

Mucosa-Associated Microbiota in Barrett's Esophagus, Dysplasia, and Esophageal Adenocarcinoma Differ Similarly Compared With Healthy Controls

Shajan Peter, MD¹, Amanda Pendergraft, PhD², William VanDerPol, MS³, C. Mel Wilcox, MD, MSPH¹, Kondal R. Kyanam Kabir Baig, MD¹, Casey Morrow, PhD⁴, Jacques Izard, PhD⁵ and Peter J. Mannon, MD¹

INTRODUCTION: Alterations in the composition of the human gut microbiome and its metabolites have been linked to gut epithelial neoplasia. We hypothesized that differences in mucosa-adherent Barrett's microbiota could link to risk factors, providing risk of progression to neoplasia.

METHODS: Paired biopsies from both diseased and nonaffected esophagus (as well as gastric cardia and gastric juice for comparison) from patients with intestinal metaplasia (n = 10), low grade dysplasia (n = 10), high grade dysplasia (n = 10), esophageal adenocarcinoma (n = 12), and controls (n = 10) were processed for mucosa-associated bacteria and analyzed by 16S ribosomal ribonucleic acid V4 gene DNA sequencing. Taxa composition was tested using a generalized linear model based on the negative binomial distribution and the log link functions of the R Bioconductor package edgeR.

RESULTS: The microbe composition of paired samples (disease vs nondisease) comparing normal esophagus with intestinal metaplasia, low grade dysplasia, high grade dysplasia, and adenocarcinoma showed significant decreases in the phylum Planctomycetes and the archaean phylum Crenarchaeota ($P < 0.05$, false discovery rate corrected) in diseased tissue compared with healthy controls and intrasample controls (gastric juice and unaffected mucosa). Genera Siphonobacter, Balneola, Nitrosopumilus, and Planctomyces were significantly decreased ($P < 0.05$, false discovery rate corrected), representing <10% of the entire genus community. These changes were unaffected by age, tobacco use, or sex for Crenarchaeota.

DISCUSSION: There are similar significant changes in bacterial genera in Barrett's esophageal mucosa, dysplasia, and adenocarcinoma compared with controls and inpatient unaffected esophagus. Further work will establish the biologic plausibility of these specific microbes' contributions to protection from or induction of esophageal epithelial dysplasia.

SUPPLEMENTARY MATERIAL accompanies this paper at <http://links.lww.com/CTG/A311>, <http://links.lww.com/CTG/A312>, <http://links.lww.com/CTG/A313>, <http://links.lww.com/CTG/A314>

Clinical and Translational Gastroenterology 2020;11:e00199. <https://doi.org/10.14309/ctg.000000000000199>

INTRODUCTION

Barrett's esophagus (BE) is characterized by the development of metaplastic columnar epithelium alongside the typical squamous epithelium found at the gastroesophageal junction. This metaplasia is believed to confer risk of neoplastic transformation, but delays in the diagnosis of esophageal adenocarcinoma (EAC), occurring only after the onset of symptoms for instance, reduce the survival significantly. Therefore, identifying the factors that

drive progression from Barrett's metaplasia to neoplasia could provide biomarkers that stratify risk and point out modifiable targets for improving outcomes.

Microbes and their metabolic byproducts have been linked to gut neoplasia including EAC presumed progressing from BE (1–5). Patients with reflux and Barrett's esophagus are reported to have an altered luminal microbe composition, with increased relative abundance of Gram-negative genera such as Fusobacterium,

¹Division of Gastroenterology and Hepatology, Department of Medicine, University of Alabama at Birmingham (UAB), Birmingham, Alabama, USA; ²Department of Biostatistics, UAB, Birmingham, Alabama, USA; ³Center for Clinical and Translational Science, University of Alabama at Birmingham, Birmingham, Alabama, USA; ⁴Department of Cell Developmental and Integrative Biology, UAB, Birmingham, Alabama, USA; ⁵Nebraska Food for Health Center, University of Nebraska, Lincoln, Nebraska, USA. **Correspondence:** Shajan Peter, MD. E-mail: shajan@uab.edu.

Received September 25, 2019; accepted June 12, 2020; published online August 3, 2020

© 2020 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of The American College of Gastroenterology

Bacteroides, and Veillonella, along with decreased relative abundance of Streptococcus species (2,3,6,7). The normal esophagus show relatively higher quantities of Streptococcus compared with patients with excess esophageal acid exposure (type 2) who harbor relatively more Gram-negative, anaerobic, and microaerophilic organisms. However, there is a lack of data on whether there is a progression of microbiome changes from healthy esophagus to gastroesophageal reflux disease (GERD) to BE to EAC (4,8). Our aim was to test for differences in the mucosa-associated microbiota that could be linked to dysplasia progression in the intestinal metaplasia (IM)-neoplasia pathway.

MATERIALS AND METHODS

Sample collection and preparation

Fifty-two patients were enrolled, which included controls patients (normal) (n = 10) and patients with IM (n = 10), low grade dysplasia (LGD) (n = 10), high grade dysplasia (HGD) (n = 10), and esophageal cancer (EAC) (n = 12). These patients were selected after endoscopic and histological confirmation of the respective BE spectrum. Normal patients with no evidence of BE were included as controls. After obtaining informed consent and before the endoscopy, demographics and clinical data were collected, including height, weight, waist, hip circumference, and smoking status.

Exclusion criteria included a history of acute onset of reflux symptoms, acute or chronic vomiting, or the use of antibiotics or nonsteroidal anti-inflammatory drugs within the preceding 4 weeks of collection of biopsy samples. History of use of antacid, H2 blocker, proton-pump inhibitor, bismuth-containing compounds, for at least 2 weeks before sample collection was documented but were not excluded. All subjects provided informed consent for obtaining study specimens, and the study was approved by the Clinical Research Ethics of the University of Alabama at Birmingham (IRB # IRB-140109003). A sterilized endoscope was introduced directly into the esophagus and stomach, avoiding the aspiration of fluid in the oral cavity and esophagus and the introduction of water into the stomach. The Barrett's segment length was measured as the distance between the distal end and the squamocolumnar junction and classified as per the Prague criteria for circumferential (C) and maximal (M) dimensions (5). The presence of a hiatal hernia was defined while measuring the distance from the diaphragmatic hiatus and the distal end of the tubular esophagus.

Biopsies (2 from each site) of the Barrett's/EAC and visibly unaffected esophagus (3 cm cephalad) along with gastric secretions and gastric cardia biopsies were obtained during medically indicated endoscopy. In the stomach, up to 1 mL of gastric fluid was aspirated for analysis in a sterile container as a comparator control to represent swallowed and refluxed microbes. Esophageal samples were collected first (Barrett's then squamous) before stomach samples (cardia and gastric juice) to minimize contamination of the esophagus with stomach bacteria. Biopsy forceps were rinsed in sterile water between samples. These collected samples were placed immediately in Xpedition lysis and stabilization buffer (Zymo Research, CA) and frozen at -80°C until DNA extraction.

