The Salivary Microbiome and Predicted Metabolite Production are Associated with Progression from Barrett's Esophagus to Esophageal Adenocarcinoma

Quinn S Solfisburg[1], Federico Baldini[2], Brittany L Baldwin-Hunter[3], Harry H Lee[2], Heekuk Park[3,4], Daniel E Freedberg[3,5], Charles J Lightdale[3], Tal Korem[2,6,7]*, Julian A Abrams[3,5,8]*

*co-corresponding authors

[1]Department of Medicine, Boston University School of Medicine, Boston, MA, USA
[2]Program for Mathematical Genomics, Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, USA
[3]Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA
[4]Microbiome and Pathogen Genomics Collaborative Center, Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA
[5]Digestive and Liver Disease Research Center, Columbia University Irving Medical Center, New York, NY, USA
[6]Department of Obstetrics and Gynecology, Columbia University Irving Medical Center, New York, NY, USA
[7]CIFAR Azrieli Global Scholars Program, CIFAR, Toronto, Canada
[8]Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY USA

Co-corresponding authors:
Tal Korem, PhD
530 W 166th Street
Alianza Building, 3rd floor
New York, NY 10032
E-mail: tal.korem@columbia.edu

Julian Abrams, MD, MS
630 W 168th Street
P&S 3-401
New York, NY 10032
E-mail: ja660@cumc.columbia.edu

Abstract:

Esophageal adenocarcinoma (EAC) is rising in incidence and associated with poor survival, and established risk factors do not explain this trend. Microbiome alterations have been associated with progression from the precursor Barrett's esophagus (BE) to EAC, yet the oral microbiome, tightly linked to the esophageal microbiome and easier to sample, has not been extensively studied in this context. We aimed to assess the relationship between the salivary microbiome and neoplastic progression in BE to identify microbiome-related factors that may drive EAC development. We collected clinical data and oral health and hygiene history and characterized the salivary microbiome from 250 patients with and without BE, including 78 with advanced neoplasia (high grade dysplasia or early adenocarcinoma). We assessed differential relative abundance of taxa by 16S rRNA gene sequencing and associations between microbiome composition and clinical features and used microbiome metabolic modeling to predict metabolite production. We found significant shifts and increased dysbiosis associated with progression to advanced neoplasia, with these associations occurring independent of tooth loss, and the largest shifts were with the genus *Streptococcus*. Microbiome metabolic models predicted significant shifts in the metabolic capacities of the salivary microbiome in patients with advanced neoplasia, including increases in L-lactic acid and decreases in butyric acid and L-tryptophan production. Our results suggest both a mechanistic and predictive role for the oral microbiome in esophageal adenocarcinoma. Further work is warranted to identify the biological significance of these alterations, to validate metabolic shifts, and to determine whether they represent viable therapeutic targets for prevention of progression in BE.

INTRODUCTION:

Esophageal adenocarcinoma (EAC) has seen a dramatic rise in incidence in the past several decades, is often diagnosed at advanced stages, and is associated with poor survival.[1,2] The factors that drive EAC remain incompletely understood. Barrett's esophagus (BE) is the precursor lesion to EAC, but the overwhelming majority of BE patients do not progress to EAC. Established EAC risk factors, including gastroesophageal reflux disease (GERD) and obesity, do not fully explain the rise in its incidence.[3,4]

Increasing evidence suggests that the microbiome plays an important role in modifying the risk of a variety of epithelial cancers[5-8] as well as in modulating the response to treatment.[9-11] Changes in the esophageal microbiome have been observed in EAC and with progression from BE to EAC,[12,13] raising the possibility that bacteria contribute to esophageal neoplasia. Reliable sampling of the esophageal microbiome, however, requires invasive procedures. A more accessible "window" to the esophageal ecosystem is the oral microbiome, which was shown to strongly influence it.[14] A small study of the tumor-associated microbiome in EAC found a high prevalence of domination by oral flora such as *Streptococcus,*[12] pointing to a link between the oral microbiome and EAC. Oral microbiome alterations have been associated with future risk of EAC[15], and differences in the oral microbiome of BE patients were described previously in a small study of 49 patients.[16] Alterations in the oral microbiome have also been associated with poor oral health[17], which was in itself associated with increased risk of EAC in a recent analysis.[18] It remains unclear how oral dysbiosis and poor oral health interact in their association with EAC. Finally, little is known with regard to oral microbiome alterations associated with neoplastic progression in BE patients. A clearer understanding of these oral microbiome changes could identify factors that may drive progression of neoplasia, representing novel therapeutic targets.

Here, we profiled the salivary microbiome from 250 patients with various stages of BE and EAC who were undergoing upper endoscopy. We identify multiple characteristics of the oral microbiome associated with neoplastic progression in BE and

show that they are independent of oral health. Using metabolic modeling, we predict metabolite profiles associated with alterations in BE, suggesting a mechanistic role for microbially produced metabolites. Finally, we show that the salivary microbiome offers a mild improvement in diagnostic accuracy compared to models based on clinical risk factors. Our results demonstrate the potential of studying the oral microbiome in the context of progression to EAC.

RESULTS:

Oral microbial composition from a large endoscopy cohort

We recruited 250 adult patients undergoing upper endoscopy and characterized their oral microbiome using 16S rRNA gene sequencing. (**Methods**) A total of 244 patients were included in the analyses: 125 controls without Barrett's esophagus (BE), and 119 BE patients (20 with non-dysplastic BE, 11 indefinite for dysplasia, 10 low

	All (n=244)	Non-BE (n=125)	BE (n=119)	P-value
Age, years; mean (SD)	57.8 (18.7)	50.9 (18.7)	65.0 (15.8)	<0.001
Sex, male; N (%)	140 (57%)	45 (36%)	95 (80%)	<0.001
Ever smokers; N (%)	103 (42%)	41 (33%)	62 (52%)	0.002
BMI mean (SD)	27.5 (6.7)	27.0 (7.9)	28.1 (5.1)	<0.001
Race, white; N (%)	222 (90%)	105 (84%)	117 (98%)	<0.001
GERD; N (%)	167 (68%)	64 (51%)	103 (87%)	<0.001
PPI use; N (%)	148 (60%)	41 (33%)	107 (90%)	<0.001
Aspirin use; N (%)	77 (31%)	25 (20%)	52 (44%)	<0.001
Oral health and hygiene				
Tooth loss; N (%)	127 (52%)	52 (42%)	75 (63%)	<0.001
Tooth brushing frequency, ≥ daily; N (%)	233 (95%)	123 (98%)	110 (92%)	0.03
Mouthwash frequency, ≥ daily; N (%)	139 (56%)	69 (55%)	70 (59%)	0.61

Table 1. Patient Characteristics. P – t-test or Fisher exact p for difference between BE and non-BE.

