# DIFFERENCES BETWEEN COMMUNITY STRUCTURE OF THE HUMAN ORAL MICROBIOME AND VAGINAL MICROBIOME IN THE CONTEXT OF COMMUNITY ASSEMBLY

RACHAEL MENDOZA
*Bard College, 30 Campus Rd, Annandale-On-Hudson, NY 12504*

MENTOR SCIENTISTS: DR. MICHAEL J. HICKERSON AND ISAAC OVERCAST
*City College of New York, New York, NY 10031*

*Abstract*. We studied the microbial community structure of the human oral cavity and urogenital tract. Our research question was "are the oral and vaginal microbial communities structured similarly in the context of community assembly?" Our study's data originated from the Human Microbiome Project (Turnbaugh et al. 2007), which is available on their online data portal. We used a recently developed massive eco-evolutionary synthesis simulation (MESS; Overcast et al.) as the framework to compare these data to predictions from the Equilibrium theory of Island Biogeography (Macarthur & Wilson 1967). This simulation model generates the species abundance distribution and genetic diversity over time, and allows for parameterizing of both neutral and non-neutral assembly processes. To make easy comparison between observed data and simulated data, data are summarized into spectra of genetic diversity, phylogenetic diversity, and Shannon's diversity that describes the shape of the species abundance distribution.

We calculated these summary statistics from about 150 of independent human oral microbiome and vaginal microbiome communities using the PYTHON programming packages. Our first hypothesis is that the oral microbiome and vaginal microbiome have different sizes under the Equilibrium theory of Island Biogeography (ETIB) due to different exposure levels that are analogous to being closer to the mainland under the ETIB. Our second hypothesis is that the oral microbiome and vaginal microbiome are at different stages towards equilibrium under the ETIB. We found that the oral microbiome and vaginal microbiome may have different exposure levels and that the two microbial communities may be at different stages towards equilibrium.

## INTRODUCTION

Our proposed research is focused on studying the eco-evolutionary dynamics of human microbiome community assembly. The microbiome is the community of microorganisms, such as bacteria, archaea, and fungi, inhabiting the same location or inhabiting a particular host. The human microbiome project (HMP), conducted in 2008, aimed to sample and sequence the microbiome at 48 different anatomical locations from 300 healthy individuals. Each individual's microbiome per anatomical location was pooled and sequenced for the 16s rRNA gene to quantify patterns of abundance, and phylogenetic diversity across the different anatomical sites that included the oral cavity, urogenital tract, gastrointestinal tract, skin, and nasal passages. The sequence data and microbiome metagenomic datasets are available online (Human Microbiome Project Data Portal). Ecological theories along with the genetic data from the human microbiome project can be used together to explain community assembly and dynamics.

Two classic ecological models that explain community assembly include Macarthur & Wilson's Equilibrium Theory of Island Biogeography (ETIB) and Hubbell's unified neutral theory of biodiversity (UNTB) and biogeography (MacArthur and Wilson 2016; Hubbell 2011). Hubbell's unified neutral theory of biodiversity and biogeography is an individual-based extension of Macarthur & Wilson's ETIB. ETIB describes a dynamic equilibrium of communities in the number of species (i.e. species richness) on an island when immigration of new species and local extinction are balanced as the system reaches dynamic

equilibrium. The immigration and extinction rates are both functions of island size and distance from the mainland. At equilibrium, the larger and closer the island is to the mainland, the higher the species richness and diversity of the community. Hubbell's ETIB share those similarities with ETIB, by also defining equilibrium when local extinction is balanced with immigration from the metacommunity. UNTB differs from ETIB in that UNTB operates on the level of individual organisms, whereas ETIB only considered species as the focal unit. The individual based nature of UNTB, allows for predicting patterns of species richness as well the abundance distribution of species within the local community.

These classic neutral models predict that species richness and abundances patterns are due to stochastic ecological drift assuming ecological equivalence between species within a trophic level. In contrast, non-neutral models predict that community structure is driven by stochastic drift as well as species interactions and species specific differences in traits. These differences can take the form of either competition between species or interaction between individuals and the local environment (often termed as 'environmental filtering'). Similar to macrobial communities, microbial communities could be structured by either neutral or non-neutral processes.

