

# Composition of Phyla Communities by Sex

Jing Li

April 2022

## 1 Background

Microorganisms occupy many areas of the human body, appropriate levels of microbiomes are essential for immunity as they provide a barrier for harmful bacteria. Diverse factors such as diet, age, and the environment can influence the status of microbial communities in humans. Phyla compositions vary according to its body site and function. The upper respiratory tract is constantly interacting with the external environment through breathing, leading to a more diverse range of microbiomes. The aim of this data analysis is to investigate factors such as sex that could potentially affect distinctive patterns of phyla compositions in the nasal cavity.

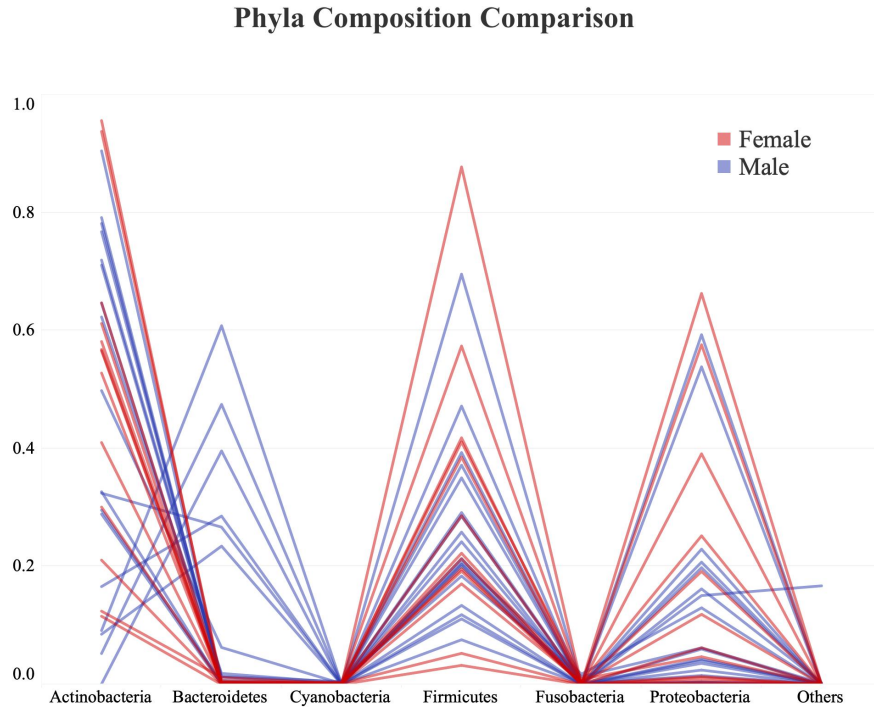
## 2 Subsetting data

Data set provided by NIH's Human Microbiome Project (HMP). 31 nasal cavity samples containing relative abundances of microbial phyla with some samples repeated for a few individuals. Metadata of each participant was provided in a separate CSV file. All files included an unique identifier for each sample, metadata was joined with the phyla data through the unique identifier to split between female and male samples for sex difference analysis.

Phyla data had a total of 22 taxonomy classifications, with most of them recorded as zeros. Actinobacteria, Bacteroidetes, Cyanobacteria, Firmicutes, Fusobacteria, and Proteobacteria were chosen as the top six for comparison given that most samples had values in place for those phyla columns. The rest is compiled to a new column with a new sum for relative abundance named 'Others'.

### 3 Overview of all samples

Fig. 1: Parallel Coordinates plot with all 31 samples separated by female and male. Microbial communities of each sample are plotted based on the top six most abundant microbes and lower level abundances as Others.



The plot above was the first approach to graphically visualize the data set. For the purpose of determining any significant differences in the composition of microbiomes by combining two categories (female and male). Line graph was used to better understand any trends or clustering of similar characteristics in the dataset.

Based on this first plot, an initial observation can be made about the relationship between males and Bacteroidetes. It appears that around half of the male sample possesses abundant levels of Bacteroidetes compared to the female counterpart. Distinct patterns can not be concluded regarding the overall phyla composition of females and males.

