

Supplemental Figures:

Figure S1: Recruitment flow diagram

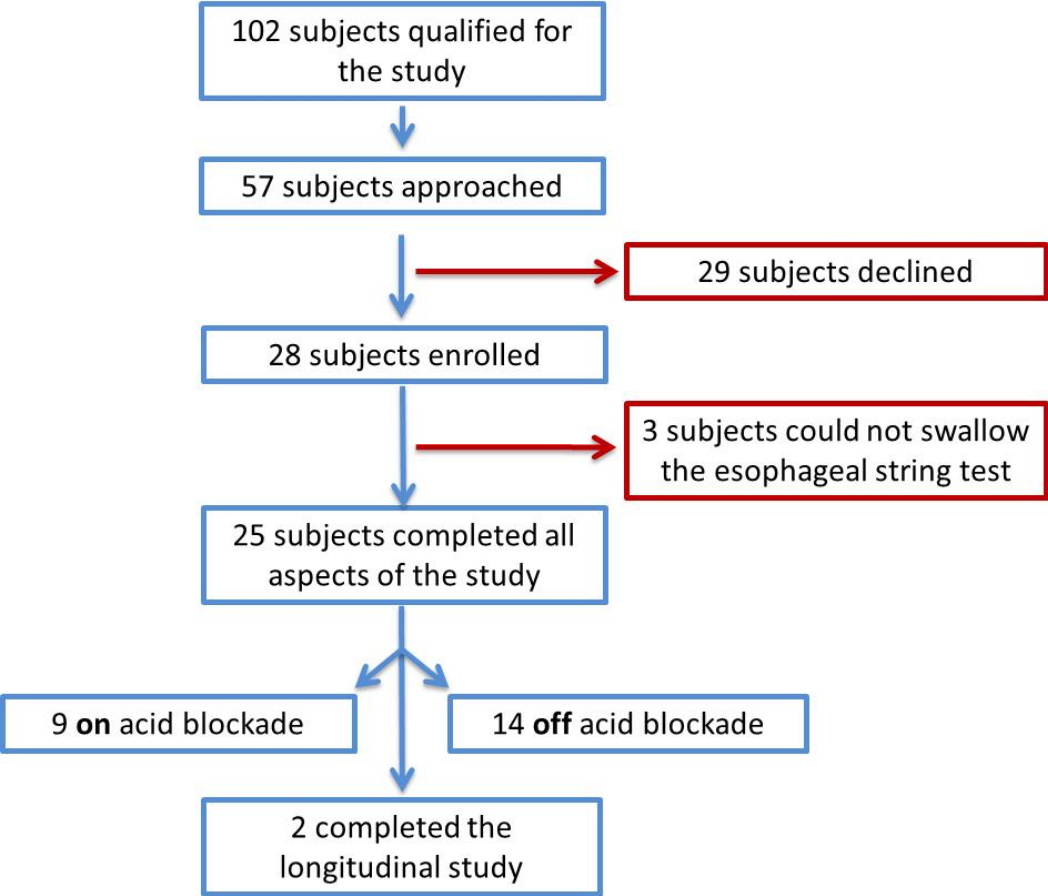


Figure S2: Multidimensional scaling. Morisita-Horn beta diversity was used as a distance metric in a multidimensional scaling plot to visualize relationship between samples based on community composition. Here the nares samples were removed from the plot. Samples continued to show poor clustering despite removal of the nares site.