grade dysplasia, 54 high grade dysplasia, and 24 intramucosal (T1a) adenocarcinoma). Patients with BE were more likely to be older (t-test p<0.001), male (Fisher exact p<0.001), white (p=0.001), or ever-smokers (defined as ≥100 lifetime cigarettes smoked) (p=0.003). They were also more likely to have GERD (p<0.001), to be treated with proton pump inhibitor (PPI; p<0.001), to take aspirin (p<0.001), and to have a higher BMI (p<0.001). (**Table 1**) There was no significant difference in the use of mouthwash between BE and non-BE patients (p=0.61), but non-BE patients were more likely to brush their teeth at least daily (98% vs 92%, p=0.03). Compared to non-BE, a significantly higher proportion of patients with BE had tooth loss (63% vs. 42%, p=0.001), largely due to an increase in tooth loss in patients with advanced neoplasia (defined as high grade dysplasia or adenocarcinoma, 66%; non-dysplastic BE, 40%; non-BE, 42%; Fisher's exact p=0.001). (**Figure 1**) Older age (per year, adjusted OR 1.06, 95% CI 1.04-1.08) and a history of smoking (adjusted OR 2.12, 95% CI 1.07-4.19) were independently associated with tooth loss. (**Supplementary Table 1**) In multivariable analyses adjusting for EAC risk factors (age, male sex, white race, and GERD), tooth loss was associated with a non-significant increased risk of advanced neoplasia (vs. non-BE, adjusted OR 1.49, 95%CI 0.96-2.47). (**Supplementary Table 2**) These results are in line with a recent analyses of data from the Nurses' Health Study, which found an association between both tooth loss and periodontal disease and risk of esophageal adenocarcinoma, but showed a decrease in association strength after adjusting for covariates.[18] Established EAC risk factors were independently associated with advanced neoplasia even after controlling for daily tooth brushing, use of mouthwash, and presence of tooth loss, whereas these measures of oral health and hygiene were not independently associated with advanced neoplasia (p=0.21, 0.57, and 0.34, respectively).

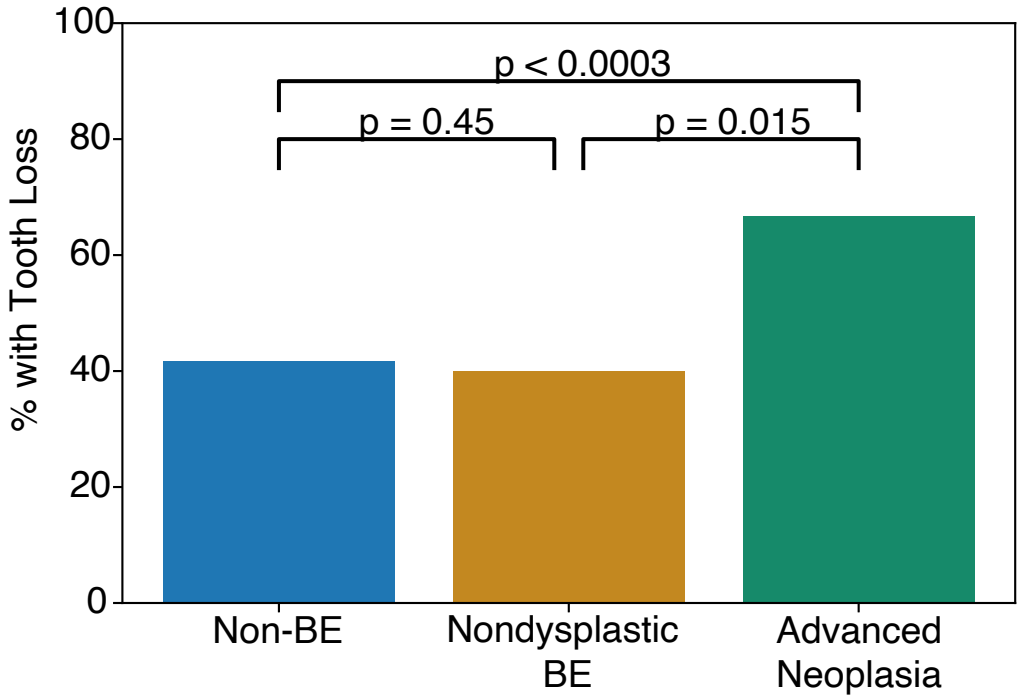

Figure 1. Tooth loss is significantly more common in advanced neoplasia. A significantly higher proportion of patients with advanced neoplasia (high grade dysplasia or esophageal adenocarcinoma) had tooth loss as compared to non-BE and non-dysplastic BE patients combined (Fisher's exact p = 0.001). P – Fisher's exact.

The oral microbiome of BE patients is progressively altered with dysplastic changes.

To assess whether the salivary microbiome is associated with neoplastic progression, we focused our analyses on comparisons between three groups: non-BE (n=125), non-dysplastic BE (n=20), and advanced neoplasia (high grade dysplasia and adenocarcinoma [HGD/EAC]; n=78). We found that neoplastic progression was associated with significantly lower alpha diversity (Shannon: Kruskal-Wallis p=0.005, **Figure 2A**; Simpson: p=0.0029, **Supplementary Figure 1**). Compared to patients without BE, the alterations in alpha diversity were more pronounced in patients with advanced neoplasia than in those with nondysplastic BE (non-BE vs. non-dysplastic BE, Mann-Whitney p=0.11; non-BE vs advanced neoplasia, p=0.0006). There was no significant difference in alpha diversity comparing nondysplastic BE with advanced

neoplasia (p=0.23). We further found that the oral microbiome from patients with advanced neoplasia tended to cluster separately than the rest of the cohort (weighted UniFrac, ANOSIM p<0.001; **Figure 2B**). Similar results were found when including all the subjects in their individual groups (non-BE, nondysplastic BE, IND, LGD, HGD, and EAC; **Supplementary Figure 2**). Our results indicate that the salivary microbiome alterations observed in advanced neoplasia are reflected both in the diversity of each individual's microbiome (alpha diversity) and in compositional differences between individuals (beta diversity).