To use these ecological theories, human body sites can be viewed as islands in space and time that are colonized by a meta-community. With this perspective, we can apply island biogeography theory such as ETIB, UNTB, and nonneutral models to understand processes underlying microbial diversity and assembly. Datasets that contain 16S rRNA-based observations about microbial diversity from the HMP may be aligned with the predictions from the neutral theory of community of assembly (Turnbaugh et al. 2007), to test whether human microbial communities structure neutrally or non-neutral interactions are important. Studying the processes that contribute to microbial community structure is important because an improved understanding of microbial community assembly and composition could illuminate mechanisms that cause differences between microbiomes of diseased individuals and healthy individuals.

Many studies have investigated the link between human microbiome structure and human health (Cho and Blaser 2012; Ma et al. 2012; Kilian et al. 2016). A previous study provided evidence that diseased individuals had a different microbial community composition than healthy individuals (Cho and Blaser 2012). For example, a previous study investigated whether human lung microbiomes are structured neutrally in healthy and diseased individuals (Venkataraman et al. 2015). They found that the composition of a healthy lung microbiome is consistent with the neutral model predictions and that the microbiome of diseased lungs harbored communities under active selection, meaning that species interactions in the community level and environment interactions were the dominant processes in community assembly. Another study investigated both the neutral and non-neutral model in the context of microbial diversity patterns in human microbial communities (Jeraldo et al. 2012). Similar to this study, Li and Ma (2016) tested the neutral theory on all of the body sites sequenced from the HMP. They found that the structure of the human microbiome across body sites is not well explained by neutral processes alone. However, this study only used the relative abundance data while ignoring DNA sequence data and non-equilibrium dynamics.

Although there has been extensive research on the urogenital microbiome and oral microbiome (Ravel et al. 2011; Ravel et al. 2012; Kilian et al. 2016), prior to the advent of the human microbiome project (HMP) variation in the structure and composition of microbial communities across body sites was poorly understood (Nemergut et al 2013; Shafquat et al. 2014). The urogenital microbiome has been shown to be a stable community, with little turnover in community structure over time (Ma et al. 2012). The vaginal microbiome has the lowest alpha diversity and the buccal mucosa microbiome has the highest median alpha diversity (Human Microbiome Project Consortium 2012). On the other hand, the oral microbiome is the second most diverse microbial community in the human body and the heterogeneous environment supports different microbial communities, meaning there is a lot of turnover over time (Kilian et al. 2016). Since the urogenital microbiome and oral microbiome have large differences in how they are assembled with regards

to sources and time scales, it is important to compare community structure and assembly processes for these two sites.

Our research question was "Are human buccal mucosa and vaginal microbiomes structured similarly? And to what extend do neutral or non-neutral processes contribute to this structure?" Based on previous studies done on the oral and urogenital microbiome, our hypothesis was that there will be differences in the spectra of abundances and genetic diversities between human buccal mucosa and vaginal microbiome communities that are consistent with the expectations of a neutral model in vaginal communities, a non-neutral model in the communities of the buccal mucosa. Specifically our two hypotheses that we tested for were that the buccal mucosa microbiome and vaginal microbiome have different exposure levels (i.e. analogous to island immigration rates) or sizes under the ETIB model, and that buccal mucosa microbiome and vaginal microbiome are at different stages towards equilibrium.

## METHODS

To characterize the two types of communities and compare with predictions under the ETIB and UNTB, we calculated nucleotide diversity, absolute divergence, species abundance, and Shannon's diversity index for each sampling location. With the datasets from the HMP, we used several PYTHON programming packages such as biopython, dendropy, pandas, skbio, and matplot to organize, visual, and analyze the observed data from the buccal mucosa and vaginal microbiome. We calculated community size, richness, Shannon's entropy, and Bray-Curtis dissimilarity based on the abundance data from the HMP. In addition to the abundance data, we calculated nucleotide diversity with the sequence data. We compared the summary statistics of buccal mucosa and vaginal microbiome to investigate the differences between the two microbiomes' structures. We utilized the two-sided t-test as the formal way of statistical analysis. From these results, we have developed an idea of where the differences lie in their community structures and an informed hypothesis of the processes of community assembly of the two microbiomes.

