Exploring a possible correlation between the human oral microbiome and body mass index

Ivy S. McDermott
ivmc6708@colorado.edu

Follow this and additional works at: http://scholar.colorado.edu/honr_theses
Part of the Biodiversity Commons, Computational Biology Commons, and the Other Public Health Commons

Recommended Citation
Exploring a possible correlation between the human oral microbiome and body mass index

By

Ivy McDermott

Department of Molecular, Cellular, and Developmental Biology
University of Colorado at Boulder

Defense date: March 28, 2016

Thesis Advisor:
Dr. Kenneth Krauter, Department of Molecular, Cellular, and Developmental Biology

Defense Committee:
Dr. Kevin Jones, Department of Molecular, Cellular, and Developmental Biology
Dr. Noah Fierer, Department of Ecology and Evolutionary Biology
# Table of Contents

Abstract.........................................................................................................................3  
Introduction..................................................................................................................4-5 
Significance..................................................................................................................5-6 
Materials and Methods...............................................................................................6-8 
Results..........................................................................................................................9-14 
Discussion...................................................................................................................15-16  
Acknowledgements.....................................................................................................16 
References.....................................................................................................................17-21  
Appendix I.....................................................................................................................22-23
Abstract:

Obesity is a growing health problem in America affecting more than a third of Americans\textsuperscript{25} and is quickly becoming a global health crisis\textsuperscript{34}. Recently there has been much interest in the possible link between the human gut microbiome and obesity as fecal transplants may serve as a potential therapeutic treatment. However, not much research has been done looking at a potential association with the human oral microbiome and obesity. In this thesis, I examined 976 individuals previously sequenced for their oral microbiome. Each sample was classified as underweight, normal, overweight, or obese, according to their BMI. I measured the microbial diversity of each sample and compared the relative diversity of each class through alpha and beta diversity. Alpha diversity measures the microbial diversity within an individual while beta diversity measures the microbial diversity between individuals. I also investigated whether there was a clear association between monozygotic twins who were discordant in phenotype. While phyla-level changes were detected in the different weight class, overall the oral microbiome does not appear to be associated with human weight.
Introduction:

The human oral microbiome is composed of a diverse community of microorganisms which is comprised of a core set of bacterial genera that are commonly shared among individuals\textsuperscript{12,18,23}. Bacteria accumulate in biofilms on various surfaces such as teeth and soft tissues like the gums\textsuperscript{5,6,23}. The oral microbiome also exhibits long-term stability\textsuperscript{18} though this stability may be altered or influenced by a variety of factors including the environment of the host such as tooth loss as well as external factors like smoking and oral hygiene. Oral diseases, such as periodontitis and dental caries\textsuperscript{2,19,22,24}, as well as other diseases such as diabetes\textsuperscript{14,24} and cancer\textsuperscript{24} have been associated with the oral microbiome. Currently there is much interest in the association between human weight and the gut microbiome, however, the role of the human oral microbiome and human weight is not well understood.

The association of the human gut microbiota and obesity has been the focus of intense investigation recently. Studies have characterized the composition and diversity of humans and mice exhibiting lean and obese phenotypes\textsuperscript{20,7}. These studies have shown that obese individuals exhibit phyla-level changes in their gut microbiome as well as a decrease of microbial diversity\textsuperscript{7,13,21}. Many papers have focused on the ratio of \textit{Firmicutes} to \textit{Bacteroidetes} in obese and lean individuals with obese individuals displaying higher levels of \textit{Firmicutes} and lower levels of \textit{Bacteroidetes}\textsuperscript{13,21,30,33,36}. Other papers have contested these results\textsuperscript{7,35}. Fecal transplant experiments in mice also demonstrate a link between the gut microbiome and obesity\textsuperscript{21}.

