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 [1]. 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 [2] 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 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 [3–5]. 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:

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

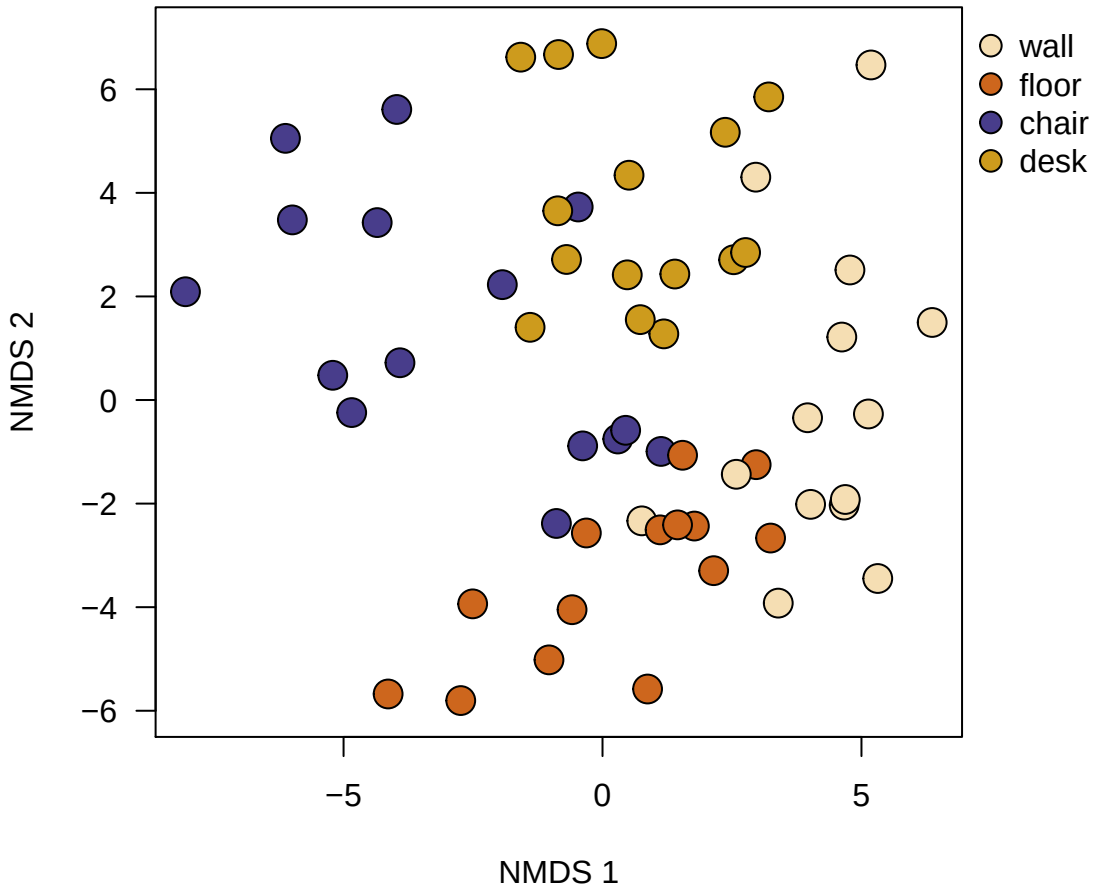
where NZ is the number of non-zero entries. Reproducible documents were created with the `knitr` package in R [6].

Results

Out of a total 1.5923×10^4 sequences that passed quality filtering, we analyzed 5800 sequences in 58 samples distributed among 964 OTUs (97% sequence similarity). The most abundant OTU in the dataset was a Cyanobacterium (3.14% of all sequences). The most abundant taxa are shown in Table 1.

	Phylum	Family	Genus	Species	RelAbu
505954	Cyanobacteria	Xenococcaceae	–	–	3.137931
4449609	Proteobacteria	Sphingomonadaceae	Sphingomonas	–	2.327586
1039477	Firmicutes	Staphylococcaceae	Staphylococcus	epidermidis	2.155172
359689	Actinobacteria	Corynebacteriaceae	Corynebacterium	–	2.017241
134470	Firmicutes	Alicyclobacillaceae	Alicyclobacillus	–	1.965517

Table 1: Most abundant taxa across all surfaces.



28

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
map$SurfaceType	3	2.196251	0.7320836	1.848144	0.0931142	0.001
Residuals	54	21.390385	0.3961182	NA	0.9068858	NA
Total	57	23.586635	NA	NA	1.0000000	NA

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

29 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 [7]. 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.

We did not find any significant coorelation between community similarity and spatial distance ($p = 0.653$; 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

1. Meadow JF, Altrichter AE, Kembel SW, Moriyama M, O'Connor TK, et al. (2014) Bacterial communities on classroom surfaces vary with human contact. Microbiome 2: 7.

2. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7: 335–336.

3. McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PloS one 8: e61217. Available: http://dx.plos.org/10.1371/journal.pone.0061217.

4. Oksanen J, Blanchet FG, Kindt R, Legendre P, O'Hara RB, et al. (2011) vegan: Community Ecology Package. Available: http://cran.r-project.org/package=vegan.

5. Roberts DW (2010) labdsv: Ordination and Multivariate Analysis for Ecology. Available: http://cran.r-project.org/package=labdsv.

6. Xie Y (2013) knitr: A Comprehensive Tool for Reproducible Research in R. In: Stodden V, Leisch F, Peng RD, editors. Implementing reproducible computational research. Chapman; Hall/CRC. Available: http://www.crcpress.com/product/isbn/9781466561595.

7. Kembel SW, Meadow JF, O'Connor TK, Mhuireach G, Northcutt D, et al. (2014) Architectural design drives the biogeography of indoor bacterial communities. PLOS ONE 9: e87093. doi:10.1371/journal.pone.0087093.