We chose the buccal mucosa and vaginal microbiome as the body sites of study because previous studies indicate they have differences in community structures, and there has been no studies done that directly compares the two. Also, the buccal mucosa and vaginal microbiome have major implications on human health (Kilian et al. 2016; Ma et al. 2012). Studying these two sites will give us insight about the differences of microbial community structure across anatomical locations. With the knowledge of the neutral and non-neutral processes, we can better understand community assembly based on location and begin to question what characteristics of the habitat causes the differences or similarities between the microbial communities. We limited our investigation to female individuals in the human microbiome project to avoid any potential confounding interactions with biological sex.

## RESULTS

The community structure of the buccal mucosa and vaginal microbiome have significant differences based on the abundance and DNA sequence data from the human microbiome project. The buccal mucosa and vaginal microbiome's richness was significantly different ($T = 9.214$, $p = 1.096 \text{ e-}17$, Fig 1). The buccal mucosa microbiome's richness mean was $862.961 \pm 391.680$, which was about 82% higher than vaginal microbiome's richness mean, $475.398 \pm 278.734$. Phylogenetic diversity between the two body sites was also significantly different ($T = 3.74$, $p = 0.000216$, Fig. 2). The buccal mucosa microbiome's phylogenetic diversity mean was $0.0011333 \pm 0.0002$, which was about 11% greater than the midvagina microbiome's phylogenetic diversity, $0.00099 \pm 0.0003$. The community size between the two body sites were not significantly different ($T=0.4335$, $p = 0.6650$, Fig. 3). The buccal mucosa microbiome's community size, $5452.8 \pm 3594.1$, was 3.5% greater than the vaginal microbiome's community size, $5266.1 \pm 3349.4$. Shannon's entropy between the two body sites were significantly different ($T= 5.308$, $p = 2.375 \text{ e-}07$, Fig.

4). The buccal mucosa microbiome's Shannon's entropy, 4.003 ± 1.006, was greater than the midvagina microbiome's Shannon's entropy, 4.003 ± 1.006. Most of the Bray-Curtis values were approximately 1 (Fig. 5). Also, difference between the buccal mucosa and midvagina microbiomes' amount of segregating sites was significantly different (T = 8.8003, p = 8.2054e-17, Fig. 6). The buccal mucosa microbiome contains 450.031±17.0740 segregating sites, which was 5.3% greater than the midvagina microbiome's amount of segregating sites, 427.66 ± 25.021.

## DISCUSSION

Our results supported our first hypothesis, which was the buccal mucosa microbiome and vaginal microbiome have different exposure levels under the ETIB model. The difference between species richness and phylogenetic diversity of the two body sites was significant (Fig. 1, Fig. 2). The buccal mucosa microbiome had a higher species richness and phylogenetic diversity than the midvagina microbiome based on their means, which means that the buccal mucosa microbiome have a higher species count and that the species are less related to each other than the midvagina's species. Assuming equilibrium under the ETIB model, size and/or exposure levels affect species richness and diversity of the community. Since community size was not significantly different (Fig. 3), we infer that the exposure levels might be different and could be correlated to the significant differences between species richness and phylogenetic diversity. In addition to these significant differences in community structure, most of the values in the Bray-Curtis dissimilarity distance matrix between the two body sites were approximately 1 (Fig. 5), which means that there is no species overlap between the two communities. There was also a significant difference between amount of segregating sites between the two communities (Fig. 6), and based on the mean the buccal mucosa microbiome have more segregating sites than the vaginal microbiome. Since segregating sites are the differences/mutations between related genes, which contributes to phylogenetic diversity, the bacterial species of the buccal mucosa community have more genetic differences between their related genes.