Microbiota analysis

Mucosa-associated microbes from paired samples were analyzed for significant differences in bacterial communities between sites (e.g., disease vs nondisease) and among diagnostic strata representing progression of disease from IM to LGD to HGD to adenocarcinoma. DNA extraction using bead beating and purification were performed using the QIAamp DNA Mini Kit.

Amplicon library for bidirectional (2×250 bp) sequencing on the Illumina MiSeq platform was constructed using universal primers 515f, 5'-GTGCCAGCMGCCGCGGTAA-3' and 806r, 5'-GGACTACHVGGGTWTCTAAT-3' targeted across 16S ribosomal ribonucleic acid (rRNA) V4 gene hypervariable regions. Multiplex polymerase chain reaction was used with bar codes for 96 samples including, in some runs, negative (no added DNA) controls. PCR products were resolved on agarose gels; DNA was isolated and purified using Qiagen kits and was then quantitated. The products were sequenced on the MiSeq platform, a single flow cell, single lane instrument that can generate approximately 9 Gb of sequence data from our paired end 250bp run (6). An average of 37,411 reads per sample were obtained after quality control steps. The sequence count table was rarefied to 18,483 sequences per sample to minimize the effects of uneven sampling.

Sequence curation and annotation

For the analysis of microbiome data, we have established an analytical pipeline based on the latest version of the Quantitative Insights into Microbial Ecology (QIIME) tool suite (7). The first step in our analysis was to assess the quality of the raw data using fast alignment quality check and then filter out low quality data using the fast alignment X windows terminal emulator toolset. A combination of tools within the QIIME suite were used for clustering reads into operational taxonomic units (OTUs) (uclust), taxa assignment (Ribosomal Database Project) classifier using the Greengenes 16S rDNA database (8–10) and as necessary, alignment and phylogenetic inference using PyNASTPython nearest Alignment Space Termination (11) and Fasttree (12). To expedite sample processing and reporting, we have built an automated pipeline, Quick Microbiome data analysis pipeline (6).

16S rRNA gene sequence data were categorized as eukarya, archaea, mitochondria, or chloroplast, and unknown kingdoms were removed. The final sequences were clustered into OTUs using a 97% identity cutoff using an average neighbor clustering algorithm. Species-level identifications for OTUs of interest were determined through searches within the collection of 16S rRNA gene sequences curated by the National Center for Biotechnological Information (database built on June 16, 2016) using default megablast parameters and minimum e-values of 1.0×10^{-5} . Putative bacterial species were reported for any OTUs with predominant sequences that had greater than 99% sequence identities with those found in the National Center for Biotechnological Information 16S rRNA gene sequence database.

Selection of differentially abundant OTUs

Before differential abundance analysis, we identified potential confounding factors that might influence results. Categorical variables were assessed using the Fisher exact test; continuous variables were assessed using the Kruskal-Wallis (KW) test or one-way analysis of variance depending on the assumption of normality as judged by the Shapiro-Wilk test (Table 1).

We performed differential abundance analysis using generalized linear models based on the negative binomial distribution and the log link function as incorporated in the R Bioconductor package *edgeR* to assess the phylum- and genus-level OTUs found to have a mean relative abundance $>0.01\%$ (13). Moderated tagwise dispersion was estimated for each OTU before model fitting through an empirical Bayesian framework with previous degrees of freedom equated to 10 shrinking toward common dispersion. Trimmed mean of M-values normalization was used

Table 1. Patient demographics and clinical characteristics

	Description	Controls	IM	LGD	HGD	Esophageal cancer
n		12	9	12	10	10
Age, mean (SD)		52.2 (11.0)	57.6 (10.1)	57.1 (13.6)	67.2 (13.7)	64.9 (9.0)
Race n (%)	AA	4 (33.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	W	8 (66.7)	9 (100.0)	12 (100.0)	10 (100.0)	10 (100.0)
Sex n (%)	F	6 (50.0)	1 (11.1)	4 (33.3)	0 (0.0)	1 (10.0)
	M	6 (50.0)	8 (88.9)	8 (66.7)	10 (100.0)	9 (90.0)
BMI, mean (SD)		30.0 (7.0)	29.2 (4.9)	31.1 (5.8)	27.7 (2.6)	28.1 (5.1)
Tobacco use, n (%)	Current	1 (8.3)	1 (11.1)	0 (0.0)	5 (55.6)	4 (40.0)
	Former	1 (8.3)	4 (44.4)	1 (9.1)	4 (44.4)	2 (20.0)
	Never	10 (83.3)	4 (44.4)	10 (90.9)	0 (0.0)	4 (40.0)
Gastrointestinal reflux disease, n (%)	No	3 (25.0)	0 (0.0)	3 (25.0)	3 (33.3)	2 (20.0)
	Yes	9 (75.0)	9 (100.0)	9 (75.0)	6 (66.7)	8 (80.0)
Hypertension, n (%)	No	7 (58.3)	4 (44.4)	6 (50.0)	5 (55.6)	6 (60.0)
	Yes	5 (41.7)	5 (55.6)	6 (50.0)	4 (44.4)	4 (40.0)
Diabetes mellitus, n (%)	No	8 (66.7)	9 (100.0)	8 (66.7)	8 (88.9)	5 (50.0)
	Yes	4 (33.3)	0 (0.0)	4 (33.3)	1 (11.1)	5 (50.0)
Hyperlipidemia, n (%)	No	6 (50.0)	5 (55.6)	7 (58.3)	5 (55.6)	4 (40.0)
	Yes	6 (50.0)	4 (44.4)	5 (41.7)	4 (44.4)	6 (60.0)
Coronary artery disease, n (%)	No	11 (91.7)	8 (88.9)	11 (91.7)	5 (55.6)	9 (90.0)
	Yes	1 (8.3)	1 (11.1)	1 (8.3)	4 (44.4)	1 (10.0)
	Yes	0 (0.0)	0 (0.0)	0 (0.0)	1 (11.1)	9 (90.0)
Hiatal hernia, n (%)	No	7 (58.3)	5 (55.6)	7 (58.3)	6 (66.7)	9 (90.0)
	Yes	5 (41.7)	4 (44.4)	5 (41.7)	3 (33.3)	1 (10.0)
Prague_Criteria_C, median, cm (IQR)		NA	2.0 (0.0, 7.0)	4.5 (2.5, 8.5)	5.0 (3.0, 6.8)	4.0 (3.0, 7.0)
Prague_Criteria_M, median, cm (IQR)		NA	3.0 (1.0, 8.0)	6.0 (3.5, 10.3)	6.0 (4.0, 7.8)	5.0 (4.5, 8.0)
Length of segment (%)	Long (>3 cm)	NA	3 (37.5)	9 (75.0)	8 (88.9)	5 (83.3)
	Short (<3 cm)	NA	5 (62.5)	3 (25.0)	1 (11.1)	1 (16.7)

HGD, high grade dysplasia; IM, intestinal metaplasia; IQR, interquartile range; LGD, low grade dysplasia.

to satisfy the assumption that most OTUs are not differentially abundant between incremental fluctuations in disease progression. To account for the multiplicity of samples drawn from individuals, unique patient identifiers were incorporated on the right-hand side of design matrices capturing within-individual variability. Diagnosis was ranked from 0 (normal) to 4 (EAC) to generate an independent variable capable of capturing the effect disease progression. Age, sex, and smoking history were incorporated into the models so to explicitly adjust for dependencies between the confounders determined to significantly discriminate groups.