## 4 Quantitative analysis

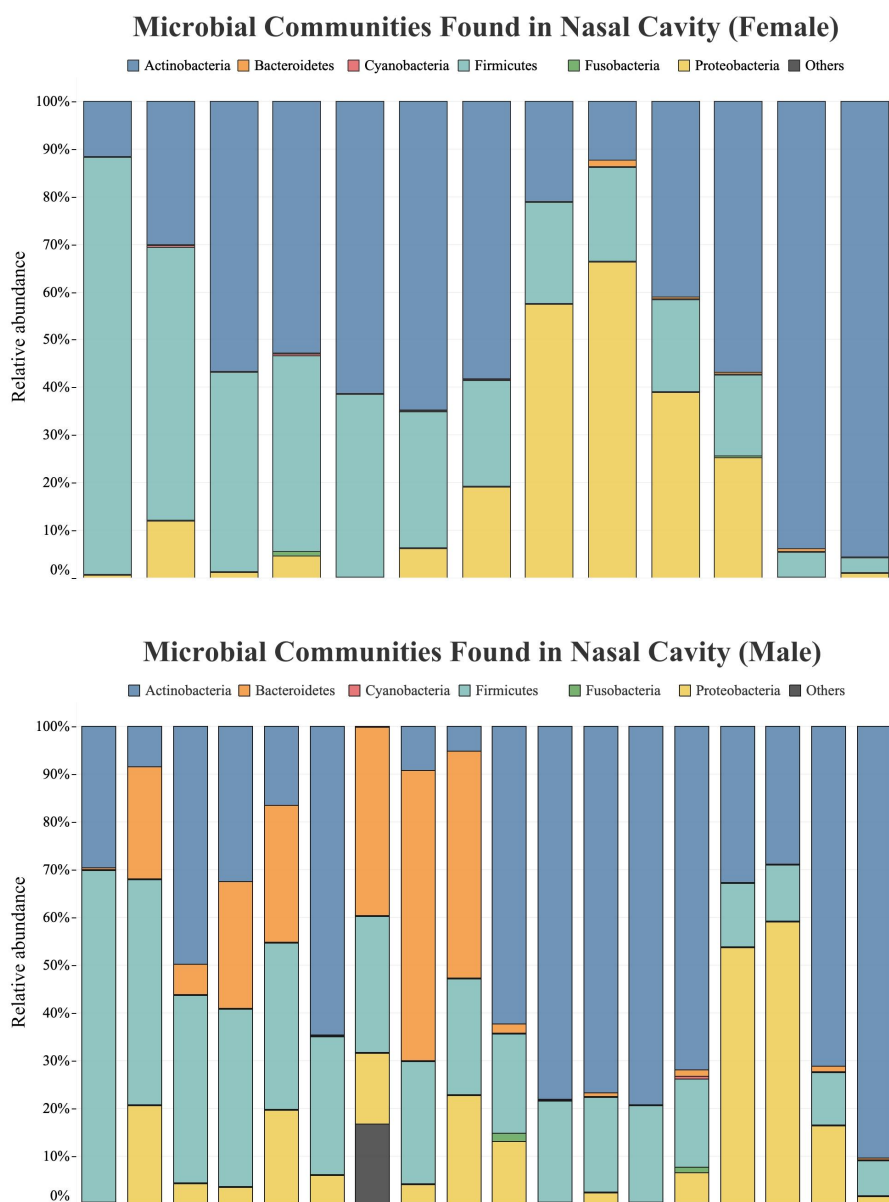
Fig. 2 (Female) and Fig. 2.1 (Male): Bacterial communities of female samples (n=13) and male samples (n=18). Values are depicted through a quantitative color scale to visualize phyla abundance.



Heat maps were created to easily direct attention to notable differences in the level of abundance. Findings were parallel to Fig.1, where Bacteroidetes presented itself as significant only in male samples. Actinobacteria, Firmicutes, and Proteobacteria were observed to have slightly greater presence in female samples compared to males.

## 5 Differences in phyla composition

Fig. 3 (Female) and Fig. 3.1 (Male): Stacked bar plot showing relative percentage of each phyla that makes up the microbial community of each sample.