Our results also supported our second hypothesis, which was the buccal mucosa microbiome and vaginal microbiome are at different stages towards equilibrium. Shannon's entropy is correlated to a community's stage towards equilibrium. The higher the community's Shannon's entropy is, the closer the community is to equilibrium (Overcast et al.). The difference between the two body sites' Shannon's entropy was significantly different (Fig. 4), and based on the mean, buccal mucosa microbiome is more even than the vaginal microbiome. Since the buccal mucosa microbiome yielded a higher mean Shannon's entropy than the vaginal microbiome, we think that the buccal mucosa microbiome is closer to equilibrium, and the two communities are at different stages towards equilibrium.

Some of the uncertainties in our study originated from not having enough data on the human subjects. There are many variables that can contribute to the differences in community structure like the participants' lifestyle, eating habits, location, age, etc. Many factors such as those listed before can affect their microbiome (Turnbaugh et al. 2007). Also, our results support both of our two hypotheses, which is contradictory because our first hypothesis assumes equilibrium, and our second hypothesis does not assume equilibrium. In order to tease out our two hypotheses, we need to run a simulation model-based approach, which simulates data at all stages equilibrium.

## FUTURE DIRECTIONS

A simulation model-based approach could allow for the comparison of the different predictions of the neutral or non-neutral models given the observed data from the human microbiome project. Approximate Bayesian computation is the formal statistical method we will use to compare the microbiome data's fit to the neutral and non-neutral model. Approximate Bayesian computation (ABC) utilizes the summary statistics of both the observed and the simulated communities to test model fit and estimate assembly

parameters. Goodness-of-fit tests can be performed post-hoc to ensure the best fit model and estimated parameters are capable of reproducing the observed data. (Csilléry et al. 2010).

With the summary statistics calculated from the observed data, we are then able to compare the observed data with the summary statistics from the ABC simulations of neutral and nonneutral simulations to see whether or not the vaginal microbiome and buccal mucosa microbiome are better fit to either of those community assembly models. An ABC model could also produces simulations of the community at all stages towards equilibrium so with this data, we will also be able to tease out our two hypotheses and make a better inference of the factors correlated to the significant differences between the buccal mucosa and vaginal microbial community structures.

## ACKNOWLEDGEMENTS

I would like to acknowledge my two mentors Dr. Michael Hickerson and Isaac Overcast as well as the REU coordinators, Dr. Aude Lochet and Dr. Alan Berkowitz. I would also like to thank the Cary Institute of Ecosystem Studies and the National Science Foundation. My research would not be possible without the help of this community.

## LITERATURE CITED

Cho, I., and M. J. Blaser. 2012. The human microbiome: at the interface of health and disease. Nature reviews. Genetics **13**:260–270.

Csilléry, K., M. G. B. Blum, O. E. Gaggiotti, and O. François. 2010. Approximate Bayesian Computation (ABC) in practice. Trends in Ecology & Evolution **25**:410–418.

Hubbell, S. P. 2011. The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32). Princeton University Press.

Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. Nature **486**:207–214.

Jeraldo, P., M. Sipos, N. Chia, J. M. Brulc, A. S. Dhillon, M. E. Konkel, C. L. Larson, K. E. Nelson, A. Qu, L. B. Schook, F. Yang, B. A. White, and N. Goldenfeld. 2012. Quantification of the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes. Proceedings of the National Academy of Sciences of the United States of America **109**:9692–9698.

Kilian, M., I. L. C. Chapple, M. Hannig, P. D. Marsh, V. Meuric, A. M. L. Pedersen, M. S. Tonetti, W. G. Wade, and E. Zaura. 2016. The oral microbiome - an update for oral healthcare professionals. British Dental Journal **221**:657–666.

Li, L., and Z. S. Ma. 2016. Testing the neutral theory of biodiversity with human microbiome datasets. Scientific Reports **6**:31448.

Ma, B., L. J. Forney, and J. Ravel. 2012. Vaginal microbiome: rethinking health and disease. Annual Review of Microbiology **66**:371–389.

MacArthur, R. H., and E. O. Wilson. 1967. The Theory of Island Biogeography. Princeton, N.J: Princeton University Press.

Mikkelson, G. M. 2005. Niche-based vs. neutral models of ecological communities. Biology and Philosophy **20**:557–566.