To compensate for multiple testing, all *P*-values were updated with a Benjamini-Hochberg (false discovery rate) adjustment to yield *q*-values. In the event, a *q*-value was found to be less than the threshold of 0.10, pairwise comparisons across diagnosis were considered to emphasize the contrasts supporting the omnibus differences in differential abundance. All differential abundance analyses were performed using R Project for Statistical Computing v 3.5.0.

Data analysis

The Mann-Whitney *U* test was performed to compare the variables of 2 sample groups. Multiple group comparisons were made using the KW test. The Fisher exact test performed on categorical variables. *P* < 0.05 was taken as statistical significance. The *P* values obtained were adjusted for multiple comparisons by the false discovery rate method. The corresponding *q* < 0.05 was taken as statistical significance. Tests were performed using R Project for Statistical Computing V.3.3.1.

A model based on the Dirichlet-multinomial distribution and the log link function as incorporated in the R Bioconductor package *Human Microbiome Project* was used to test for variability in the composition of phylum- and genus-level OTUs found to have a relative abundance > 0.01% (14). This model contrasts OTUs via a generalized Wald-type test statistic so to identify groups or sites in which the vectors of estimated mean relative abundances not equal.

Visualizations of alpha diversity (diversity within the sample) using different metrics including Shannon, Chao, and Simpson

Table 2. 16S sequence library size among patient groups

	Cancer	HGD	IM	LGD	Control	Shapiro-wilks (site)	P	Test
Abnormal (1)								
Mean (SD)	17,112.9 (12,716.0)	17,368.2 (11,228.0)	17,709.0 (7,259.3)	13,153.0 (6,759.9)	NA	0.0008	0.57	KW
Median [IQR]	13,686.5 [6,692–49,547]	17,400.5 [3,524–38,901]	16,203 [7,633–30,140]	11,372 [6,470–29,523]	NA			
Abnormal (2)								
Mean (SD)	13,518.8 (5,783.0)	18,782.9 (9,633.0)	13,737.3 (9,382.3)	14,314.5 (6,983.0)	NA	0.2922	0.46	ANOVA
Median [IQR]	13,766 [4,269–24,127]	17,212 [6,437–38,808]	11,320 [3,174–28,900]	14,962 [2,057–22,297]	NA			
Cardia								
Mean (SD)	12,112.9 (8,038.3)	16,345.8 (17,628.5)	19,593.2 (13,726.4)	11,402.9 (7,756.0)	16,106.0 (9,417.8)	<0.0001	0.69	KW
Median [IQR]	12,692 [3,723–25,162]	9,688 [4,366–67,934]	17,612.5 [3,831–40,438]	9,319 [1,347–25,060]	15,104 [3,392–31,834]			
Gastric juice								
Mean (SD)	17,271.9 (12,365.1)	16,476.2 (8,681.2)	15,478.9 (10,562.9)	21,690.3 (7,339.2)	15,323.4 (12,126.3)	0.0582	0.67	ANOVA
Median [IQR]	15,415.5 [4,435–40,057]	17,978 [2,668–31,670]	14,392.5 [3,190–37,279]	20,953 [12,183–31,714]	12,145.5 [1,986–43,029]			
Normal (1)								
Mean (SD)	18,016.8 (10,885.7)	15,106.8 (7,932.8)	10,824.4 (4,791.6)	14,467.7 (9,122.2)	9,854.8 (7,386.3)	0.0113	0.27	KW
Median [IQR]	16,737.5 [5,463–33,210]	12,995 [4,596–27,492]	10,119 [5,082–19,923]	12,954 [72–27,480]	7,581.5 [1,005–26,679]			
Normal (2)								
Mean (SD)	14,072.3 (10,875.8)	19,644.3 (7,676.9)	17,988.4 (10,022.8)	16,877.1 (5,411.3)	12,904.7 (4,881.7)	0.1561	0.29	ANOVA
	8,415 [2,126–33,698]	17,469 [6,982–34,215]	19,084 [4,714–36,203]	17,186 [6,231–24,615]	11,814 [6,504–22,512]			
Shapiro-Wilks (group)	0.0003	0.0017	0.5061	0.0013	<0.0001			
P	0.7854	0.5452	0.1195	0.5396	0.256			
Test	KW	KW	ANOVA	KW	KW			
Shapiro-Wilks test of normality (omnibus)	<0.0001							
KW P-value for group diff	0.3323							
Wilcoxon rank sum P-values for pairwise comparisons and group differences (BH adjustment)								
	Cancer		Controls		HGD		IM	
Controls		0.72		—		—		
HGD		0.72		0.99		—		—
IM		0.72		0.99		0.99		—
LGD		0.72		0.39		0.44		0.39
KW P-value for sample site diff		0.382						

Table 2. (continued)

Wilcoxon rank sum <i>P</i> -values for pairwise comparisons for sample site diff (BH adjustment)					
	ABNL (1)	ABNL (2)	Cardia	Gastric juice	Normal (1)
ABNL (2)	0.87	—	—	—	—
Cardia	0.68	0.76	—	—	—
Gastric juice	0.87	0.76	0.68	—	—
Normal (1)	0.68	0.68	0.87	0.68	—
Normal (2)	0.87	0.76	0.68	0.94	0.68

ABNL, abnormal; HGD, high grade dysplasia; IM, intestinal metaplasia; KW, Kruskal-Wallis; IQR, interquartile range; LGD, low grade dysplasia.

indices and beta diversity (diversity between samples) using Bray Curtis and UniFrac are also presented. Samples were grouped by user defined variables and significant differences between groups are determined by performing a per mutational multivariate analysis of variance test on each of the beta diversity indices. Furthermore, a KW test is performed to identify key taxa whose changes in relative abundances between groups are playing a significant role in driving the overall group differences. These statistical tests were performed using tools within the QIIME package (7).

RESULTS

Patient demographics and clinical characteristics

The BE and cancer patients were predominantly white men, although the control group was split evenly between men and women, providing some samples in the control group to assess the effect of sex on the esophageal microbiota (Table 1). Patients with HGD and esophageal cancer tended to be older on average. There were no significant differences among groups by body mass index, the presence of GERD, hypertension, diabetes, coronary artery disease, or hyperlipidemia. However, the absence of tobacco use was significantly different among the groups, being high in controls (83% never smokers) and low in certain dysplasia and cancer groups (0%–40% never smokers), although LGD group subjects had a 91% rate of never smoker status. The characteristics of the BE (prague criteria-segment length) also were not significantly different between the groups (although these data may not be normally distributed). Given some significant differences in exposures such as tobacco and age among the groups, these variable effects on esophageal microbiota composition were tested separately (Figure 4).