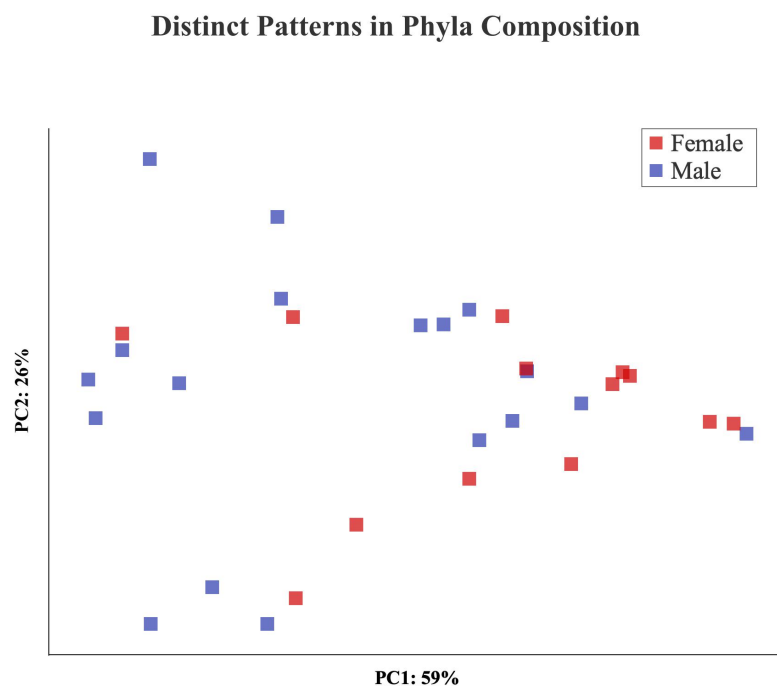


Results from the bar plots are similar to previous figures, using stacked bar plots to view each phyla as parts to a whole with percentages emphasizes the

presence of each phyla according to sex. Correlations can be observed within the phyla community rather than sex considering the high levels of variance in both categories. Phyla composition appears to be distinctive among individuals more than sex.

## 6 Finding distinct patterns

Fig. 4: PCA plot to show patterns of clustering based on similarities of microbial communities within the sex of the sample.



PCA was used to graphically present multivariate data to discover clusters and outliers. Overlapping of the data points indicate similarities between the two categories, clustering was very minimal in this case. Phyla composition does not seem to be characterized definitively according to sex.

## 7 Comparisons within phyla community

Fig. 5: Scatter plot representing the negative relationship between Actinobacteria and Firmicutes in all samples.