Nemergut, D. R., S. K. Schmidt, T. Fukami, S. P. O'Neill, T. M. Bilinski, L. F. Stanish, J. E. Knelman, J. L. Darcy, R. C. Lynch, P. Wickey, and S. Ferrenberg. 2013. Patterns and processes of microbial community assembly. Microbiology and Molecular Biology Reviews **77**:342–356.

Overcast, I., B. C. Emerson, and M. J. Hickerson. 2019. An integrated model of population genetics and community ecology. Journal of Biogeography **46**:816–829.

Petkova, D., J. Novembre, and M. Stephens. 2016. Visualizing spatial population structure with estimated effective migration surfaces. Nature Genetics **48**:94–100.

Ravel, J., P. Gajer, Z. Abdo, G. M. Schneider, S. S. K. Koenig, S. L. McCulle, S. Karlebach, R. Gorle, J.

Russell, C. O. Tacket, R. M. Brotman, C. C. Davis, K. Ault, L. Peralta, and L. J. Forney. 2011. Vaginal microbiome of reproductive-age women. Proceedings of the National Academy of Sciences of the United States of America **108**:4680–4687.

Rosindell, J., and L. J. Harmon. 2013. A unified model of species immigration, extinction and abundance on islands. Journal of Biogeography **40**:1107–1118.

Shafquat, A., R. Joice, S. L. Simmons, and C. Huttenhower. 2014. Functional and phylogenetic assembly of microbial communities in the human microbiome. Trends in Microbiology **22**:261–266.

Turnbaugh, P. J., R. E. Ley, M. Hamady, C. M. Fraser-Liggett, R. Knight, and J. I. Gordon. 2007. The human microbiome project. Nature **449**:804–810.

Ursell, L. K., J. L. Metcalf, L. W. Parfrey, and R. Knight. 2012. Defining the human microbiome. Nutrition Reviews **70**:S38–S44. doi:10.1111/j.1753-4887.2012.00493.x

Venkataraman, A., C. M. Bassis, J. M. Beck, V. B. Young, J. L. Curtis, G. B. Huffnagle, and T. M. Schmidt. 2015. Application of a neutral community model to assess structuring of the human lung microbiome. mBio **6**:e02284-14. https://doi.org/10.1128/mBio.02284-14

## APPENDIX

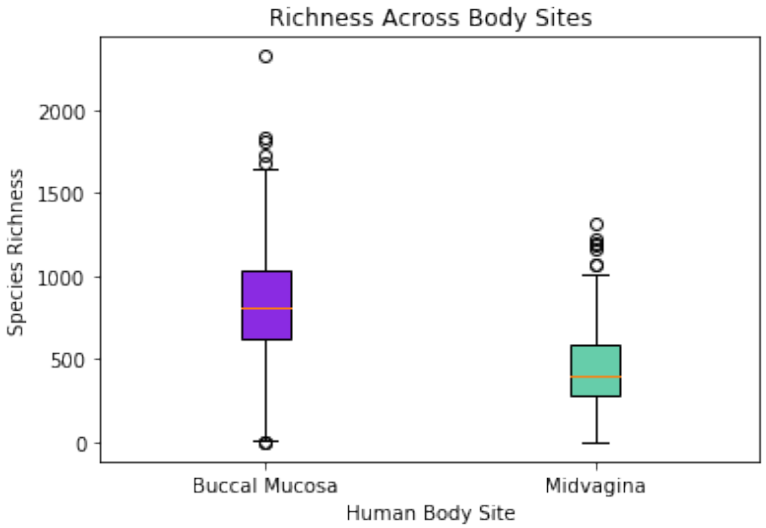


**FIGURE 1**. The comparison between buccal mucosa and midvagina microbiome's alpha diversity. The difference between richness was significant (T= 9.214, p = 1.096 e-17). The midvagina microbiome's richness mean (475.398 ± 278.734) is approximately 82% lower than buccal mucosa microbiome's richness (862.961 ± 391.680).