Microbiota composition between biopsy sites within individuals and within groups

Because we had replicate samples of unaffected and diseased mucosa from each subject, we were able to measure the inherent variability in our data because of our methodology and design. The variability of 16S sequencing data was compared between replicate biopsies from individual subjects. There were no significant differences in library size among biopsies from sample sites within patient groups (Table 2) or specific to sample sites themselves (unaffected vs diseased within individual patients). Of note, the overall relatively low library sizes (compared with stool samples for instance) reflect the smaller numbers of mucosa-associated microbes that are

recovered from individual endoscopic biopsies and small volumes of gastric secretions.

The composition of the mucosa-associated microbiota based on 16S sequencing in paired unaffected biopsies within the same patient was stable at the phylum level within the same histologic groups. For instance, across phyla, 50% of absolute count differences were less than or equal to 0.0% and 75% of absolute differences were less than or equal to 31.0. Based on this finding of low variability between replicate samples, we combined paired unaffected and diseased samples within an individual patient rather than analyze each of the replicate biopsy samples separately.

Overall measures of diversity, including alpha diversity (including Shannon and inverse Simpson index and Chao1 estimates) and beta diversity (Bray-Curtis) showed no significant differences between sites or groups (see Figures 1 and 2, Supplementary Digital Content 1 and 2, <http://links.lww.com/CTG/A311> and <http://links.lww.com/CTG/A312>). At the phylum level, the bacterial composition was similar to that reported for the oral cavity with the largest mean percentages belonging to Firmicutes (47.81%), Proteobacteria (20.67%), Bacteroidetes (16.93%), Actinobacteria (5.57%), and Fusobacteria (4.76%) (Figure 1). At the genus level, upper gastrointestinal tract genera like *Lactobacillus*, *Streptococcus*, and *Prevotella* were dominant but the high frequency of *Tissierella*, an isolate usually associated with fecal sources, was unexpected yet consistent throughout samples including healthy controls. The largest mean percentages by composition were *Tissierella soehngenia* (16.67%), *Lactobacillus* (7.15%), *Streptococcus* (7.27%), *Acinetobacter* (5.80%), and *Prevotella* (5.24%) (Figure 2).

Microbiota composition differences among control, Barrett's metaplasia, and esophageal cancer biopsies

Analysis of OTUs obtained by sequencing of 16S rRNA genes from controls to IM to dysplasia and finally cancer provided a number of phyla and genera, which were down or up, regulated among the histologic groups as per the disease gradient. Using a cutoff of 0.1 for the *q*-value assigned to results with *P*-values <0.05, there were 2 phyla containing 3 genera, which had continuous, unidirectional negative-fold changes from IM, LGD, HGD, and EAC compared with the healthy control group. Phylum Planctomycetes (containing genus Planctomyces, neither Gram-positive or negative, lacking cell wall peptidoglycan and displaying anaerobic metabolism) was decreased in all groups compared with controls but significantly so in HGD and cancer.

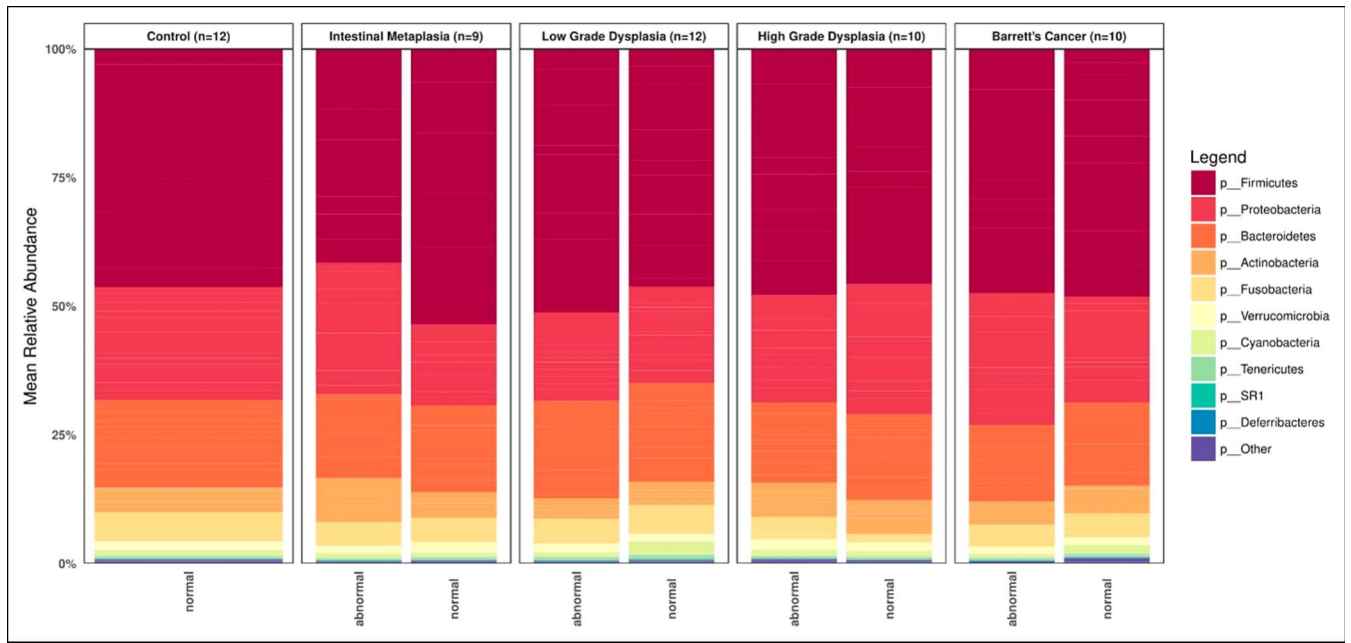


Figure 1. Compositional bar plots of mean relative abundance at the phylum-level of taxonomy. Normal and abnormal tissue samples are presented from patients diagnosed with Barrett’s cancer, high grade dysplasia, low grade dysplasia, and intestinal metaplasia along with normal tissue samples from controls. The largest percentages of observations on average were determined to be Firmicutes (47.81%), Proteobacteria (20.67%), Bacteroidetes (16.93%), Actinobacteria (5.57%), and Fusobacteria (4.76%).

Phylum Crenarchaeota (containing genus *Nitrosopumilus*) was similarly decreased. Genus *Balneola* (an aerobic Gram-negative bacterium in the family Crenotrichaceae but in the phylum Bacteroidetes) was also found to be decreased across disease groups, significantly in the HGD samples. When looking at

relative abundance of microbes among groups, the trend was similar to that seen in fold changes of absolute numbers of microbes (see Figure 3 and Table 1, Supplementary Digital Contents 3 and 4, <http://links.lww.com/CTG/A313>, <http://links.lww.com/CTG/A314>). These same fold changes in microbial phyla and

