**TITLE:** Relationship of the Esophageal Microbiome and Tissue Gene Expression and Links to the Oral Microbiome: A Randomized Clinical Trial

**AUTHORS:** Medini K. Annavajhala, Michael May, Griselda Compres, Daniel E. Freedberg, Roseanna Graham, Stephania Stump, Jianwen Que, Tal Korem, Anne-Catrin Uhlemann, Julian A. Abrams

**SUPPLEMENTAL METHODS**

***Study Design***

 Subjects were excluded for any of the following reasons: use of proton pump inhibitors (PPI) or H2-receptor antagonists beginning ≤1 month prior to enrollment (standing acid suppression medication use >1 month prior to enrollment was permitted); history of upper gastrointestinal cancer or histologically proven Barrett’s esophagus; history of antireflux, bariatric, or other gastric or esophageal surgery; use of an antimicrobial mouth rinse within 1 month of enrollment; use of antibiotics or immunosuppressant medications within 3 months of enrollment; uncontrolled HIV or other immunosuppressed states; pregnant or breast feeding; inability to give informed consent.

 Data were collected on patient demographics, medical history, medications, and smoking history. A history of reflux symptoms was assessed using a modified version of the Mayo Gastro-Esophageal Reflux Questionnaire. [1] Subjects also completed a food frequency questionnaire derived from the National Health Interview Survey that is validated for the assessment of fat and fiber intake over the preceding 4 weeks. [2, 3]

 All subjects provided written informed consent. The study was approved by the Columbia University Institutional Review Board and was performed in accordance with all relevant guidelines and regulations. The study was registered with clinicaltrials.gov (NCT02513784).

***Adverse Effects***

 At the end of the treatment period patients in the treatment arm were asked about any side effects experienced while using chlorhexidine. Yellow tooth discoloration can occur with use of chlorhexidine. For subjects in the chlorhexidine arm, photographs were taken of the subjects’ teeth at baseline and after the treatment period. The images were reviewed by a dentist (RG), who was blinded to whether the images were pre- or post-treatment, and discoloration was scored from 0 (none) to 3 (severe yellowing/discoloration). A change in discoloration score was calculated for each subject, and worsening in discoloration was defined as a change in score ≥1.

***Biosample Collection***

At baseline (2 weeks prior to the endoscopy), saliva was collected and stored in Oragene DNA OG-500 collection kits (DNA Genotek). The adherent oral microbiome was sampled using oral swabs (Epicentre Catch-All Sample collection swabs) by broadly sampling 5 distinct sites (right and left buccal lining, tongue dorsum, hard palate, and superior labial frenulum). On the day of endoscopy, saliva and oral swabs were again collected. At the beginning of the upper endoscopy, the scope channel was flushed with 20 mL sterile water. The esophageal squamous microbiome was sampled with brushings (Endoscopy Cytology Brush, model G22174; Cook Medical), by passing the brush back and forth 10 times in each of four quadrants. Brushings were taken from the distal esophagus, 3 cm above the squamo-columnar junction, and avoiding any mucosal abnormalities. Brush tips were cut using sterile wire cutters and placed in sterile Eppendorf tubes. Biopsies were taken from the same region of the esophagus, avoiding abrasions from the brushings, and placed in Qiagen Allprotect®. All samples were then stored at -80˚ C.

***Microbiome Analyses***

Saliva was added to garnet bead tubes with lysis solution (PowerSoil kit, Qiagen), shaken, and centrifuged. DNA was extracted from the resulting supernatant using the Mo Bio UltraClean Protocol (Qiagen). DNA from oral swabs and esophageal and cardia brushings was isolated using the Mo Bio PowerSoil kit (Qiagen). The V3/V4 region of the 16S rRNA gene was amplified using established primers with Illumina Nextera adaptors. [4] Libraries were multiplexed using Illumina Nextera XT Index kits and normalized and pooled with 10% PhiX. Sequencing was performed on an Illumina MiSeq with a v3 kit and 600 cycles. Sequence data were uploaded to the NCBI Sequence Read Archive (SUB5135615).

16S rRNA sequences were processed using USEARCH v11.0.667 and R v3.3.0. USEARCH was used for quality-filtering, trimming, de-replication, and chimeric sequence filtering of FASTQ sequences. Clustering of operational taxonomic units (OTUs) was also performed in USEARCH, and taxonomic classifications were assigned using the ribosomal database project (RDP) database [5] through the SINTAX method [6] within USEARCH. OTUs with an average relative abundance of <0.005% across all samples were filtered using the *phyloseq* v1.19.1 [7] package via Bioconductor [8] in R. Microbiome diversity metrics (Shannon and Chao α-diversity; Bray-Curtis, unweighted, and weighted UniFrac -diversity) were calculated using *phyloseq*. Based on α-diversity rarefaction, we applied minimum cutoffs of aligned counts for inclusion in analysis (minimum 2,500 for esophageal and cardia swabs, 5,000 for oral swabs and saliva). This resulted in 40 saliva, 38 oral swab, 17 esophageal squamous, and 12 cardia samples with sufficient counts for analysis. Further analyses were not performed on the cardia in light of the low number of samples that were above the cutoff.