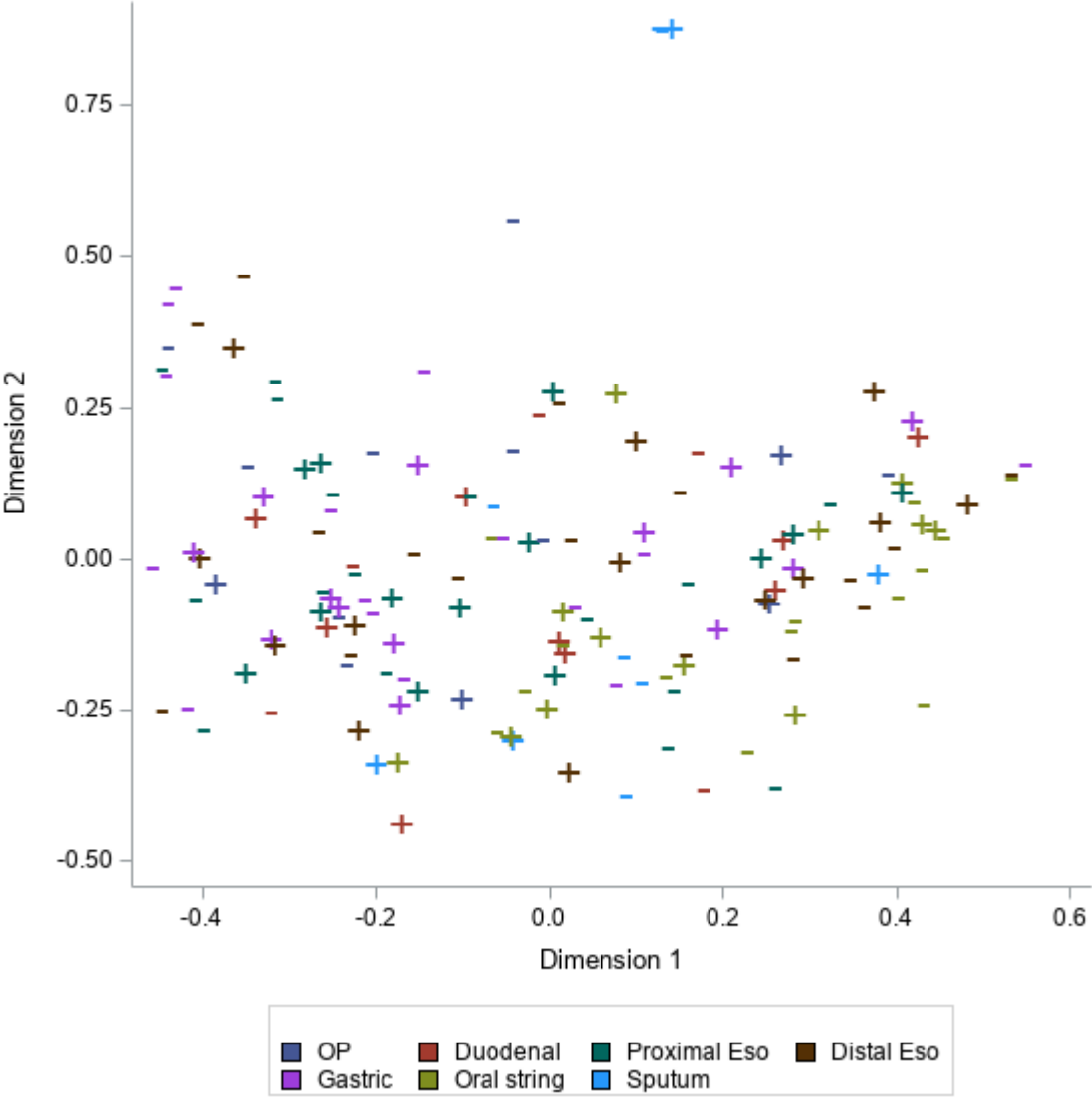


Figure S3: Comparison of sputum composition to oral, proximal and distal string. Ten individuals had sputum and string collected. The agreement appears to be similar between sputum composition and the different string locations except for subjects 4, 6, and 22.