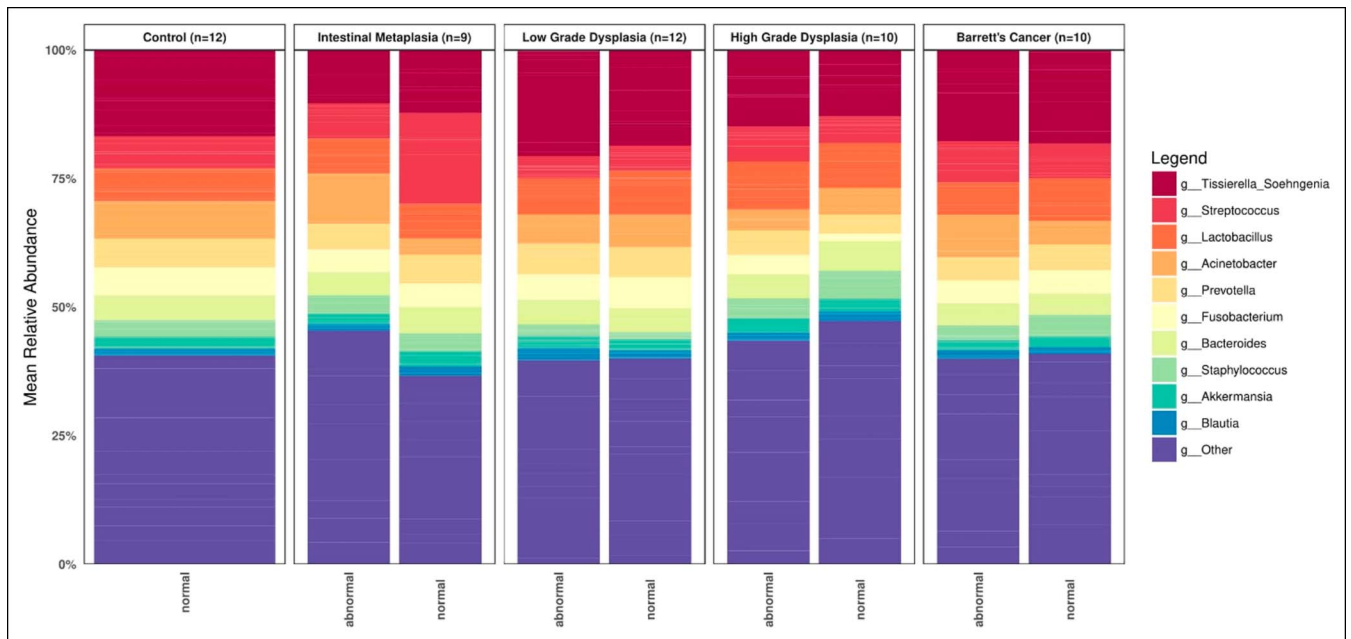


Figure 2. Compositional bar plots of mean relative abundance at the genus-level of taxonomy. Normal and abnormal tissue samples are presented from patients diagnosed with Barrett’s cancer, high grade dysplasia, low grade dysplasia, and intestinal metaplasia along with normal tissue samples from controls. The largest percentages of observations on average were determined to be *Tissierella Soehngenia* (16.67%), *Lactobacillus* (7.15%), *Streptococcus* (7.27%), *Acinetobacter* (5.80%), and *Prevotella* (5.24%).

genera were also seen when similar tests were performed between the diseased and unaffected esophageal biopsies within the same patient; in addition, these changes persisted when disease biopsies were compared with gastric juice samples and gastric cardia biopsies (data not shown) (Figure 3).

Effects of smoking, disease status, and sex on mucosa-associated microbiota in controls and Barrett's metaplasia

We adjusted for potential confounders and factors believed to influence the progression from metaplasia to neoplasia including age, sex, and smoking history. Figure 4 shows the effects of individual variable stratification on the changes in the phylum-level microbial compositions from controls to EAC. After stratifying the data according to tobacco use, decreases in both phyla were again noted at the *P* and *q* value cutoff (and significant decreases in multiple other phyla are also noted) while adjusting for age somewhat decreases the significance levels of their differences at the *q*-value cutoff. Phylum Planctomycetes seems to be more sensitive to stratification by age than phylum Crenarchaeota because only the latter phylum was still significantly decreased compared with control samples (data not shown for sex). The significant decrease in the 3 genera (Balneola, Nitrosopumilus, and Planctomyces) was also unaffected by adjusting for tobacco use history (data not shown).

DISCUSSION

The most widely accepted risk factors for BE and EAC are chronic gastroesophageal reflux and associated conditions, such as hiatal hernia or esophagitis (15). Chronic reflux can induce esophageal mucosal damage that may create inflammatory microenvironments that promote dysplasia and carcinogenesis. This microenvironment is further influenced by various factors such as diet, race, body mass index, and medications (16).

Interestingly, the increasing incidence of BE and EAC has been associated with improved sanitation and the consequent decrease in gastric *Helicobacter pylori*, mainly in the West (17,18). Although the inverse correlation between EAC incidence and *H. pylori* infection may relate strictly to decreased acid output (hypochlorhydria) and reflux, *H. pylori* colonization also significantly alters the native gastric and esophageal flora. Moreover, the bacterial community of the normal esophagus changes in patients with reflux-related disorders including BE and exposure to proton-pump inhibitor therapy (19). We did not document the *H. pylori* status before enrollment in the study. The composition of the human gut microbiota changes with, and may influence, many human diseases, including gastrointestinal cancer development. The esophagus contains a complex population of resident microbes that changes in patients with reflux-related disorders or BE compared with healthy individuals (18). These findings supported a proposal to group the microbiome composition into 2 categories labeled as "type I and type II". Type I, the normal esophagus, was populated with high quantities of Streptococcus compared with patients with excess esophageal acid exposure, type II, which contained a higher proportion of Gram-negative, anaerobic, and microaerophilic organisms. This information suggests that Gram-positive bacteria dominate the healthy esophageal microbiome, whereas GERD produces a shift in the microbiome to increased numbers of Gram-negative and anaerobic bacteria (20). A study using the Cytosponge technique including other tissue samples to evaluate the microbial profile at different stages of Barrett's carcinogenesis showed decreased microbial diversity and altered community composition associated with disease: decreases in Gram-negative (*Veillonella*, *Megasphaera*, and *Campylobacter*) and Gram-positive (*Granulicatella*, *Atopobium*, *Actinomyces*, and *Solobacterium*) taxa with increased *Lactobacillus fermentum* (2). Analysis of microbial samples