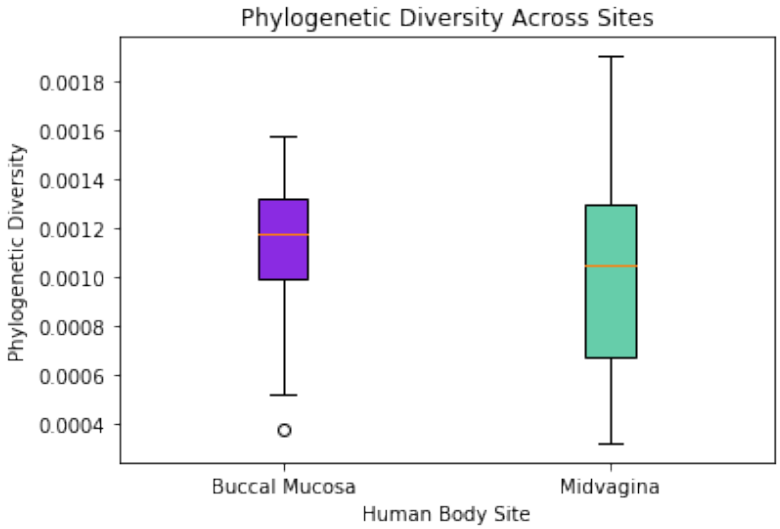


**FIGURE 2**. The comparison of phylogenetic diversity between buccal mucosa and midvagina microbiomes. The difference between the phylogenetic diversity of the two communities was significant (T = 3.74, p = 0.000216). The midvagina microbiome's phylogenetic diversity mean was 0.00099 ± 0.0003, which was less than buccal mucosa microbiome's phylogenetic diversity mean, 0.0011333 ± 0.0002.

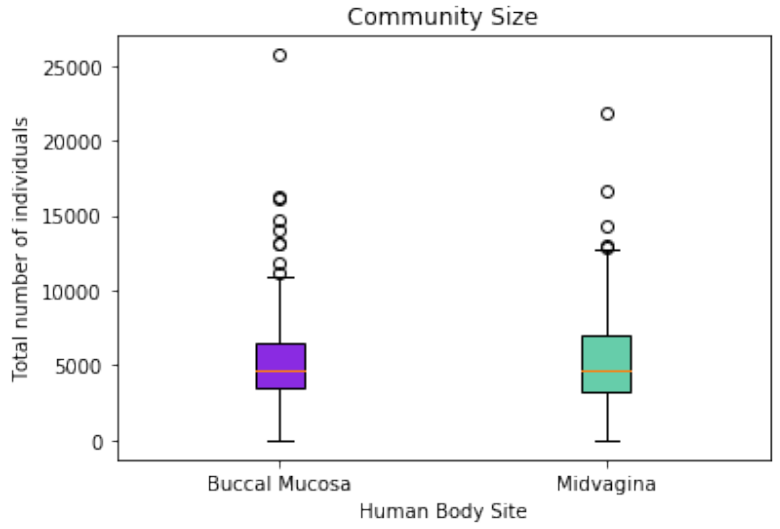


**FIGURE 3**. The comparison of community size between buccal mucosa and midvagina microbiomes. The difference between the community size of the two communities was significant (T=0.4335, p = 0.6650). The buccal mucosa community size's mean is 5452.8±3594.1, and the midvagina community size's mean is 5266.1±3349.4.