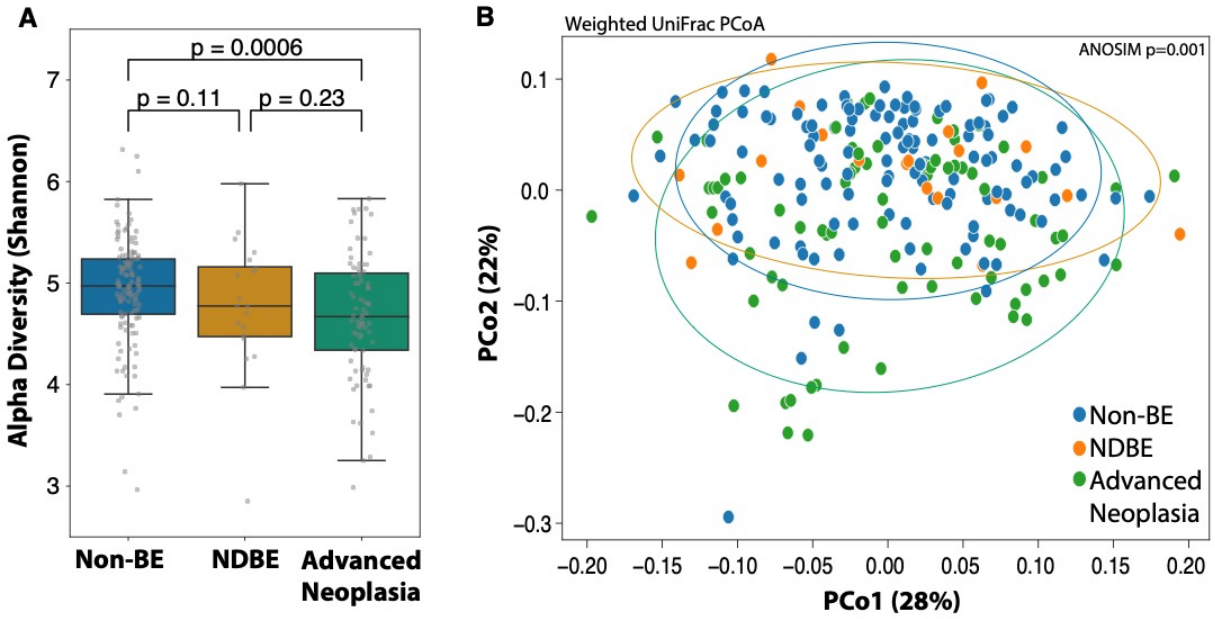

Figure 2. A microbial signature of BE. (A) Patients with advanced neoplasia have significantly reduced alpha diversity compared to non-BE patients. Kruskal-Wallis overall p-value=0.005. p – Mann-Whitney U test. (B) Weighted UniFrac PCoA demonstrated significant clustering of patients with advanced neoplasia (ANOSIM p=0.001). NDBE, non-dysplastic BE; advanced neoplasia – high grade dysplasia or esophageal adenocarcinoma; ellipse, 2 standard deviation sigma ellipse.

We next checked whether specific microbes were associated with BE progression. We therefore compared relative abundance of different OTUs between all BE patients vs. non-BE using ALDEx2.[19] (**Methods**) A total of 26 OTUs were identified as differentially abundant ($p < 0.05$, FDR corrected at 0.1; **Figure 3**). To assess whether the dysbiotic signature associated with BE is more pronounced with dysplastic

changes, we checked whether these 26 OTUs were correlated with progression across the neoplastic spectrum, from no dysplasia to EAC (Methods). There was a significant association ($p<0.05$, FDR corrected at 0.1) for 23 of the 26 taxa, with a clear shift in composition with neoplastic progression, notably in the transition from low grade dysplasia (LGD) to high grade dysplasia (HGD). (**Figure 3**) This transition in composition from LGD to HGD is consistent with our prior observations of esophageal microbiome alterations with progression to EAC.[20] The taxonomic alterations associated with progression were notable for increased relative abundance of several *Streptococcus* species. Streptococci form biofilms in the oral cavity[21-23].

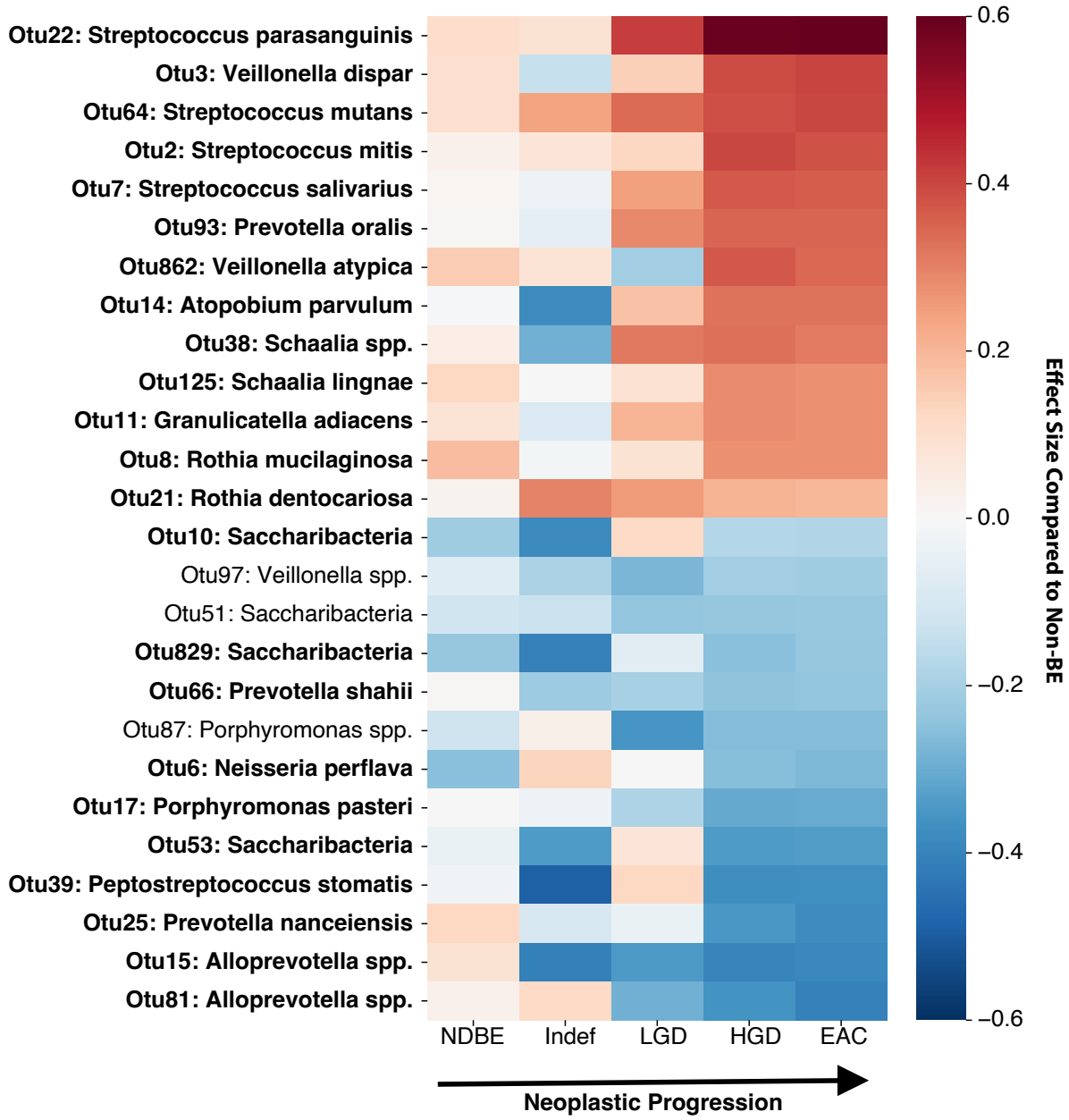

Figure 3. Increased oral dysbiosis with progressive dysplasia. Shifts in the oral microbiome compared to non-BE patients (shown as ALDEx2 effect sizes) were more pronounced with progression from no dysplasia to EAC, particularly notable in patients with high grade dysplasia and EAC. Bolded OTUs were significantly associated with neoplastic progression. ($p < 0.05$, FDR corrected at 0.1; **Methods**) NDBE, nondysplastic Barrett's esophagus; Indef, indefinite for dysplasia; LGD, low grade dysplasia; HGD, high grade dysplasia; EAC, esophageal adenocarcinoma.

The salivary microbiome is associated with advanced neoplasia even when controlling for tooth loss.