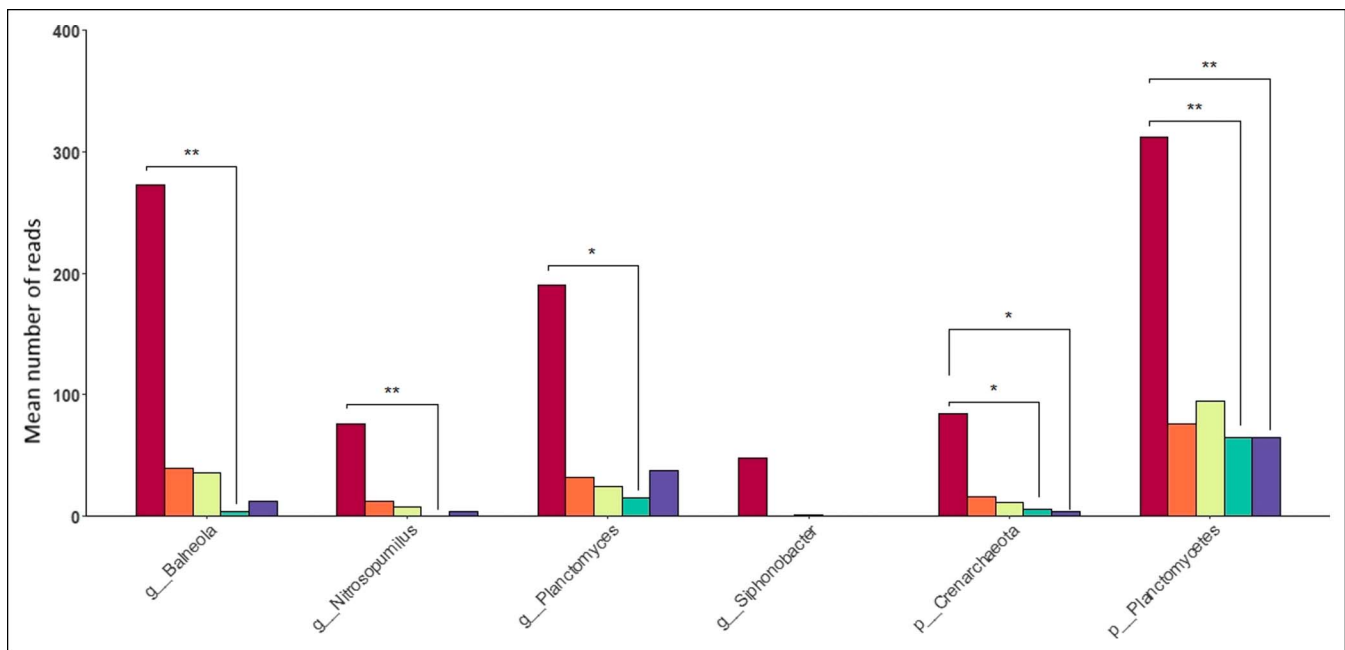


Figure 3. Comparison of mean number of reads observed in samples of the OTUs by histologic diagnosis found across increasing histologic severity based on a *q*-value <0.10. The diagnosis groups are represented as control (red), intestinal metaplasia (orange), low grade dysplasia (yellow), high grade dysplasia (green), and EAC (purple). **q*-value <0.10, ***q*-value <0.05, ****q*-value <0.01. EAC, esophageal adenocarcinoma; OTU, operational taxonomic unit.

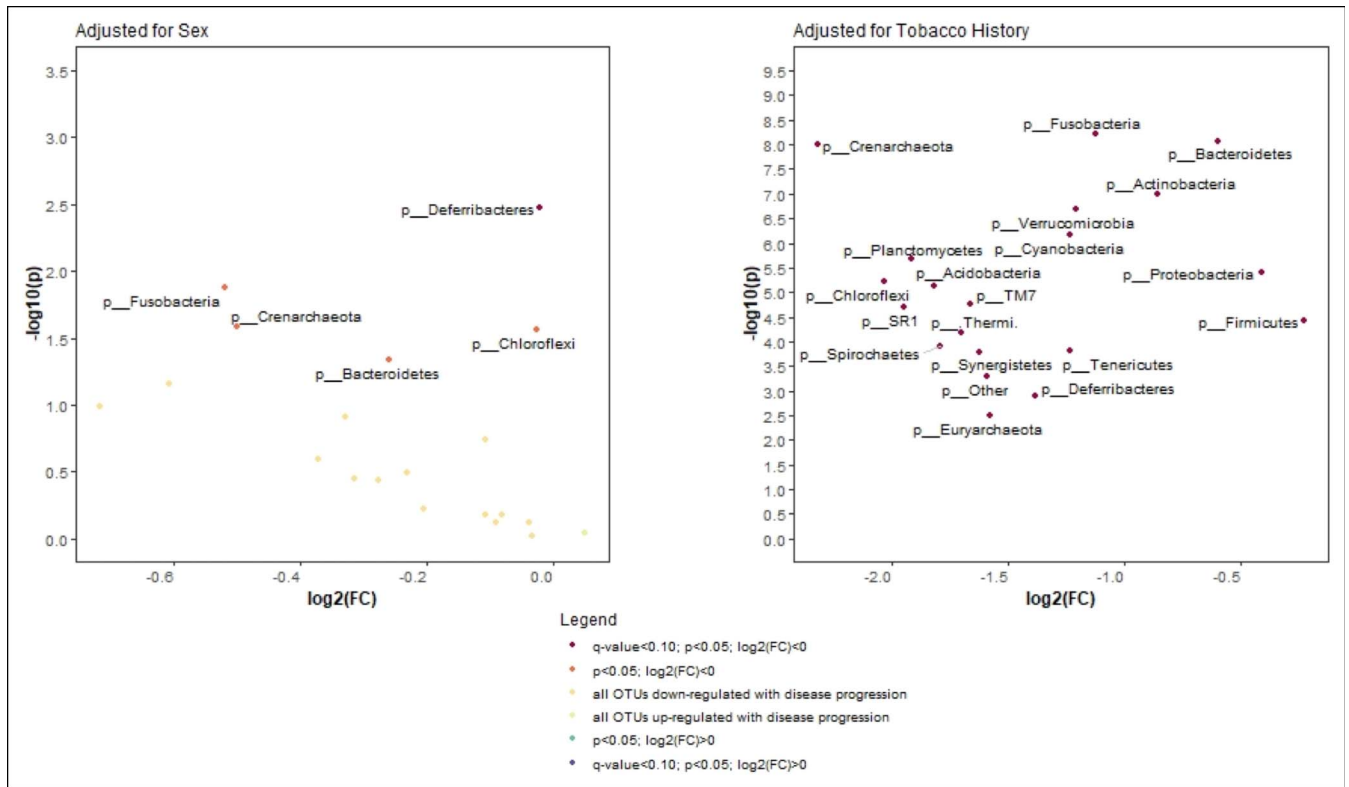


Figure 4. Downregulation of taxa (by log fold-change) across disease progression at phylum-levels of taxonomy phylum-level taxa adjusted for sex (left), tobacco use (right). log2FC, log2 fold changes; OTU, operational taxonomic unit.

obtained using the Cytosponge seems to differ from biopsy- or brush-obtained samples and may represent background by both gastric and oral bacteria. However, these studies have not shown significant differences in the microbiota from samples of BE or EAC compared with nearby unaffected esophageal tissue (21,25). Our data show significant decreases in a small group of organisms in Barrett's esophageal metaplasia, but these changes do not intensify with progression along the path of worsening dysphagia and frank neoplasia. There have been several proposed mechanisms by which the esophageal microbiota could affect the progression of IM and the subsequent development of dysplasia to EAC. The levels of acid exposure could have an indirect influence on microbial composition of the esophagus because different bacteria thrive in unique, niche microenvironments. On the other hand, an individual's unique microbiome and the associated microbial milieu of the lower esophagus may predispose certain people to develop esophagitis or IM in the presence of excessive levels of refluxed stomach acid (3,20). Microbial-derived metabolic products influencing the host immunologic response has also been proposed as a mechanism by which the microbial environment exerts a proinflammatory effect on the esophageal lining. Environmental or dietary factors related to obesity can result in change in the microbiome of esophagus affecting the stem cell biology and thereby susceptibility to carcinogenesis (22). The paradigm of microbial-promoting gastrointestinal tumorigenesis is focused on microbial-epithelial-immune interaction and the effects of microbial metabolism producing protumor effects (21).