The oral microbiome may serve as a biomarker for diseases related to the gut such as irritable bowel syndrome\textsuperscript{12,28} which suggests that the two microbiomes are correlated. Previous research has shown that both the human oral and gut microbiomes were correlated to biomarkers of atherosclerosis\textsuperscript{14} and the oral microbiome has been linked to irritable bowel syndrome\textsuperscript{26}. The
mouth is also a “gateway” to the rest of the body. Bacteria that accumulate on the different intraoral tissues slough off into the saliva. The microorganisms in the saliva may also move along contiguous epithelial surfaces within the mouth and into the gut.

**Significance**

Understanding the role of the microbiome in human health and disease is crucial for the development of disease diagnostics, therapeutics, and potentially personalized medicine. The oral microbiome has been implicated in oral and systemic diseases as particularly severe forms of oral diseases may cause systemic diseases. Periodontitis has been positively correlated with both diabetes and cardiovascular disease (CVD) suggesting that the oral microbiome may also be related to these systemic diseases. Since obesity has been linked to diabetes and CVD the oral microbiome may also be associated with weight. It has been previously suggested that oral bacteria may be related to the pathology of obesity, however this finding has been contested in other papers.

I investigated a possible correlation between the human oral microbiome and body mass index (BMI). I chose to use BMI because weight alone is not an adequate measurement of obesity since BMI also factors in an individual’s height. This research holds the promise that if successful, the oral microbiome may prove a useful surrogate to studies of the gut microbiome that are more complicated to carry out. My plan was to investigate the microbial diversity in relation to four BMI phenotypes: underweight, normal, overweight, and obese. I intended to look at this correlation in all samples with height and weight data available, controlling for age and genetics. I also wanted to examine discordant twin pairs as well as twins whose phenotype
changed over two time points. I was able to analyze the first group as well as the discordant monozygotic twins, but I did not have time to investigate the other twin groups.

**Materials and Methods:**

**Samples**

Saliva samples were received from the Center of Antisocial Drug Dependence as well as the Colorado Twin Registry. DNA from samples was isolated in the laboratory as described previously.42

**Sequencing of Bacteria**

The V4 region of the bacterial 16S small ribosomal subunit rDNA was sequenced via Illumina MiSeq as previously described.4 The 16S rDNA region has constant and variable regions that can be sequenced to determine OTUs present in each sample. Samples were prepared in triplicate and amplified via Polymerase Chain Reaction (PCR) with barcoded primers to amplify the 16S region of rDNA. Primers were designed as previously described1. The triplicates were pooled together after PCR to limit variation due to technique such as pipetting error and PCR machine variability. The samples were then run on a 2% e-gel for 20 minutes to verify that samples were amplified. PCR products were quantified using a picogreen assay according to manufacturer’s instructions. The PCR products from each barcoded subject were pooled in equimolar amounts (240ng/sample) by a graduate student, Brittany Demmitt, according to the Earth Microbiome1 protocol and then sent to the CU DNA Sequencing facility for sequencing on an Illumina MiSeq instrument.
Sample Selection

Data sets were assembled using RStudio (Version 0.99.441). The body mass index (BMI) was calculated according to a formula provided by the CDC as the weight in pounds divided by the height in inches multiplied by a constant \((\frac{\text{Lbs}}{\text{in}^2} \times 703)\). and subjects were divided into obese, underweight, overweight or normal classes according to age-appropriate target BMI\(^{15}\). Samples collected only via a scope collection method were used (n=2110). I took the complete set and removed any duplicate family members. In other words, I only kept one member of each family in order to control for genetics. For retention, I prioritized subjects who were classified, in order, as obese, underweight, overweight, and normal. This left me with 1020 subjects. Finally, each weight class was manipulated so that the mean age of each group was statistically similar to each other using a pair-wise t-test (n=976). I considered controlling for sex, but there’s no evidence that sex influences the oral microbiome\(^{18}\).