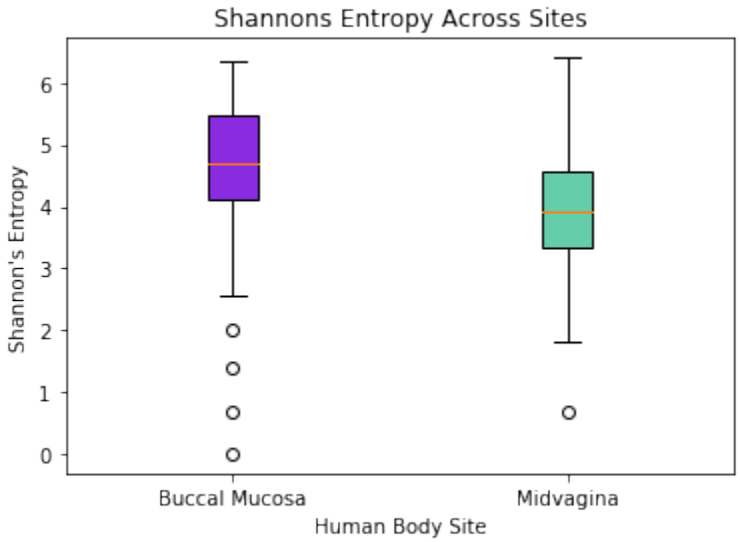


**FIGURE 4**. The comparison between buccal mucosa and midvagina microbiomes' Shannon's entropy. The difference between Shannon's entropy values were significant (T= 5.308, p = 2.375 e-07). Midvagina microbiome's Shannon's entropy mean was 4.003 ± 1.006, which is less than buccal mucosa microbiome's Shannon's entropy mean, 4.682 ± 1.056.

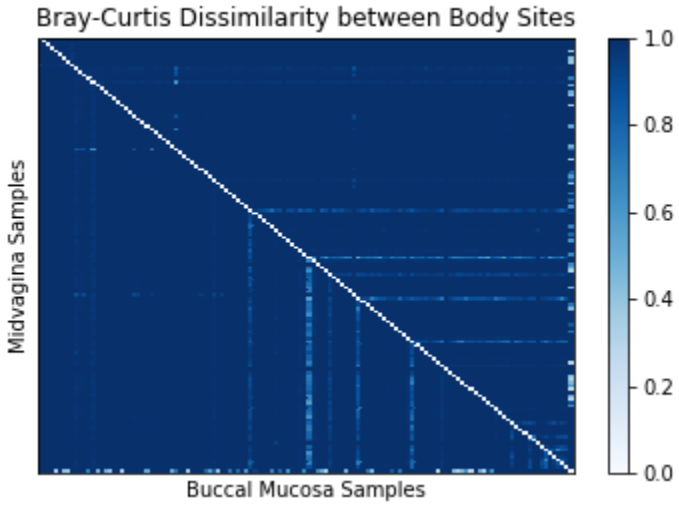


**FIGURE 5**. Bray-Curtis Dissimilarity between buccal mucosa and midvagina sequence data.

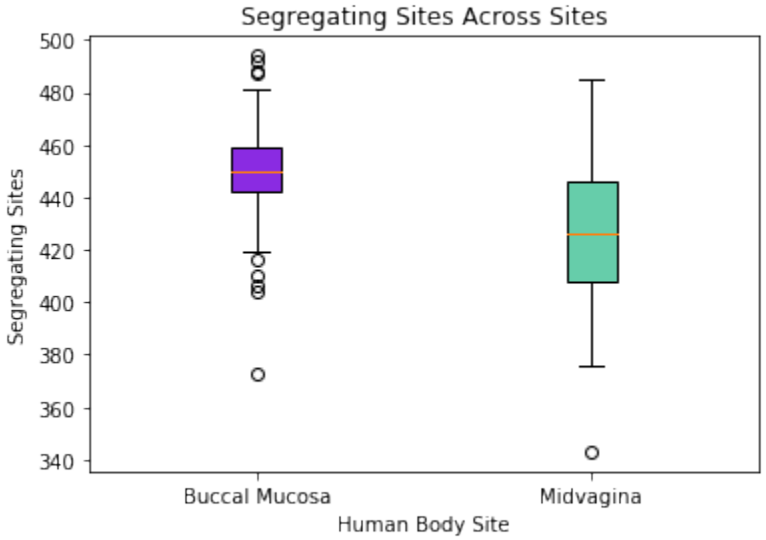


**FIGURE 6**. The comparison of segregating sites between buccal mucosa and midvagina microbiomes. The difference between the segregating sites between the two communities was significant (T = 8.8003, p = 8.2054e-17). The midvagina microbiome's segregating sites mean was 427.66 ± 25.021, which was less than buccal mucosa microbiome's segregating sites mean, 450.031±17.0740.