Toll-like receptors (TLRs) mediate some of the interactions between the immune system and the esophageal microbiome potentially driving a chronic inflammatory response. (20,23)

Increased expression of TLRs 1–3,6,7,9 may act to mediate pathogenic microbial interaction of the innate immune systems as seen in a rat model for EAC, whereas TLRs 1,2,4,6,9 (particularly TLR-4) have been found to be overexpressed in human EAC samples (23–29). Flagellated bacteria could activate TLR5. Bacterial DNA has also found to promote invasion in *in vitro* models. The release of biochemical mediators that affect the tonicity of the lower esophageal sphincter is a proposed mechanism by which the esophageal microbiome can exert a direct effect on the development of BE. Bacterial invasion of the esophagus and gut in general could induce an inflammatory response that involves up-regulation of cytokines, such as IL-17 and IL-23, which can promote tumor development (20). Lipopolysaccharide is found in the outer membrane of Gram-negative bacteria and has been shown to be increased in the lower esophagus of patients with esophagitis and BE (type II microbiome) (30). Lipopolysaccharide can upregulate gene expression via the TLR-4 and NF- κ B pathways leading to increased expression of inflammatory cytokines (31). This can also indirectly exacerbate reflux via induction of nitric oxide synthesis (for relaxation of the lower esophageal sphincter) and production of cyclooxygenase 2 (for delayed gastric emptying) (32–34).

Our data have identified a small number of genera and phyla that are consistently associated with decreased absolute number of organisms when comparing normal control tissue with IM, dysplasia, and frank neoplasia. For instance, there was a decrease in frequency of phylum Planctomycetes/genus Planctomyces (an environmentally derived commensal) and phylum Crenarchaeota (containing genus Nitrosopumilus, ammonia oxidizing archaean) was similarly decreased. Planctomycetes microbes

have been isolated at low levels in the human gut microbiota in a few studies. The factors driving this diversity is unknown and could be related to a resistance to antibiotics such peptidoglycan inhibitors (35). Interestingly, this phylum seemed more prevalent in noncancer specimens compared with those with cancer (rectum) signifying an inverse relationship to tumorigenesis, consistent with its apparent loss in the samples of BE to EAC perhaps (36). The finding of Crenarchaeota in the human gut microbiome has also been reported, especially the species *Sulfolobus*, an obligate aerobe which can reduce sulfur particularly under acidic conditions (37–40); the anaerobic methane-producing Crenarchaeota may also be included in this group because they have been found in the oral cavity from subgingival plaques. The oxidation of ammonia to nitrite and nitrate is also of potential interest because these compounds may have pleiotropic effects in carcinogenesis so that a loss could be proposed to introduce an imbalance in locally regulated inhibitory compounds. Phylum Planctomycetes seems to be more sensitive to stratification by age and sex than phylum Crenarchaeota. We did not notice a difference by adjusting for tobacco use history and thereby pointing out to an irrelevant correlation in this cohort. Despite some limitations by the small sample size, we obtained samples under direct observation from patients with biopsy proven categories of BE to EAC representing the disease progression spectrum. We tried to avoid contamination and used samples of gastric juice and adjacent squamous esophagus from the same cases to control for background microbial communities that would help identify disease-associated changes. We used the normal controls and replicate samples to facilitate removal of false OTUs that could lead to false discovery especially when considering the areas studied with expected low microbial biomass (41). One of the limitations of the data analysis is that given the low numbers of bacteria isolated from esophageal endoscopic biopsies, the conventional means of adjusting for false discovery may not adequately address the low numbers of reads and have the potential to overweight scarce taxa. The small sample size limited our ability to adjust for all the patient level factors mentioned.

Overall, our data identify significant changes in a small number of genera and phyla that track across the progression from normal esophageal mucosal microbial communities through IM and dysplasia to neoplasia. These observations will need to be confirmed and their biologic plausibility tested by accompanying studies of the transcriptome and metabolome of these communities and *in vitro* exposure using relevant gnotobiotic collections with epithelial models.

CONFLICTS OF INTEREST

Guarantor of the article: Shajan Peter, MD.

Specific author contributions: S.P.—design, coordination, patient enrollment acquisition and interpretation of data, literature review, analysis, interpretation, and drafting of the article.

A.P.—interpretation of data and statistical analysis. WVDP—data informatics and analysis. C.M.W.—patient enrollment and sample data collection. K.R.K.K.B.—patient enrollment and sample data collection. C.M.—intellectual content, design and coordination of the study, interpretation of data, and revision of manuscript for intellectual content. J.I.—intellectual content and critical review of analysis. P.M.—intellectual content, design and coordination of the study, acquisition, analysis, and interpretation of data and final approval of the manuscript.

Financial support: University of Alabama internal research funding.

Potential competing interests: None to report.

ACKNOWLEDGEMENTS

The following are acknowledged for their support of the Microbiome Resource at the University of Alabama at Birmingham: Comprehensive Cancer Center (P30AR050948), Center for Clinical Translational Science (UL1TR001417), University Wide Institutional Core, Heflin Center for Genomic Sciences and Microbiome Center. University of Alabama Health Service Foundation General endowment fund (UABHSF GEF).

Study Highlights

WHAT IS KNOWN

- ✓ The prevalence of BE and progression to EAC has been increasing.
- ✓ BE proceeds from advancing dysplasia to neoplasia.
- ✓ Microbes may contribute to induction of or protection from neoplasia.

WHAT IS NEW HERE

- ✓ Two oral and gastrointestinal tract microbial phyla are significantly decreased in BE metaplasia and esophageal adenocarcinoma.
- ✓ The findings are well controlled by using adjacent unaffected tissue and independent healthy controls.
- ✓ The methods focus on mucosal-adherent microbes directly obtained.

TRANSLATIONAL IMPACT

- ✓ These organisms have known metabolite production that can affect carcinogenesis.
- ✓ These microbes can be studied *in vitro* for effects on esophageal tissue/organoids to discover mechanisms of action and test interventions.
- ✓ Preventing progression of BE to neoplasia will change the natural history of this disease.

REFERENCES

1. Pei Z, Bini EJ, Yang L, et al. Bacterial biota in the human distal esophagus. *Proc Natl Acad Sci U S A* 2004;101:4250–5.
2. Elliott DRF, Walker AW, O'Donovan M, et al. A non-endoscopic device to sample the oesophageal microbiota: A case-control study. *Lancet Gastroenterol Hepatol* 2017;2:32–42.
3. Yang L, Lu X, Nossa CW, et al. Inflammation and intestinal metaplasia of the distal esophagus are associated with alterations in the microbiome. *Gastroenterology* 2009;137:588–97.
4. Snider EJ, Freedberg DE, Abrams JA. Potential role of the microbiome in Barrett's esophagus and esophageal adenocarcinoma. *Dig Dis Sci* 2016;61:2217–25.
5. Sharma P, Dent J, Armstrong D, et al. The development and validation of an endoscopic grading system for barrett's esophagus: The Prague C & M criteria. *Gastroenterology* 2006;131:1392–9.
6. Kumar R, Eipers P, Little RB, et al. Getting started with microbiome analysis: Sample acquisition to bioinformatics. *Curr Protoc Hum Genet* 2014;82:18.8.1–29.
7. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335–6.
8. Wang Q, Garrity GM, Tiedje JM, et al. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007;73:5261–7.