While doing this initial investigation, I noticed that there were a large number of discordant twins. I identified 24 discordant monozygotic twin pairs, meaning they had identical genomes but differing phenotypes. I chose only pairs where one cotwin was obese and the other was normal weight. The discordant phenotypes in these twins suggests that a factor other than genetics may be driving their weight classes. The microbial diversity of these discordant twins were compared to the microbial diversity of 16 pairs of normal weight concordant monozygotic cotwins and 16 pairs of obese concordant cotwins. There were considerably more normal concordant cotwins than obese cotwins, 16 pairs of normal concordant cotwins were randomly chosen out of the larger set.
**Statistical Analyses**

All statistical analyses were performed in QIIME\textsuperscript{29} (1.9.1) and RStudio\textsuperscript{37}. The sequences from the Illumina MiSeq run were categorized into different operational taxonomic units (OTU) which is a categorization of the microbes based solely on their DNA sequence similarities. Brittany Demmitt created a table that contained the counts of the OTUs in all samples sequenced. I matched the sample IDs of the data sets described above to the OTU table in order to find the frequency of OTUs of each individual. The filtered OTU table was then rarefied to a depth of 2500 reads/sample. OTUs were then filtered for abundance. The OTUs were filtered to those present in 50\% of the respective samples and were present a minimum of 5 times. Alpha diversity, which measures the microbial diversity within an individual, was performed using a PD whole tree metric. This metric is based on a phylogenetic tree and uses branch length to determine diversity\textsuperscript{39}. Beta diversity analysis, a measure of microbial diversity between individuals, using the Bray-Curtis metrics were used on all data sets. Bray-Curtis quantifies the compositional diversity between individuals based on OTU counts\textsuperscript{40}. Results were plotted on a principal component plot via emperor\textsuperscript{38}. A linear regression analysis was performed on the total sample set (n=976) of the principal components, independent variable, and the weight classes as well as the raw BMI scores, dependent variable. Other covariates (age, sex, MiSeq run, and sample collection year) were all controlled for in all regression analyses. An alpha diversity analysis was performed for both data sets as well, and a comparison of the frequency of different phyla-level OTUs was performed on each data set.
Results

Correlation between weight and the oral microbiome in unrelated individuals

I looked at the alpha diversity of the different weight classes using the PD whole tree metric. Unlike the results seen in the gut microbiome between lean and obese individuals the alpha diversity between all phenotypes was similar (Figure 1). This suggests the microbial diversity between all the weight classes is the same, and there are no significant differences in the oral microbiome between any of the four phenotypes.

![Alpha diversity curve showing the four different weight classes. Alpha diversity was measured using PD_whole_tree metrics. Red is normal, blue is obese, orange is overweight and green is underweight.](image)

**Figure 1: Alpha diversity of all samples**
Alpha diversity curve showing the four different weight classes. Alpha diversity was measured using PD_whole_tree metrics. Red is normal, blue is obese, orange is overweight and green is underweight.

I next performed a beta diversity analysis using the non-tree based Bray-Curtis metrics on all the unrelated individuals (n=976). Figure 2 represents a 2D plot of the beta diversity analysis of the first three principal components. No clear clustering of the different weight classes was observed from beta diversity, suggesting that the oral microbiome is not sufficiently distinct between obese and lean individuals. However, I noticed that many of the obese samples were
near principal coordinate 1 (PC1) which explains about 20% of the variation in the sample and principal coordinate 2 (PC2) which explains about 8% of the variation in the sample. I then performed a multivariate linear regression against PC1 and PC2 to see if either of these components were significantly related to the obese phenotype.

Figure 2: Beta diversity of all unrelated individuals
2D plots of the first three principal components of Bray-Curtis beta diversity analysis. Each point represents an individual and is plotted in a distance matrix relatively to the diversity of each other. Points are colored by their respective weight class. Red is normal, blue is obese, orange is overweight, and green is underweight.
Linear regression of PC1 and PC2

I performed a linear regression of multiple possible covariates against PC1, PC2 and PC3. I regressed out the weight classifications using normal weight as the reference phenotype against other variables (sex, age, MiSeq run, and sample collection year). I also performed a regression of the principal coordinates for the raw BMI numbers, regressing out the same covariates as above. For PC1 the beta coefficient of the obese weight class is -0.056 (Table 1) which was significant (P<0.001). On average individuals of the obese phenotype were 0.056 units below individual’s of normal phenotype along PC1. When only the raw BMI scores were used I retained significance (P<0.001). This suggests that the distribution of heavier individuals along PC1 which explains about 20% of the variation of the population was significantly different compared to normal weight individuals.

