Bacterial communities on classroom surfaces

Manuscript demo

James F Meadow[1]

[1] Biology and the Built Environment Center, Institute of Ecology and Evolution, University of Oregon, Eugene, OR USA, jfmeadow@gmail.com

---

Introduction

The data used here are a small subset (first 20,000 quality-filtered sequences) of those previously published (Meadow et al. 2014). This demo illustrates a few basic multivariate analysis methods with a sample dataset. In the original manuscript, we investigated the sources of microbes on classroom surfaces, and whether those microbial communities reflect common human contact with indoor surfaces.

Methods

This sequence dataset was processed using QIIME 1.8 (Caporaso et al. 2010) with a default MacQIIME installation (http://www.wernerlab.org/software/macqiime). Scripts for processing raw data are in the `../QIIME/` folder. To pick OTUs in that folder, you will execute the `pickTheseOTUs.sh` script sitting in that folder. This script wants to run MacQIIME, so if you are not using MacQIIME, you'll probably need to alter the top line to reflect your system.

For statistical analyses, we primiarily used the `phyloseq` package to handle QIIME output files, and `vegan` and `labdsv` for multivariate ecology stats (Mc-Murdie and Holmes 2013; Oksanen et al. 2011; Roberts 2010). All sequences were rarefied to an equal sampling depth (100 sequences per sample) prior to analysis. Beta-diversity was calculated using the Canberra taxonomic metric. The Canberra metric is defined as:

$$d_{jk} = \frac{1}{NZ} \sum \frac{x_{ij} - x_{ik}}{x_{ij} + x_{ik}}$$

where *NZ* is the number of non-zero entries. Reproducible documents were created with the `knitr` package in R (Xie 2013).

Results

Out of a total 1.5923×104 sequences that passed quality filtering, we analyzed 5800 sequences in 58 samples distributed among 966 OTUs (97% sequence similarity). The most abundant OTU in the dataset was a Cyanobacterium (2.67% of all sequences). The most abundant taxa are shown in Table 1.

|	Phylum	Family	Genus	Species	RelAbu
505954	Cyanobacteria	Xenococcaceae	-	-	2.67
1039477	Firmicutes	Staphylococcaceae	Staphylococcus	epidermidis	2.52
4449609	Proteobacteria	Sphingomonadaceae	Sphingomonas	-	2.40
359689	Actinobacteria	Corynebacteriaceae	Corynebacterium	-	2.16
4482309	Proteobacteria	Acetobacteraceae	-	-	2.14

Table 1: Most abundant taxa across all surfaces.
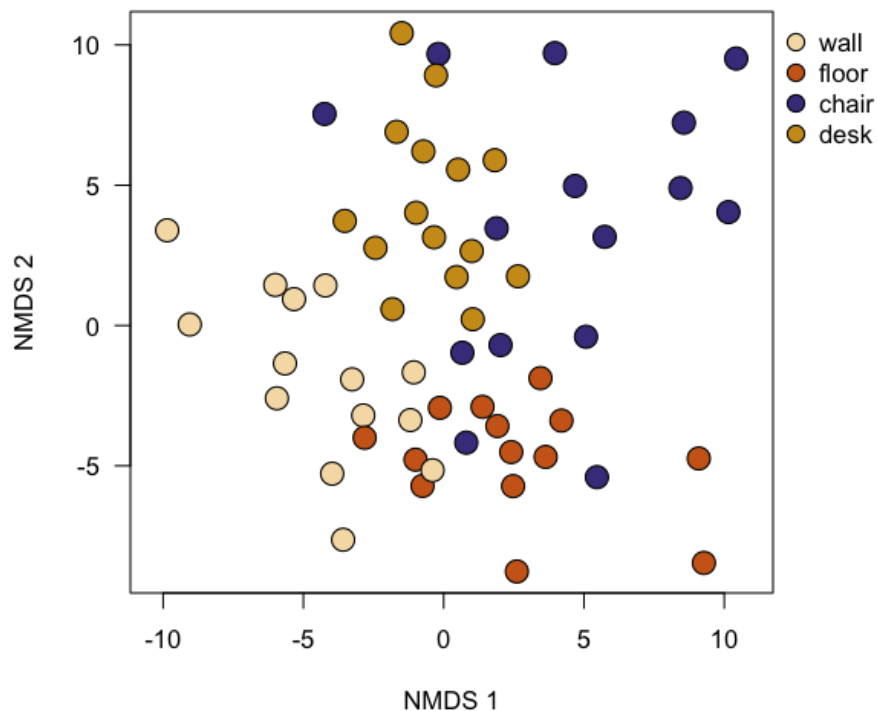
Figure 1: Samples cluster by the type of surface.

We found that surface type explained a significant amount of community variation (p = 0.001; from PERMANOVA on Canberra distances).

Next, we tested for a quasi-distance-decay relationship. This is the sort of pattern we see in just about every ecosystem with most forms of life. We even found this to be a stong predictor in the dust sampled from the entire building (Kembel et al. 2014). So we can use the x and y coordinates as a map of samples, and then calculate the Euclidean pairwise distance between all samples. Then that goes through a mantel test to determine if these distance are correlated with the community distances.

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
map$SurfaceType	3	2.14	0.71	1.80	0.09	0.001
Residuals	54	21.39	0.40		0.91	
Total	57	23.52			1.00	

Table 2: Surface type explains a significant amount of variation among communities.

We did not find any significant coorelation between community similarity and spatial distance (p = 0.456; from Mantel test) when considering all samples together. Likewise, individual sample types tested alone showed no relationship with spatial distance (p > 0.1 for all four sample types).

Discussion

So it looks like the type of surface, potentially as a proxy for human contact, explains a significant amount of variation, in the microbial communities on those surfaces, but their proximity to each other around the room doesn't seem to matter at all.

References

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, et al. 2010. "QIIME allows analysis of high-throughput community sequencing data." *Nature Methods* 7: 335–336.

Kembel, Steven W., James F. Meadow, Timothy K. O'Connor, Gwynne Mhuireach, Dale Northcutt, Jeff Kline, Maxwell Moriyama, G. Z. Brown, Brendan J. M. Bohannan, and Jessica L. Green. 2014. "Architectural design drives the biogeography of indoor bacterial communities." *PLOS ONE* 9 (01): e87093. doi:10.1371/journal.pone.0087093.

McMurdie, Paul J., and Susan Holmes. 2013. "phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data." Ed. Michael Watson. *PloS one* 8 (4) (jan): e61217. doi:10.1371/journal.pone.0061217. http://dx.plos.org/10.1371/journal.pone.0061217.

Meadow, James F., Adam E. Altrichter, Steven W. Kembel, Maxwell Moriyama, Timothy K. O'Connor, Ann M. Womack, G. Z. Brown, Jessica L. Green, and Brendan J. M. Bohannan. 2014. "Bacterial communities on classroom surfaces vary with human contact." *Microbiome* 2: 7.

Oksanen, Jari, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, and Helene Wagner. 2011. *vegan: Community Ecology Package*. http://cran.r-project.org/package=vegan.

Roberts, David W. 2010. *labdsv: Ordination and Multivariate Analysis for Ecology*. http://cran.r-project.org/package=labdsv.

Xie, Yihui. 2013. "knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, ed. Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman and Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.