Tooth loss is known to be strongly associated with oral microbiome composition[17], and there was an increased proportion of advanced neoplasia patients with tooth loss. (**Figure 1**) Comparing patients who did and did not have all or most of their natural adult teeth, we found that those with tooth loss had lower alpha diversity (Shannon, Mann-Whitney p=0.001; **Supplementary Figure 3A**), and that oral microbiomes from both groups clustered separately (weighted UniFrac, ANOSIM p=0.003; **Supplementary Figure 3B**). We also identified 29 OTUs that had significantly different abundance between patients with and without tooth loss. (ALDEx2 $p < 0.05$, FDR corrected at 0.1; **Figure 4**) As poor oral health is associated with oral dysbiosis, this raised the question of whether the oral microbiome is associated with advanced neoplasia independent of tooth loss.

We first examined whether salivary microbiome composition as a whole is associated with advanced neoplasia independently of tooth loss. We therefore calculated microbiome principal coordinates (PCos) using weighted UniFrac distances and used the top five PCos, which represented two thirds of the variance in microbiome composition. We then used multivariable logistic regression and found that PCo2 (explaining 22% of microbiome variation) and PCo4 (4.6%) were independently associated with advanced neoplasia. ($p<0.001$ and $p=0.004$, respectively) We then added the major EAC risk factors (age, sex, race, BMI, GERD history, and smoking history) to the model, and found that PCo2 remained independently associated with advanced neoplasia ($p=0.004$), suggesting that salivary microbiome composition represents a potential novel independent risk factor for EAC. Adding tooth loss to the model did not alter the association between PCo2 and advanced neoplasia ($p=0.004$), and in this model tooth loss was not independently associated with advanced neoplasia ($p=0.12$). Our results suggest that the association of tooth loss with advanced neoplasia is mediated through the oral microbiome.

We next assessed whether associations between specific oral taxa and advanced neoplasia are independent of tooth loss. Of the 33 taxa associated with

advanced neoplasia (N=78) vs. non-BE (N=125; ALDEx2 $p < 0.05$; FDR corrected at 0.1), 18 were also associated with tooth loss. (**Figure 4**) After adjusting for tooth loss in a generalized linear model, 20 of these taxa remained significantly associated with advanced neoplasia. (ALDEx2 $p < 0.05$, FDR corrected at 0.1) Notably, the four OTUs with the greatest increase in relative abundance in advanced neoplasia were all assigned to the genus *Streptococcus*, and the increased abundance of these *Streptococcus* OTUs in advanced neoplasia was independent of tooth loss. This corresponds with previous studies that found that the tumor-associated microbiome in EAC is often dominated by *Streptococcus* species.[12]

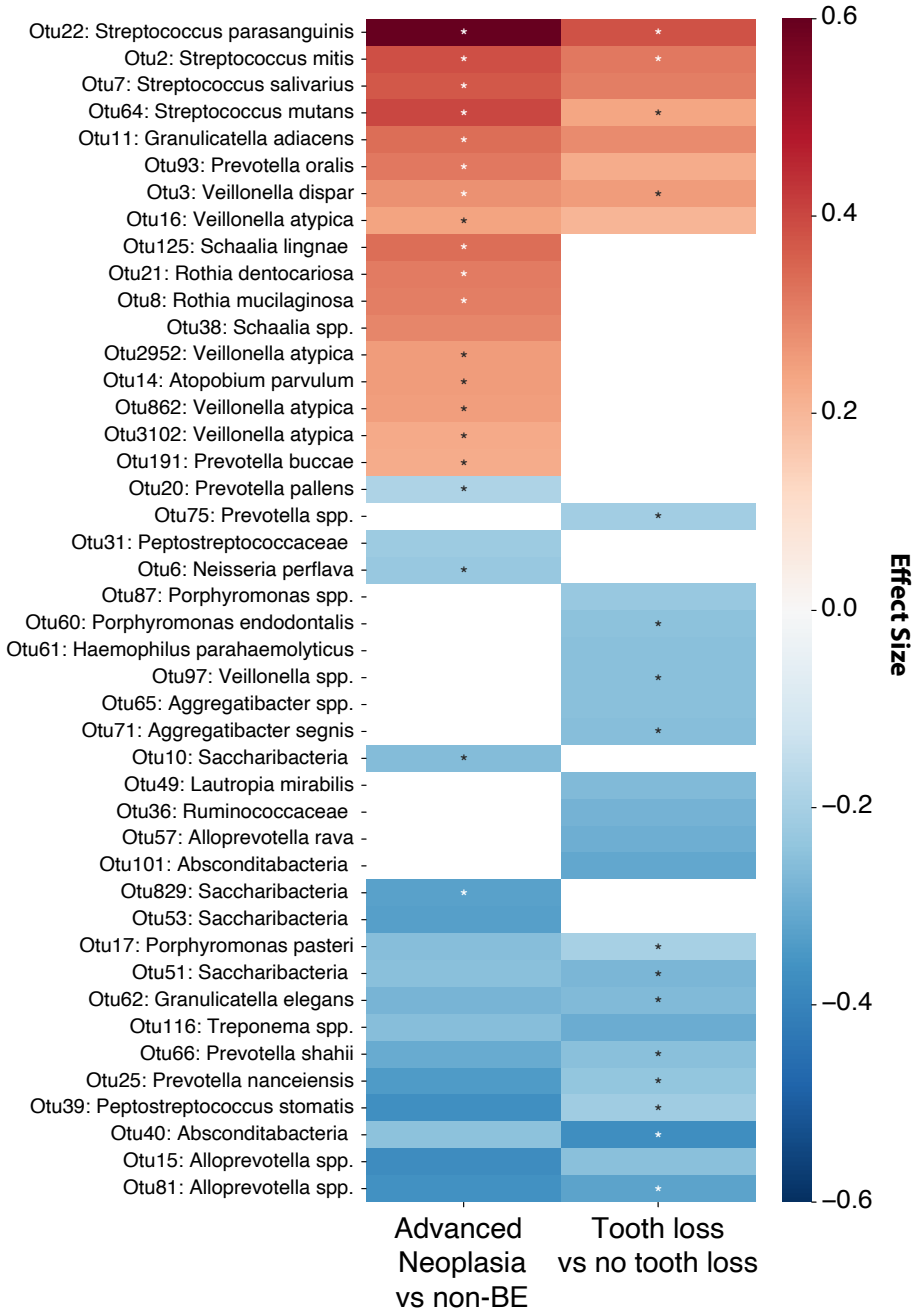

Figure 4. Oral microbes are independently associated with advanced neoplasia and tooth loss. Differentially abundant taxa in patients with advanced neoplasia (high-grade dysplasia or esophageal adenocarcinoma) compared to non-BE (left) and in patients with and without tooth loss (right). Many taxa were associated with both neoplasia and tooth loss, yet most of these remained significantly differentially abundant after adjusting for tooth loss and advanced neoplasia, respectively (ALDEx2 $p < 0.05$, FDR corrected at 0.1; denoted by asterisks).

170

Metabolic modeling predicts distinct metabolic secretion capabilities in advanced neoplasia.