### Inverse Correlation of Actinobacteria and Firmicutes

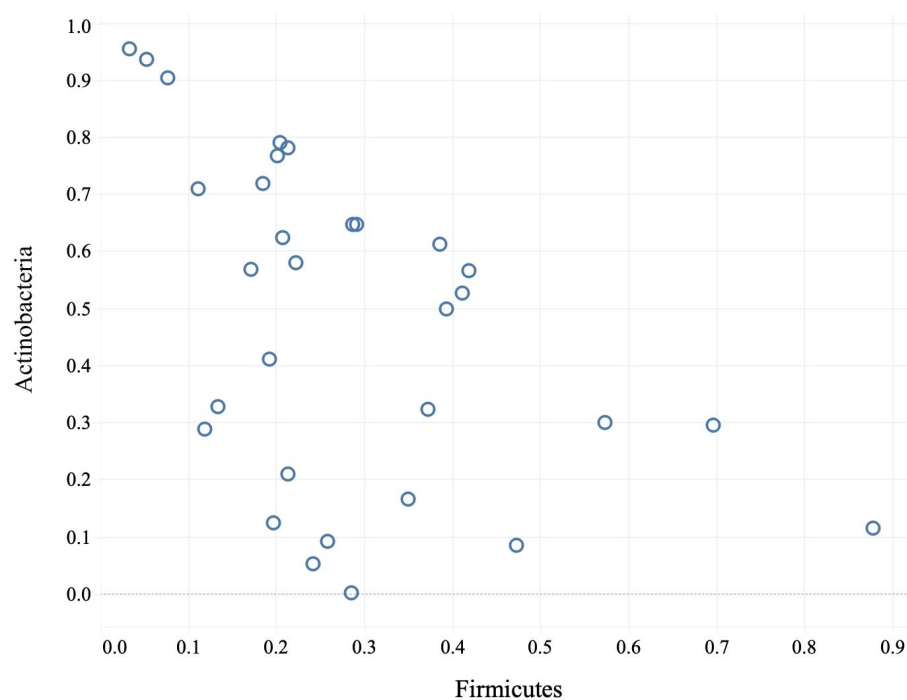
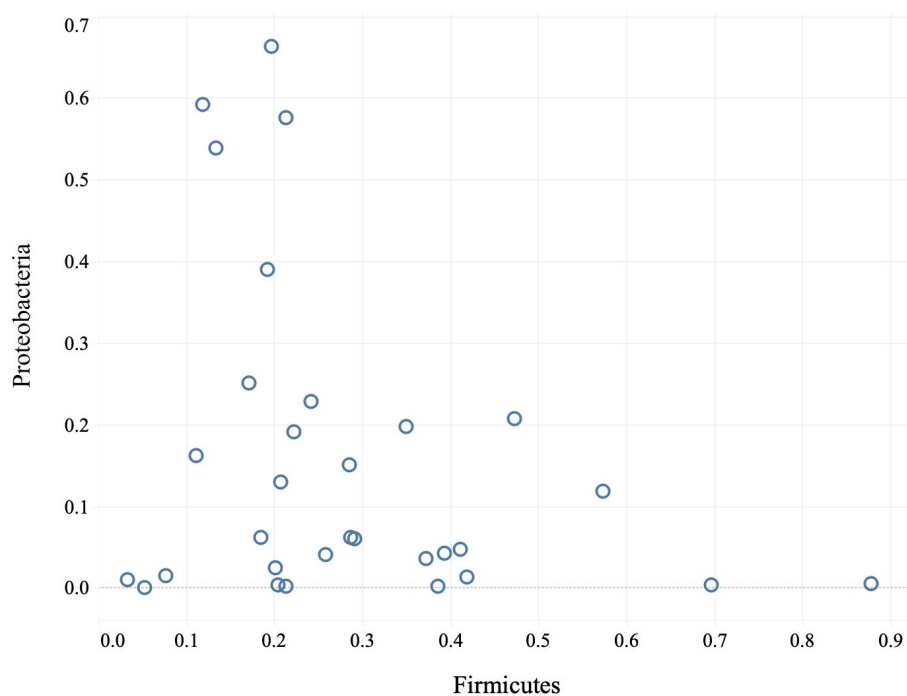


Fig. 5.1: Scatter plot representing the negative relationship between Proteobacteria and Firmicutes in all samples.

## Inverse Correlation of Proteobacteria and Firmicutes



From the figures made above to determine sex differences in phyla composition, patterns between microbiomes were also observed. Using all 31 samples, Actinobacteria exist in low levels as Firmicutes were heavily abundant. Same analysis can be made about Proteobacteria and Firmicutes. A negative correlation is found among those three microbiomes.

## 8 Conclusion

Sex differences in the phyla composition of the nasal cavity is found to be insignificant. There was great variation between female and male, results of microbial communities were not exclusive to either. However, it can be assumed that certain phyla can exist in higher levels depending on the sex. Relationships within the phyla community was apparent, but patterns may be distinct based on individuals. Considering that samples were collected from the nasal cavity, environmental factors can drastically change depending on the particles that are exchanged during the respiratory process. Further research with a larger sample size is needed to prove that sex can be a factor in distinct phyla composition.

## 9 Reference

Integrative Human Microbiome Project: NIH's Human Microbiome Project (HMP). <https://portal.hmpdacc.org/>

de Steenhuijsen Piters, Wouter A A et al. "The role of the local microbial ecosystem in respiratory health and disease." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* vol. 370,1675 (2015): 20140294. doi:10.1098/rstb.2014.0294