When weight classes were regressed against PC2, the estimate for obese compared to normal was -0.028 (Table 1) which was also significant (P<0.01). However, when the raw BMI numbers were regressed against PC2 significance was lost (P>0.1), suggesting that overall weight did not significantly contribute to plotting along PC2. Both phenotype and BMI scores were regressed against PC3, but no significance was found (Table 1).
<table>
<thead>
<tr>
<th></th>
<th>(β) Coefficient</th>
<th>P value</th>
<th>R^2=0.0816</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obese</td>
<td>-0.0560</td>
<td>0.0013**</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>-0.0200</td>
<td>0.1716</td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>0.0020</td>
<td>0.9137</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>-0.0040</td>
<td>0.0011**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(β) Coefficient</td>
<td>P value</td>
<td>R^2=0.0829</td>
</tr>
<tr>
<td>Obese</td>
<td>-0.0284</td>
<td>0.0148**</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>-0.0051</td>
<td>0.6052</td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>-0.0039</td>
<td>0.7869</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>-0.0012</td>
<td>0.1291</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(β) Coefficient</td>
<td>P value</td>
<td>R^2=0.0189</td>
</tr>
<tr>
<td>Obese</td>
<td>0.0007</td>
<td>0.4510</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>-0.0039</td>
<td>0.9420</td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>0.0037</td>
<td>0.6260</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>0.0236</td>
<td>0.3680</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Linear regression analysis of principal components 1 to 3.
Weight classifications and raw BMI scores were regressed against principal components 1 through 3. Multiple covariates were regressed out as well. Weight classes were all baseline to the normal phenotype. β coefficients are the average distance of the class as compared to the normal phenotype. *P<0.05, **P<0.01, ***P<-0.001.

Frequency of phyla-level OTUs
I next investigated whether there were any phyla-level changes of OTUs in the different weight classes of unrelated individuals. The fractional means were obtained via a Kruskal-Wallis nonparametric test41. While in the gut, the relative abundance of Firmicutes and Bacteroidetes have often been shown to have a correlation with human weight, in the oral microbiome Actinobacteria appeared to be significantly less abundant in obese individuals compared to the other weight classes (Table 2). To ensure that the difference between underweight and obese individuals was not driving this significance, I next compared only normal and obese individuals (Table 3). Again, I found that Actinobacteria was significantly less abundant in obese individuals compared to normal individuals (P <0.001). The relative abundance of Actinobacteria maintains its significance even with the Bonferroni and FDR corrected p values (Table 2 and 3).
<table>
<thead>
<tr>
<th>Phyla</th>
<th>Test-Statistic</th>
<th>P</th>
<th>FDR_P</th>
<th>Bonferroni_P</th>
<th>obese</th>
<th>normal</th>
<th>overweight</th>
<th>underweight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinobacteria</td>
<td>16.67</td>
<td>0.0008***</td>
<td>0.0041**</td>
<td>0.0041**</td>
<td>0.10</td>
<td>0.12</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td>Proteobacteria</td>
<td>5.94</td>
<td>0.1148</td>
<td>0.287</td>
<td>0.574</td>
<td>0.14</td>
<td>0.13</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td>Fusobacteria</td>
<td>2.8</td>
<td>0.4232</td>
<td>0.7053</td>
<td>1.000</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>0.87</td>
<td>0.8318</td>
<td>0.9301</td>
<td>1.000</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>0.45</td>
<td>0.9301</td>
<td>0.9301</td>
<td>1.000</td>
<td>0.56</td>
<td>0.56</td>
<td>0.56</td>
<td>0.55</td>
</tr>
</tbody>
</table>