Metabolite production by microbial communities is an important modality by which the microbiome affects the host. In order to assess if microbially produced metabolites might play a role as a driver or biomarker of neoplasia, we used microbiome community-scale metabolic models to predict metabolite secretion by the microbiome for every sample. (**Methods**) We found significant clustering of predicted metabolite profiles comparing advanced neoplasia cases with non-BE subjects (PERMANOVA p=0.001). Using principal components analysis, we found notable shift in the second component (15% explained variance; Mann-Whitney p=0.0003). (**Figure 5A**) Forty-four predicted metabolites had significantly altered abundance ($p < 0.05$, FDR corrected at 0.1) in advanced neoplasia. (**Figure 5B**) Notable alterations included increased predicted levels of L-lactic acid (p=0.023), a by-product of aerobic glycolysis, a hallmark of cancer which can contribute to neoplasticity;[24] and 2-ketobutyric acid (p=0.033), previously reported to support mitochondrial respiration and cell proliferation.[25] We also predicted that advanced neoplasia features a decrease in butyric acid (p=0.0089), a key promoter of gut homeostasis that was previously shown to be depleted in colon cancer and inflammatory bowel disease;[26-28] and a decrease in L-tryptophan (p=0.0017). (**Figure 5C**) Circulating levels of tryptophan have been inversely associated with colon cancer risk[29], and melatonin, a by-product of L-tryptophan metabolism, is under investigation for EAC prevention.[30] The pattern of the shifts we predict is therefore consistent with the potential promotion of proliferation, inflammation, and cancer.

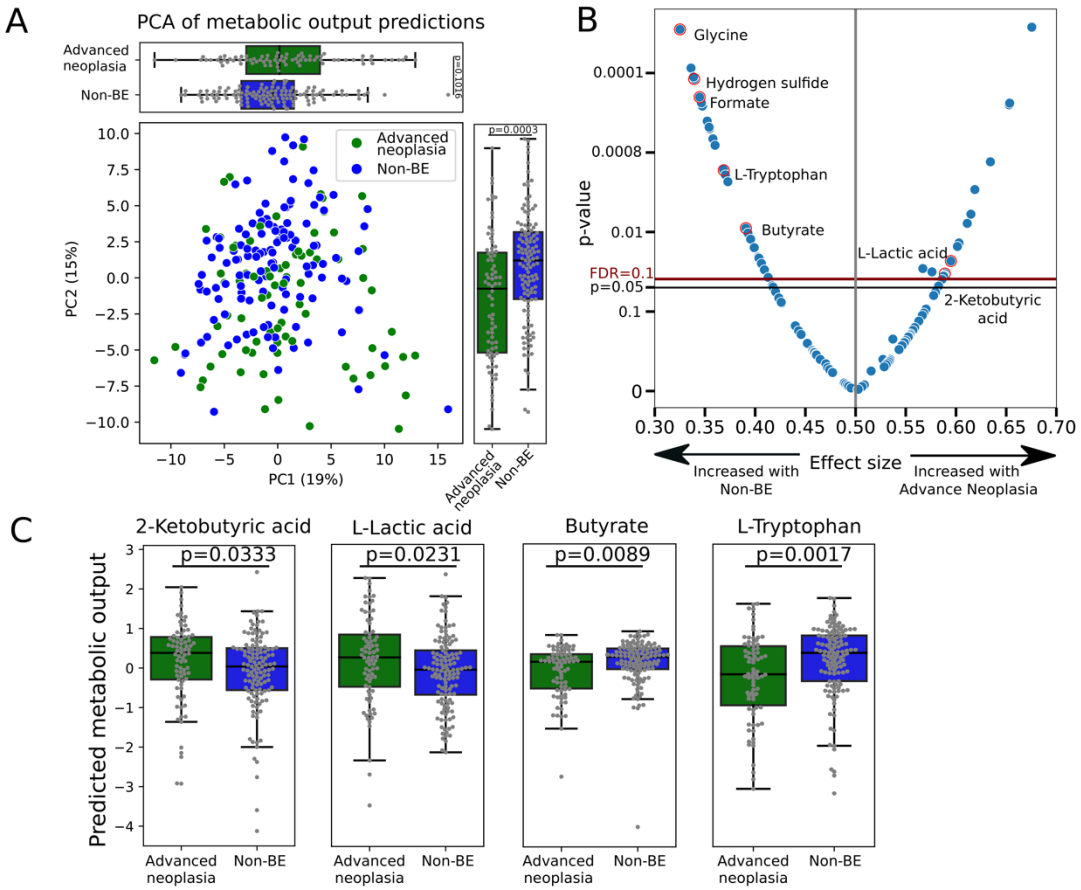

Figure 5. The predicted metabolic profile is altered in patients with advanced neoplasia. (A) Significant clustering by advanced neoplasia status on principal components analysis (PERMANOVA p=0.001), with pronounced shifts in PC2 (p=0.0003). (B) Volcano plot demonstrating differentially abundant metabolites in advanced neoplasia. (C) Significantly altered predicted levels of L-lactic acid, 2-ketobutyric acid, butyric acid, and L-tryptophan in advanced neoplasia. Plot capped at -4 for butyrate. P, Mann-Whitney test

Salivary microbiome data improves on clinical risk factors-based prediction of advanced neoplasia.

We next performed an exploratory analysis to determine whether salivary microbiome features in this cohort could be used to distinguish advanced neoplasia from non-BE patients. As a baseline, we first trained a gradient boosted decision trees model which uses clinical risk factors for EAC and tooth loss to classify advanced neoplasia. The classifier was tested in cross-validation on patients not seen in the training of that model and achieved an area under the receiver operating characteristic

curve (AUROC) of 0.84 (95%CI 0.79-0.89). The same process was then used to train a classifier using microbiome data, whereas, within each training fold, 10 OTUs were selected based on a Kruskal-Wallis test. This classifier had an AUROC of 0.72 (95%CI 0.65-0.79). Finally, a model trained on the combination of both microbiome data and EAC risk factors resulted in somewhat higher model accuracy, producing an AUROC of 0.88 (95%CI 0.83-0.91; vs. clinical risk factor model AUROC 0.84, 95%CI 0.79-0.89; DeLong p=0.053). (**Supplementary Figure 4**) A combined microbiome and clinical risk factor model with the outcome limited to HGD and excluding intramucosal EAC showed similar results with an AUROC 0.86 (95%CI 0.80-0.92), compared to an AUROC of 0.83 (95%CI 0.77-0.89) using only clinical risk factors for the same task. (**Supplementary Figure 5**)

DISCUSSION:

In this cross-sectional study of patients with and without BE, we detected marked shifts in the salivary microbiome with progression to EAC, with changes that appeared to be most pronounced in patients with advanced neoplasia. These changes included reduced diversity as well as significantly increased relative abundance of several taxa in the genus *Streptococcus*. As in previous studies, we found that tooth loss is more common in patients with advanced neoplasia. However, we show that many of the salivary microbiome associations observed in BE and advanced neoplasia persisted even when accounting for it. Further, we used metabolic modeling to identify distinct predicted metabolic secretion capabilities in advanced neoplasia.