9. Cole JR, Wang Q, Cardenas E, et al. The ribosomal database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 2009;37:D141–5.
10. McDonald D, Price MN, Goodrich J, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 2012;6:610–8.
11. Caporaso JG, Bittinger K, Bushman FD, et al. PyNAST: A flexible tool for aligning sequences to a template alignment. *Bioinformatics* 2010;26:266–7.
12. Price MN, Dehal PS, Arkin AP. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;5:e9490.
13. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–40.
14. La Rosa PS, Brooks JP, Deych E, et al. Hypothesis testing and power calculations for taxonomic-based human microbiome data. *PLoS One* 2012;7:e52078.
15. Quante M, Abrams JA, Wang TC. The rapid rise in gastroesophageal junction tumors: Is inflammation of the gastric cardia the underwater iceberg? *Gastroenterology* 2013;145:708–11.
16. Cook MB, Wild CP, Forman D. A systematic review and meta-analysis of the sex ratio for Barrett's esophagus, erosive reflux disease, and nonerosive reflux disease. *Am J Epidemiol* 2005;162:1050–61.
17. Rubenstein JH, Shaheen NJ. Epidemiology, diagnosis, and management of esophageal adenocarcinoma. *Gastroenterology* 2015;149:302–17.e1.
18. Gall A, Fero J, McCoy C, et al. Bacterial composition of the human upper gastrointestinal tract microbiome is dynamic and associated with genomic instability in a Barrett's esophagus cohort. *PLoS One* 2015;10:e0129055.
19. Amir I, Konikoff FM, Oppenheim M, et al. Gastric microbiota is altered in oesophagitis and Barrett's oesophagus and further modified by proton pump inhibitors. *Environ Microbiol* 2014;16:2905–14.
20. Nadatani Y, Huo X, Zhang X, et al. NOD-like receptor protein 3 inflammasome priming and activation in Barrett's epithelial cells. *Cell Mol Gastroenterol Hepatol* 2016;2:439–53.
21. Quante M, Graham TA, Jansen M. Insights into the pathophysiology of esophageal adenocarcinoma. *Gastroenterology* 2018;154:406–20.
22. Munch NS, Fang HY, Ingermann J, et al. High-fat diet accelerates carcinogenesis in a mouse model of Barrett's esophagus via interleukin 8 and alterations to the gut microbiome. *Gastroenterology* 2019;157:492–506.e2.
23. Zaidi AH, Kelly LA, Kreft RE, et al. Associations of microbiota and toll-like receptor signaling pathway in esophageal adenocarcinoma. *BMC Cancer* 2016;16:52.
24. Helminen O, Huhta H, Leppanen J, et al. Nuclear localization of Toll-like receptor 5 in Barrett's esophagus and esophageal adenocarcinoma is associated with metastatic behavior. *Virchows Arch* 2016;469:465–70.
25. Helminen O, Huhta H, Lehenkari PP, et al. Nucleic acid-sensing toll-like receptors 3, 7 and 8 in esophageal epithelium, Barrett's esophagus, dysplasia and adenocarcinoma. *Oncoimmunology* 2016; 5:e1127495.
26. Kohtz PD, Halpern AL, Eldeiry MA, et al. Toll-like receptor-4 is a mediator of proliferation in esophageal adenocarcinoma. *Ann Thorac Surg* 2019;107:233–41.
27. Diakowska D, Nienartowicz M, Grabowski K, et al. Toll-like receptors TLR-2, TLR-4, TLR-7, and TLR-9 in tumor tissue and serum of the patients with esophageal squamous cell carcinoma and gastro-esophageal junction cancer. *Adv Clin Exp Med* 2019;28:515–22.
28. Huhta H, Helminen O, Lehenkari PP, et al. Toll-like receptors 1, 2, 4 and 6 in esophageal epithelium, Barrett's esophagus, dysplasia and adenocarcinoma. *Oncotarget* 2016;7:23658–67.
29. Kauppila JH, Selander KS. Toll-like receptors in esophageal cancer. *Front Immunol* 2014;5:200.
30. Yang L, Francois F, Pei Z. Molecular pathways: Pathogenesis and clinical implications of microbiome alteration in esophagitis and Barrett esophagus. *Clin Cancer Res* 2012;18:2138–44.
31. Abdel-Latif MM, O'Riordan J, Windle HJ, et al. NF-kappaB activation in esophageal adenocarcinoma: Relationship to Barrett's metaplasia, survival, and response to neoadjuvant chemoradiotherapy. *Ann Surg* 2004;239:491–500.
32. Fan YP, Chakder S, Gao F, et al. Inducible and neuronal nitric oxide synthase involvement in lipopolysaccharide-induced sphincteric dysfunction. *Am J Physiol Gastrointest Liver Physiol* 2001;280: G32–42.
33. Morris CD, Armstrong GR, Bigley G, et al. Cyclooxygenase-2 expression in the Barrett's metaplasia-dysplasia-adenocarcinoma sequence. *Am J Gastroenterol* 2001;96:990–6.
34. Wilson KT, Fu S, Ramanujam KS, et al. Increased expression of inducible nitric oxide synthase and cyclooxygenase-2 in Barrett's esophagus and associated adenocarcinomas. *Cancer Res* 1998;58:2929–34.
35. Cayrou C, Samba B, Armougom F, et al. Molecular diversity of the Planctomycetes in the human gut microbiota in France and Senegal. *APMIS* 2013;121:1082–90.
36. Thomas AM, Jesus EC, Lopes A, et al. Tissue-associated bacterial alterations in rectal carcinoma patients revealed by 16S rRNA community profiling. *Front Cel Infect Microbiol* 2016;6:179.
37. Rieu-Lesme F, Delbes C, Sollelis L. Recovery of partial 16S rDNA sequences suggests the presence of Crenarchaeota in the human digestive ecosystem. *Curr Microbiol* 2005;51:317–21.
38. Bang C, Schmitz RA. Archaea associated with human surfaces: Not to be underestimated. *FEMS Microbiol Rev* 2015;39:631–48.
39. Nkanga VD, Henrissat B, Drancourt M. Archaea: Essential inhabitants of the human digestive microbiota. *Hum Microbiome J* 2017;3:1–8.
40. Gaci N, Borrel G, Tottey W, et al. Archaea and the human gut: New beginning of an old story. *World J Gastroenterol WJG* 2014;20:16062.
41. Salter SJ, Cox MJ, Turek EM, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014;12:87.

Open Access This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work, provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.