**Table 2: Relative abundance of different phyla in different weight classes**
Table lists five core phyla of the human oral microbiome and their mean fractional abundance in each weight class. Analysis performed on all unrelated individuals (n=976) A Kruskal-Wallis analysis was done to calculate the p value. *P<0.5, **P<0.01, ***P<0.001

<table>
<thead>
<tr>
<th>Phyla</th>
<th>Test-Statistic</th>
<th>P</th>
<th>FDR_P</th>
<th>Bonferroni_P</th>
<th>obese</th>
<th>normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinobacteria</td>
<td>15.65</td>
<td>7.62E-05***</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.10</td>
<td>0.12</td>
</tr>
<tr>
<td>Proteobacteria</td>
<td>4.58</td>
<td>0.0323*</td>
<td>0.0808</td>
<td>0.1615</td>
<td>0.14</td>
<td>0.13</td>
</tr>
<tr>
<td>Fusobacteria</td>
<td>2.12</td>
<td>0.1450</td>
<td>0.2417</td>
<td>0.7251</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>0.64</td>
<td>0.4230</td>
<td>0.5288</td>
<td>1.0000</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>0.05</td>
<td>0.8306</td>
<td>0.8306</td>
<td>1.0000</td>
<td>0.56</td>
<td>0.56</td>
</tr>
</tbody>
</table>

**Table 3: Relative abundance of phyla in normal and obese individuals**
Table lists five core phyla of the human oral microbiome and mean fractional abundance. Only individuals classified as obese or normal were used (n=702). P value calculated from Kruskal-Wallis analysis. *P<0.5, **P<0.01, ***P<0.001

**Microbial diversity of discordant monozygotic twins**

Next I investigated whether the oral microbiome was associated with discordant weight phenotypes in monozygotic twins. If the oral microbiome plays a role in the discordant phenotypes, I would expect to see the discordant cotwins plotted more distantly from each other on a principal coordinate plot, and the concordant cotwins to be closer. I performed a beta diversity measure on normal weight and obese discordant twins (Figure 3A) in order to visualize the microbial diversity between twin pairs. The discordant twins generally were plotted with observable distance between them, however I did not have the time to do the work necessary to quantify the vectors. They were then compared to concordant normal and concordant obese twin pairs (Figure 3B-C). Visually, the concordant twins also appear far apart on the principal component plot indicating that the oral microbiome does not appear to associate with the weight
classes in monozygotic twins. Again, due to time the vectors were not quantified, so it is hard to determine if the discordant twin pairs were significantly more distal to each other than the concordant twin pairs.

Figure 3: Beta diversity PCoA of discordant and concordant twin pairs
Beta diversity was performed using non phylogenetic tree based Bray-Curtis metric. Vectors connect twin pairs to each other. (A) Discordant twin pairs. Blue indicates obese weight class, red indicates normal weight class. (B) Concordant normal weight class twin pairs. (C) Concordant obese weight class twin pairs.
Discussion

Although there has been plenty of evidence correlating the gut microbiome to human weight and obesity, the oral microbiome does not appear to have an association with human weight as has been previously discussed\(^3\). Although some oral diseases have been linked to systemic diseases that are related to obesity, I did not see a clear correlation of oral microbial diversity and BMI. This may be due to the fact that I did not have information on the oral health of these individuals, and they were not collected for oral diseases. Drug usage may have also affected the results as many drugs such as tobacco and alcohol may affect the oral microbiome.

I did find significance in the distribution of the obese weight class and BMI raw numbers along PC1, indicating that the obese phenotype in this population is clustering around this principal component that explains most of the variation in the population as a whole. However, looking at the beta and alpha diversity, the association between the oral microbiome and body weight ultimately appears unrelated in this population.