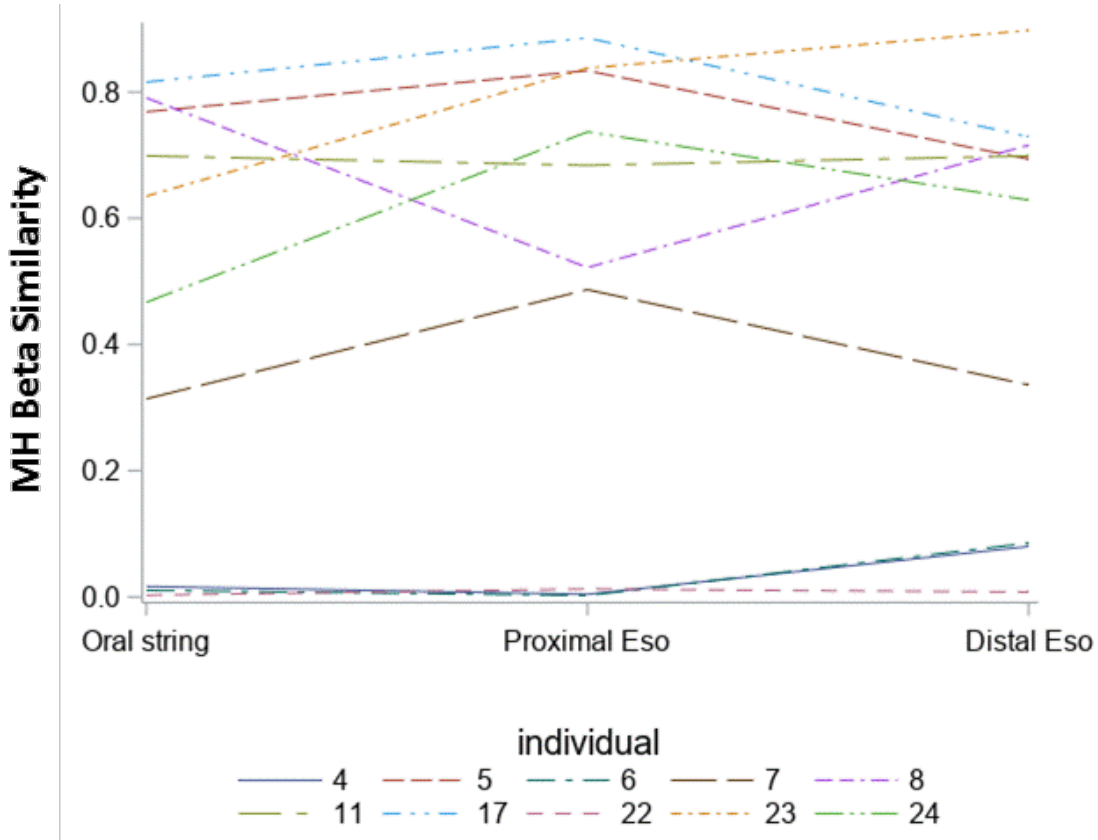
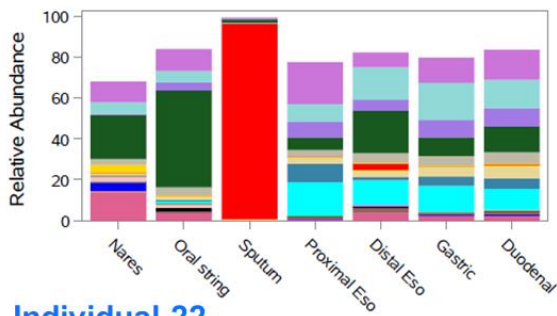
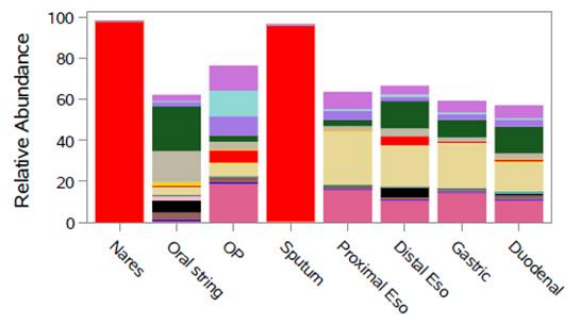


Figure S4: Microbial community composition in subjects with poor agreement between sputum and EST communities. Bar graphs indicate relative abundance of taxon for each individual at each site sampled. Each taxon is indicated in a different color. The height of the bar corresponds to the relative abundance. In all of these individuals, the low agreement in community structures is primarily due to the prominence of *Staphylococcus aureus* in the sputum sample.

Individual 4



Individual 6



Individual 22

