individual rooms while the trend line is a moving average of the data. (**B**) Heat map of the predictive accuracy of SourceTracker models that used samples taken from the first day of a patient's stay (source sample) to predict which patient a day 2 sample was taken from (sink sample).

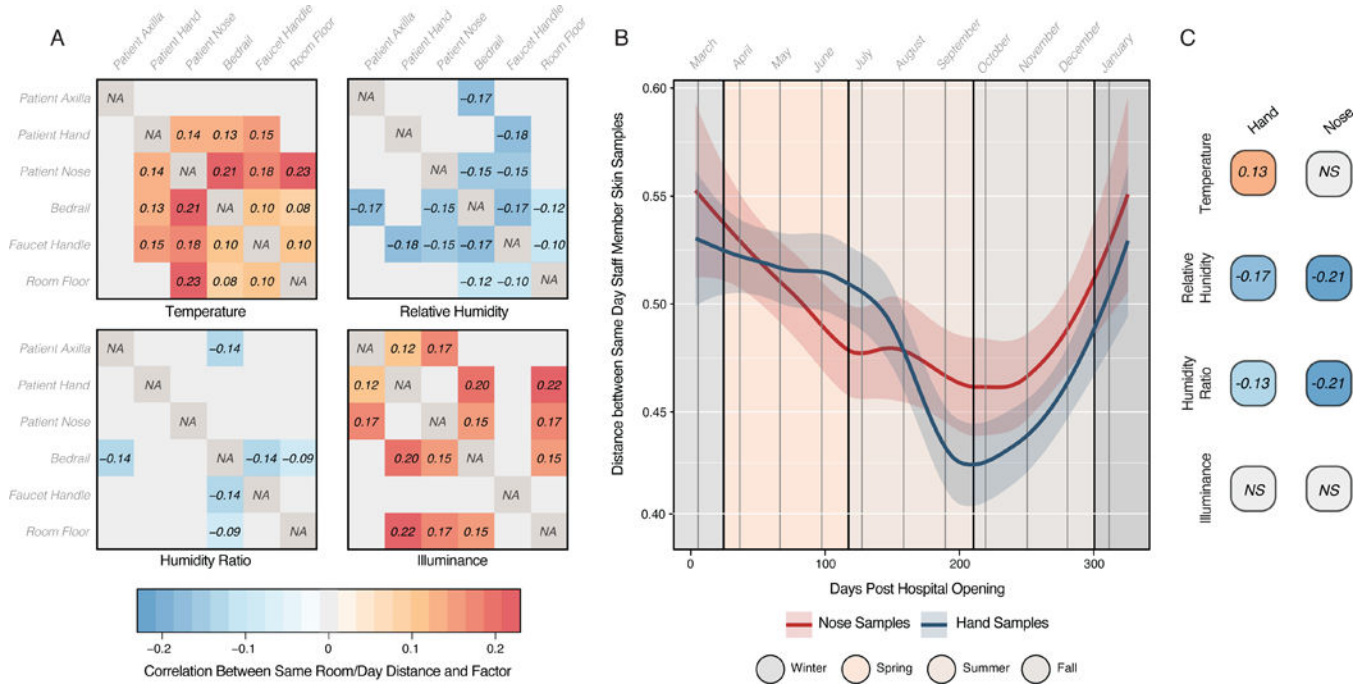

Figure 4. Effect of Environmental Factors on Microbial Transmission

(**A**) Heat maps of the correlation between environmental factors and the weighted UniFrac distance between samples taken from the same room on the same day. (**B**) Seasonal change in the distances between the hands and noses of nurses working on the same floor. Trend lines are a smoothed moving average of the data. (**C**) Correlations between environmental factors and the distances of hand and nose samples for nurses working on the same floor on the same day. Color scheme is as in (A).