 To determine how closely related the oral and esophageal microbial communities were within individuals, within-individual overall community structure (β-diversity) and relative abundance of specific taxa were compared in the esophagus and oral swabs. First, pairwise UniFrac distances were compared between paired esophagus-oral swab samples within each individual to pairwise distances between esophagus-oral swab samples randomly paired across individuals, using the Wilcoxon rank-sum test. In order to determine whether the relative abundances of specific taxa in the esophagus were associated with the relative abundances for the same taxa in oral swabs, the subset of OTUs were selected that were present in both oral swabs and the esophagus for each individual, independently. Log-transformed relative abundance values were then plotted for each selected OTU in the esophagus versus paired oral swabs from the same individual or randomly paired oral swabs from different individuals. Subsequently, for each individual Spearman correlation rho values were calculated between relative abundance of each OTU in the esophagus vs. oral swab. Spearman rho values were then calculated for the correlation of relative abundances of each OTU in the esophagus with the relative abundance of that OTU in each of the randomly selected other patient’s oral swab. The Wilcoxon rank-sum test was then used to compare paired (i.e. within-individual) vs. “randomly” paired (i.e. across-individual) rho values between the esophagus and oral swabs. All of the above analyses were then repeated for paired esophagus-saliva samples.

Correlation of individual OTUs between sampling sites (esophagus-oral swab and esophagus-saliva) within individuals was also assessed. OTUs were identified with mean relative abundance ≥1% and also with non-zero reads in ≥50% across the three sampling sites. Pearson correlation coefficients were calculated, with statistical significance defined as p<0.05 adjusted for multiple comparisons using Bonferroni correction.

 Kruskal-Wallis tests were used to assess for significant differences in α-diversity (Shannon, Chao) after intervention based on treatment arm. In addition, for oral samples Kruskal-Wallis tests were used to assess within-individual change in α-diversity (after vs. before) comparing treatment arms. For β-diversity analyses, permutational ANOVA (PERMANOVA) using the *adonis* function in the R *vegan* package was used to assess differences in overall microbial community structure (Bray-Curtis, unweighted and weighted UniFrac). DESeq2[9] was used to test for differential abundance of specific OTUs from filtered count tables as described above, with comparisons made across and within treatment arms. Statistical significance was defined as an adjusted-p (FDR) <0.1 calculated using the Benjamini-Hochberg method.

***RNA Sequencing and Gene Expression Analyses***

 Esophageal tissue gene expression analyses were performed using RNA-Seq. Esophageal biopsies were homogenized in Qiazol on Tissuelyser II (Qiagen) at frequency 25/s for 2 minutes, twice. Total RNA was then purified using the miRNeasy micro kit (Qiagen) following the kit protocol. RNA was eluted with 15l RNase-free water. Quantitation was done by Nanodrop and Bioanalyzer. Poly-A pull-down was used to enrich mRNAs from total RNA samples (200ng-1ug per sample, RIN>8 required), and library preparation was performed by using the Illumina TruSeq RNA prep kit. Libraries were then sequenced using Illumina NovaSeq 6000. RTA (Illumina) was used for base calling and bcl2fastq2 (version 2.17) for converting BCL to fastq format, coupled with adaptor trimming. The reads were mapped to a reference genome (NCBI/build37.2) using STAR(2.5.2b) and featureCounts(v1.5.0-p3). The resulting count tables were imported into R for differential abundance analysis using DESeq2 after filtering for genes with at least two reads in at least two samples. [9] The sigclust2 package in R (<https://github.com/pkimes/sigclust2>) was used to test for significant hierarchical clustering of samples based on the variance-stabilized RNA-Seq count matrix. [10] The v.measure() function from the infotheo R package (<https://cran.r-project.org/web/packages/infotheo/index.html>) was then used to test for clustering of RNA-Seq profiles by treatment group, based on homogeneity (whether each cluster defined by sigclust2 contained only samples from a single treatment group), completeness (whether all samples from a treatment group belong to the same cluster), and v-measure scores (harmonic mean between homogeneity and completeness).

 To link the presence and relative abundance of specific microbiota with tissue gene expression data, Mantel tests were run on full OTU count tables and count tables for differentially expressed genes as identified above, across treatment groups at each site. The Mantel test explores whether samples that are similar based on one variable (e.g. tissue gene expression) are also similar based on a second variable of interest (e.g. microbiome β-diversity). Here, in order to test for correlation between gene expression profiles and microbiome community structure, we ran Mantel tests comparing a Euclidean distance matrix of variance-stabilized gene expression levels from RNA-Seq data with UniFrac, weighted UniFrac, and Bray-Curtis β-diversity pairwise distance matrix from esophageal microbiome data. Reported p-values are empirically determined across 999 permutations, and test the null hypothesis that there is no relationship between the two distance matrices (i.e. distances across objects in one matrix are not correlated with distances across these objects in a second matrix). In exploratory analyses, esophageal tissue gene expression was compared between individuals with high (above the median) and low (below the median) relative abundance of OTU29 using DESeq2 and by examining principal coordinate plots.

**REFERENCES**

1. Locke GR, Talley NJ, Weaver AL, et al. A new questionnaire for gastroesophageal reflux disease. Mayo Clin Proc 1994;69:539-47.

2. Thompson FE, Midthune D, Subar AF, et al. Performance of a short tool to assess dietary intakes of fruits and vegetables, percentage energy from fat and fibre. Public Health Nutr 2004;7:1097-105.

3. Thompson FE, Midthune D, Subar AF, et al. Dietary intake estimates in the National Health Interview Survey, 2000: methodology, results, and interpretation. J Am Diet Assoc 2005;105:352-63; quiz 487.

4. Klindworth A, Pruesse E, Schweer T, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 2013;41:e1.

5. Cole JR, Wang Q, Fish JA, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 2014;42:D633-42.

6. Edgar RC. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. 2016 [cited 2019 May 1]; Available from: <https://www.biorxiv.org/content/10.1101/074161v1>

7. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 2013;8:e61217.

8. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004;5:R80.

9. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

10. Kimes PK, Liu Y, Neil Hayes D, et al. Statistical significance for hierarchical clustering. Biometrics 2017;73:811-821.