Similar to the gut microbiome, phyla-level changes were observed in the different weight classes. It has been previously proposed that higher levels of *Actinobacteria*\(^{20}\) in the gut microbiome is associated with obesity. However, in the oral microbiome, obesity was related to a decrease in the total fractional mean of *Actinobacteria* compared to all other phenotypes and the normal phenotype. *Actinobacteria* is one of the core phylum of the oral microbiome, but whether one species of bacteria from this phyla may act as a marker remains unknown. It has also been previously found that the presence of *Selenomonas noxia* in the oral microbiome was linked to overweight women\(^8\). The *selenomonas* genus was not present in this data set, so I could not look at it as a biomarker.
The results of the monozygotic twins remain unclear though the oral microbiome does not appear to be related to the discordant phenotype in twin pairs. Qualitatively, the beta diversity of the coordinate and discordant twin pairs appear similar. However, quantitatively it is unknown whether any differences are observed between the different twin groups.

While I believe my results do not show a clear correlation between the human oral microbiome and BMI, the correlation between the obese phenotype and PC1 as well as the phyla-level changes detected in the obese weight class are promising. A larger and better-defined sample would clarify this. Considering the possible link between periodontitis and other systemic diseases correlated to obesity, this sample population may not have revealed a clear correlation due to the lack of medical information available. These samples were not collected for periodontitis or other diseases that have a link to obesity. Further studies that control for the potential link between oral diseases and other systemic diseases may reveal an association with the oral microbiome and human weight.

**Acknowledgements**

I would like to sincerely thank my thesis advisor, Dr. Kenneth Krauter, for mentoring during my undergraduate career and for guiding me on this thesis project. I would also like to extend my gratitude to the other members of my thesis committee, Dr. Kevin Jones and Dr. Noah Fierer, for volunteering their time to participate in my thesis defense. I appreciate your time and commitment to be on my committee. And for her help and support, I would like to thank graduate student Brittany Demmitt for guiding me in this project and providing support and mentorship. Thank you all so much.
References


Appendix I

Verifying imputations

Along with this thesis project, I verified single nucleotide polymorphism (SNP) genotypes via sanger sequencing from imputations. An imputation is a prediction of a individual’s SNP genotype based on haplotype blocks. Each sample received a dosage score which expressed confidence for the imputation. I verified these imputation predictions by sequencing individual’s to assess their genotype for the SNP. I sequenced two SNPs, rs7444887 (n=53) and rs77693952 (n=13). Primers were designed through Primer Design and Tools provided by BiSearch (Figure 1A-B) and oligos were ordered from IDT Inc. Primers were optimized for polymerase chain reaction (PCR) before samples were run on the optimized PCR program and then run on a 1% agarose gel to ensure that all samples were amplified.

To prepare the samples for sequencing, they were purified via the QIAquick PCR purification kit according to manufacturer’s instructions (Qiagen). Purified samples were quantified for DNA concentration [ng/µl] on a Nanodrop2000 spectrometer. Samples were then diluted and prepared for sequencing by ACGT Inc. according to their low-cost option (LCO) single-pass DNA sequencing standards. Samples were diluted down to 3ng in 5 µl and the forward and reverse primers were custom mixed in at a concentration of 10pmol/µl. Prepared samples were then sent for sequencing at ACGT Inc. Examples of the results can be seen in Figure 1C-D. Of the 53 samples for rs7444887 all controls validated, and only 3 individuals with low dosage scores (<0.5) did not validate. For rs77693952 all controls validated, but 1 individual with a high dosage score (>0.99) did not validate.
Figure 1: Primers and sequencing results for imputation samples
(A) The forward and reverse primers designed for rs7444887. (B) The forward and reverse primers designed for rs77693952. (C) Example sequencing results for two different samples using rs7444887 primers. Reference allele is A/T. Highlighted peak is the SNP. Top picture shows a homozygous genotype call while bottom picture shows a heterozygous genotype call. (D) Example sequencing results for two different samples using rs77693952 primers. Reference allele is G/C. Highlighted peak is the SNP. Top picture shows a homozygous genotype call while bottom picture shows a heterozygous genotype call.