Our findings add to the growing body of evidence that the oral microbiome is linked to the esophageal microbiome and may contribute to esophageal neoplasia. In a case-control study of patients with EAC, BE, and controls, the EAC-associated microbiome had significantly reduced alpha diversity, similar to our observations in saliva.[12] Interestingly, in that study 5/15 of the EAC tumors were dominated by *Streptococcus* spp. (relative abundance 69%-98%). In our salivary microbiome analyses, 4 of the 5 taxa most strongly associated with advanced neoplasia were also

Streptococcus spp. Our group conducted a randomized controlled trial and found that an antimicrobial mouth rinse can produce esophageal microbiome and tissue gene expression changes, highlighting the relevance of oral bacteria to esophageal disease.[31] In another cohort study analyzing mouth rinse samples from patients enrolled in two large cancer prevention studies, oral microbiome alterations were noted to precede an EAC diagnosis by several years.[15] In a small study of 49 patients we previously noted marked salivary microbiome alterations associated with BE and also with advanced neoplasia.[16]

Our study features the use of microbiome metabolic models to identify broad shifts in metabolites predicted to be produced by the saliva microbiome. Many of the predicted changes to metabolite outputs correspond with existing knowledge. Lactic acid, for example, was predicted to be increased in advanced neoplasia. Lactic acid can serve as a major energy source for proliferative cancer cells, and is known to activate hypoxia inducible factors, which in turn contribute to proliferation, angiogenesis, and other neoplastic features.[24] Our findings could therefore support the hypothesis that the oral and esophageal microbiota promotes EAC development and progression via production of metabolites.[32] Our findings further correspond with a previous study detecting lactic-acid bacteria in many esophageal adenocarcinomas.[12] However, the biological significance of predicted metabolite production is unclear, and future studies are needed to validate these predictions and to elucidate the biological effects of specific bacterial metabolites on esophageal neoplasia.

Prior work has associated tooth loss and periodontal disease with increased risks of esophageal squamous cell cancer and gastric cancer, and a recent study found an association between both tooth loss and periodontal disease and risk of esophageal adenocarcinoma.[18] This indicates a potential confounding effect, as tooth loss is also associated with major alterations in oral microbiome composition.[17] Our study offers an explanation for these associations, demonstrating that salivary microbiome composition is independently associated with advanced neoplasia, even when adjusting for EAC risk factors and for tooth loss. These findings suggest that the association between tooth

loss and esophageal neoplasia is mediated by changes in the salivary microbiome, and that the salivary microbiome may represent a novel independent risk factor for EAC.

We performed exploratory analyses to assess whether the salivary microbiome could discriminate patients at highest EAC risk. The salivary microbiome is highly suitable for diagnostics, as it is stable over time[33-35], especially compared to other body sites[36], and is resistant to perturbations.[37] Addition of a microbiome-based classifier to EAC risk factors resulted in modest improvement in discrimination. However, the current study was not specifically designed to address this question, and future studies should explore further the salivary microbiome as a potential biomarker for advanced neoplasia.

Important strengths of the current study include the relatively large sample size and the inclusion of oral health and hygiene information from patients. The large sample size allowed for the detection of significant microbiome alterations, even when correcting for multiple comparisons. Previous studies of the oral microbiome in BE and EAC have not included oral health and hygiene data, key potential confounders. The patients were well characterized, with data collected on key EAC risk factors including GERD history, BMI, and smoking, which permitted microbiome analyses adjusting for these variables. The BE patients in the study were demographically similar to BE populations from other studies, enhancing the generalizability of the findings. Lastly, novel methods for predicted microbiome metabolic profiling allowed for insights into functional correlates of the salivary microbiome alterations.

The study does have certain limitations. There were a relatively small number of non-dysplastic BE patients, limiting analyses in this subgroup. Analyses did not incorporate dietary intake; however, previous studies suggest that diet has minimal impact on salivary microbiome composition.[38-40] No conclusions can be drawn with regard to temporality in this cross-sectional study. It is possible that the observed salivary microbiome alterations were caused by BE-associated advanced neoplasia, although we believe that this is unlikely. Tooth loss was self-reported rather than measured, and periodontal disease was not directly assessed. Community-scale metabolic models also have notable limitations. Our analysis was based on 16S rRNA

gene sequencing, which does not allow us to tailor models to specific strains or genetic potential present in each sample. Additionally, while genome-scale models have been curated for common gut commensals, to our knowledge, such efforts have not been done for oral microbes. Consequently, some models may be missing, while existing ones may lack representation of niche-specific metabolic capacity. Despite these limitations, these models allow a systematic application of biochemical and genetic knowledge to our analysis and raise interesting hypotheses that could be experimentally validated.

In conclusion, patients with BE-associated advanced neoplasia have a markedly altered salivary microbiome, and analyses of taxonomic alterations associated with stages of progression from BE to EAC appear to indicate that these changes are most notable at the transition from low- to high-grade dysplasia. Increased tooth loss was also observed with progression to EAC, although the salivary microbiome alterations were largely independent of tooth loss, suggesting that the association of tooth loss with advanced neoplasia is mediated through the oral microbiome. There were marked increases in various taxa in the genus *Streptococcus* in advanced neoplasia, possibly pointing to a biological contribution of these bacteria to neoplastic progression. In addition to the microbiome alterations, progression to EAC was associated with numerous changes to predicted bacterial metabolite production, with notable alterations that suggest possible proneoplastic effects related to these shifts. Further work is warranted to identify the biological significance of the microbiome alterations, to validate metabolic shifts, and to determine whether they represent viable therapeutic targets for prevention of progression in BE.

METHODS:

Study Design

A total of 250 patients with and without BE undergoing upper endoscopy at Columbia University Irving Medical Center (New York, NY) were prospectively enrolled from February 2018 through February 2019. Patients were ≥18 years old and

scheduled to undergo endoscopy for clinical indications. Patients were excluded if they had a concurrently scheduled colonoscopy, had a history of gastric or esophageal surgery, a history of esophageal squamous cell cancer, or use of antibiotics, steroids, or other immunosuppressants in the 3 months prior to the procedure. This study was approved by the Columbia University Institutional Review Board. All patients provided written informed consent.

Data were collected on patient demographics and anthropometrics (to calculate BMI) as well as clinical information including medical history, history of gastro-esophageal reflux disease (GERD; defined as experiencing frequent heartburn or fluid regurgitation), medication use at time of enrollment (with specific notation of daily use of proton pump inhibitors (PPIs), histamine-2 receptor antagonists, statins, and daily use of aspirin and non-steroidal anti-inflammatory drugs), alcohol history, and smoking history (ever smoking defined as having smoked >100 lifetime cigarettes). Data were collected on self-reported oral health and hygiene. Tooth loss was assessed using categories adapted from Borningen et al.[17]: all or most of natural adult teeth, partial plates or implants, full upper dentures or implants, full lower dentures or implants, full upper and lower dentures or implants. Data were also collected on tooth brushing and mouthwash use.

Patients did not eat or drink after midnight prior to the endoscopy and saliva collection; saliva was collected prior to the endoscopy. Patients were categorized as BE if they had a history of endoscopically suspected BE with intestinal metaplasia on esophageal biopsies. BE patients were further categorized based on the highest degree of neoplasia ever (no dysplasia (NDBE), indefinite for dysplasia (IND), low grade dysplasia (LGD), high grade dysplasia (HGD), adenocarcinoma (EAC)).

Microbiome Sequencing and Analysis

The 16S rRNA V3-V4 region was amplified using Illumina adapter-ligated primers.[41] The Illumina Nextera XT v2 index sets A-D were used to barcode sequencing libraries. Libraries were sequenced on an Illumina MiSeq using the v3 reagent kit (600 cycles) and a loading concentration of 12 pM with 10% PhiX spike-in.

Sequences were assigned to operational taxonomic units (OTUs) using USEARCH[42] with ≥ 97% sequence homology. Taxonomic assignments for the OTUs were based on the Human Oral Microbiome Database (HOMD).[43] Any subsequently unassigned OTUs were assigned by referencing the Ribosomal Database Project (RDP).[44] Samples were subsampled to 10,000 reads to compare across even sequencing depths while minimizing data loss. Five patients were excluded after sequencing because of relatively low sequencing depth with <10,000 total reads per sample. The median read count for the full cohort was >33,000. One patient was excluded because of a history of both EAC and esophageal squamous cell carcinoma.

Microbiome metabolic modeling of oral microbial communities

Microbiome metabolic modeling was performed using the Microbiome Modeling Toolbox (COBRA toolbox commit: 71c117305231f77a0292856e292b95ab32040711) [45,46] and the AGORA metabolic models (AGORA 1.02).[47] All computations were performed in MATLAB version 2019a (Mathworks, Inc.), using the IBM CPLEX (IBM, Inc.) solver. We first matched species detected by our microbial sequencing analysis with the ones present in AGORA.[47] Because AGORA metabolic models are available at the strain level, we generated species-level models using the createPanModels.m function of the Microbiome Modeling Toolbox (MMT)[45] as previously described.[48] To increase the number of species represented in our microbiome models we chose genus-level representative models for abundant microbes present in the oral cavity with >5% relative abundance in more than 10 samples. There were six species without a corresponding metabolic model, and these were either grouped with similar species or excluded from the analyses (See **Supplementary Table 3** for details).

We then used the mgPipe.m automated pipeline of the MMT to build and interrogate sample-specific microbiome metabolic models. Briefly, for each sample, personalized microbiome models are created by joining species-level metabolic models using the compartmentalization technique[49]; a lumen compartment enabling microbial metabolic interactions is added, as well as additional input and output compartments, allowing microbiome intake and secretion of metabolites. Altogether our microbiome

models included 160 microbial species with an average of 50 species for each sample and a maximum of 69. As constraint-based metabolic modeling benefits from a specification of the metabolic environment such as media and carbon source availability[50], we applied a "western diet"[51] to each sample in the form of constraints on the metabolites uptake reactions.[51] Finally, to obtain metabolic predictions, we used the Net Maximal Production Capabilities (NMPCs) through the mgPipe pipeline[45] to provide predictions of the metabolite secretion profile of each sample. To detect significant changes in NMPCs distributions between cases and controls a Mann-Whitney U test was performed for each retained NMPCs. Only NMPCs which were present in at least 10% of the cases and had at least a value of 0.01 were retained for the significance analysis. FDR correction using the Benjamini–Hochberg procedure was applied.

Statistical Analysis

The primary groups of comparison were BE patients with advanced neoplasia (HGD or EAC), non-dysplastic BE (NDBE) and non-BE controls. Grouping high grade dysplasia and intramucosal adenocarcinoma together as advanced neoplasia reflects common practice as well as clinical guidelines for treatment.[52] There is extremely low inter-observer agreement (even among expert gastrointestinal pathologists) for the diagnosis of LGD[53-55], as inflammation-induced cytologic atypia mimics the findings of LGD. As a result, while estimates of cancer risk for LGD are relatively low on average,[54] these estimates vary widely, thus making interpretations of findings for this group challenging. Patients with low grade dysplasia or indefinite for dysplasia were included in analyses assessing for alterations in the oral microbiome across the entire BE neoplastic spectrum. While there was no *a priori* reason to suspect that endoscopic therapy would have altered the salivary microbiome, comparisons were made between those patients with LGD or worse who had (n=78) and had not (n=10) received prior endoscopic therapy. There were no differences in alpha diversity (p=0.16), no evidence of clustering on beta diversity analyses (ANOSIM p=0.13), and no differentially abundant taxa. Thus, treated and untreated patients were grouped together for all analyses.

Categorical variables were compared across groups using Fisher's exact tests. Continuous variables were analyzed using t-tests or rank sum tests as appropriate, with ANOVA and Kruskal Wallis tests for ≥3 groups. For purposes of analyses, tooth loss was dichotomized as having all or most of natural adult teeth (yes/no). Multivariable logistic regression was performed to assess the association between tooth loss and advanced neoplasia, adjusted for known EAC risk factors (age, sex, GERD, body mass index (BMI), smoking).

Alpha diversity was evaluated using the Shannon diversity index and beta diversity using weighted UniFrac[56] distances. Groups were compared using both permutational multivariate analysis of variance (PERMANOVA) for predicted metabolite profiles and analysis of similarities (ANOSIM) for microbial compositions. To find differential abundances between study groups, the ALDEx2[19] R package was used. For differential abundance analyses, only OTUs present in at least 5% of all samples were included to allow for more meaningful comparisons. ALDEx2 was used to compare worst histological grades of BE as an ordinal variable in a generalized linear model and to assess correlation of BE-associated OTUs with neoplastic progression using aldex.corr to treat worst histological grade as a continuous variable. ALDEx2 was also used to find significance for differentially abundant taxa in a multivariate model with both advanced neoplasia and tooth loss.

Generalized linear models were used to assess differential relative abundance of bacterial taxa in advanced neoplasia, adjusted for tooth loss. Multivariable logistic regression was performed to detect associations between advanced neoplasia and microbiome composition (represented by its top five principal coordinates), adjusted for EAC risk factors (age, sex, race, BMI, smoking, GERD). Supervised machine learning was used to classify patients with advanced neoplasia using the LightGBM package.[57] Three models were created: 1) EAC risk factors alone (age, sex, race, BMI, smoking, GERD); 2) microbiome features alone; and 3) EAC risk factors and microbiome features together. Model parameters were optimized per fold in 10-fold cross-validation, with strict train-test sterility. The output of the models were predicted probabilities of whether

a patient has advanced neoplasia or no BE, with the goal of identifying the patients at highest risk of mortality from EAC.

All statistical analyses were performed in Python or R. Statistical significance was defined as $p<0.05$. Differential abundance analyses were corrected for multiple comparisons using the Benjamini-Hochberg procedure, and corrected statistical significance was defined as $p<0.1$. 95% confidence intervals for AUCs were calculated using the DeLong method using pROC.[58]

Data Availability: 16S rRNA gene sequencing files were uploaded to NCBI Sequence Read Archive (PRJNA785879).

References:

1. Abrams JA, Sharaiha RZ, Gonsalves L, Lightdale CJ, Neugut AI. Dating the rise of esophageal adenocarcinoma: analysis of Connecticut Tumor Registry data, 1940-2007. *Cancer Epidemiol Biomarkers Prev*. Jan 2011;20(1):183-6. doi:10.1158/1055-9965.EPI-10-0802
1055-9965.EPI-10-0802 [pii]
2. Hur C, Miller M, Kong CY, et al. Trends in esophageal adenocarcinoma incidence and mortality. *Cancer*. Mar 15 2013;119(6):1149-58. doi:10.1002/cncr.27834
3. Kong CY, Nattinger KJ, Hayeck TJ, et al. The impact of obesity on the rise in esophageal adenocarcinoma incidence: estimates from a disease simulation model. *Cancer Epidemiol Biomarkers Prev*. Nov 2011;20(11):2450-6. doi:10.1158/1055-9965.EPI-11-0547
4. Friedenberg FK, Hanlon A, Vanar V, et al. Trends in gastroesophageal reflux disease as measured by the National Ambulatory Medical Care Survey. *Dig Dis Sci*. Jul 2010;55(7):1911-7. doi:10.1007/s10620-009-1004-0
5. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. *Cell host & microbe*. Aug 14 2013;14(2):195-206. doi:10.1016/j.chom.2013.07.012
6. Tomkovich S, Dejea CM, Winglee K, et al. Human colon mucosal biofilms from healthy or colon cancer hosts are carcinogenic. *J Clin Invest*. Mar 11 2019;130:1699-1712. doi:10.1172/JCI124196
7. Kadosh E, Snir-Alkalay I, Venkatachalam A, et al. The gut microbiome switches mutant p53 from tumour-suppressive to oncogenic. *Nature*. Oct 2020;586(7827):133-138. doi:10.1038/s41586-020-2541-0
8. Pleguezuelos-Manzano C, Puschhof J, Rosendahl Huber A, et al. Mutational signature in colorectal cancer caused by genotoxic pks(+) E. coli. *Nature*. Apr 2020;580(7802):269-273. doi:10.1038/s41586-020-2080-8
9. Gopalakrishnan V, Spencer CN, Nezi L, et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science*. Jan 5 2018;359(6371):97-103. doi:10.1126/science.aan4236
10. Matson V, Fessler J, Bao R, et al. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science*. Jan 5 2018;359(6371):104-108. doi:10.1126/science.aao3290
11. Routy B, Le Chatelier E, Derosa L, et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science*. Jan 5 2018;359(6371):91-97. doi:10.1126/science.aan3706
12. Elliott DRF, Walker AW, O'Donovan M, Parkhill J, Fitzgerald RC. A non-endoscopic device to sample the oesophageal microbiota: a case-control study. *Lancet Gastroenterol Hepatol*. Jan 2017;2(1):32-42. doi:10.1016/s2468-1253(16)30086-3
13. Snider EJ, Compres G, Freedberg DE, et al. Alterations to the Esophageal Microbiome Associated with Progression from Barrett's Esophagus to Esophageal Adenocarcinoma. *Cancer Epidemiol Biomarkers Prev*. Aug 29 2019;doi:10.1158/1055-9965.EPI-19-0008

14. Annavajhala MK, May M, Compres G, et al. Relationship of the Esophageal Microbiome and Tissue Gene Expression and Links to the Oral Microbiome: A Randomized Clinical Trial. *Clinical and Translational Gastroenterology*. 2020;11(12):e00235. doi:10.14309/ctg.0000000000000235
15. Peters BA, Wu J, Pei Z, et al. Oral Microbiome Composition Reflects Prospective Risk for Esophageal Cancers. *Cancer research*. Dec 1 2017;77(23):6777-6787. doi:10.1158/0008-5472.CAN-17-1296
16. Snider EJ, Compres G, Freedberg DE, et al. Barrett's esophagus is associated with a distinct oral microbiome. *Clin Transl Gastroenterol*. Feb 20 2018;9(3):135. doi:10.1038/s41424-018-0005-8
17. Bornigen D, Ren B, Pickard R, et al. Alterations in oral bacterial communities are associated with risk factors for oral and oropharyngeal cancer. *Sci Rep*. Dec 15 2017;7(1):17686. doi:10.1038/s41598-017-17795-z
18. Lo CH, Kwon S, Wang L, et al. Periodontal disease, tooth loss, and risk of oesophageal and gastric adenocarcinoma: a prospective study. *Gut*. Mar 2021;70(3):620-621. doi:10.1136/gutjnl-2020-321949
19. Fernandes AD, Reid JNS, Macklaim JM, McMurrough TA, Edgell DR, Gloor GB. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome*. 2014/05/05 2014;2(1):15. doi:10.1186/2049-2618-2-15
20. Snider EJ, Compres G, Freedberg DE, et al. Alterations to the Esophageal Microbiome Associated with Progression from Barrett's Esophagus to Esophageal Adenocarcinoma. *Cancer Epidemiol Biomarkers Prev*. Oct 2019;28(10):1687-1693. doi:10.1158/1055-9965.Epi-19-0008
21. Hall-Stoodley L, Costerton JW, Stoodley P. Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol*. Feb 2004;2(2):95-108. doi:10.1038/nrmicro821
22. Froeliger EH, Fives-Taylor P. Streptococcus parasanguis fimbria-associated adhesin fap1 is required for biofilm formation. *Infect Immun*. Apr 2001;69(4):2512-9. doi:10.1128/IAI.69.4.2512-2519.2001
23. Lee SF, Li YH, Bowden GH. Detachment of Streptococcus mutans biofilm cells by an endogenous enzymatic activity. *Infect Immun*. Mar 1996;64(3):1035-8. doi:10.1128/iai.64.3.1035-1038.1996
24. Certo M, Tsai CH, Pucino V, Ho PC, Mauro C. Lactate modulation of immune responses in inflammatory versus tumour microenvironments. *Nat Rev Immunol*. Mar 2021;21(3):151-161. doi:10.1038/s41577-020-0406-2
25. Sullivan LB, Gui DY, Hosios AM, Bush LN, Freinkman E, Vander Heiden MG. Supporting Aspartate Biosynthesis Is an Essential Function of Respiration in Proliferating Cells. *Cell*. Jul 30 2015;162(3):552-63. doi:10.1016/j.cell.2015.07.017
26. O'Keefe SJ. Diet, microorganisms and their metabolites, and colon cancer. *Nature reviews Gastroenterology & hepatology*. Dec 2016;13(12):691-706. doi:10.1038/nrgastro.2016.165
27. Bultman SJ. Molecular pathways: gene-environment interactions regulating dietary fiber induction of proliferation and apoptosis via butyrate for cancer prevention. *Clinical cancer research : an official journal of the American Association for Cancer Research*. Feb 15 2014;20(4):799-803. doi:10.1158/1078-0432.CCR-13-2483

28. Smith PM, Howitt MR, Panikov N, et al. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science*. Aug 2 2013;341(6145):569-73. doi:10.1126/science.1241165
29. Papadimitriou N, Gunter MJ, Murphy N, et al. Circulating tryptophan metabolites and risk of colon cancer: Results from case-control and prospective cohort studies. *International journal of cancer Journal international du cancer*. Jul 1 2021;doi:10.1002/ijc.33725
30. Majka J, Wierdak M, Brzozowska I, et al. Melatonin in Prevention of the Sequence from Reflux Esophagitis to Barrett's Esophagus and Esophageal Adenocarcinoma: Experimental and Clinical Perspectives. *Int J Mol Sci*. Jul 13 2018;19(7)doi:10.3390/ijms19072033
31. Annavajhala MK, May M, Compres G, et al. Relationship of the Esophageal Microbiome and Tissue Gene Expression and Links to the Oral Microbiome: A Randomized Clinical Trial. *Clin Transl Gastroenterol*. Dec 2020;11(12):e00235. doi:10.14309/ctg.0000000000000235
32. Krautkramer KA, Fan J, Backhed F. Gut microbial metabolites as multi-kingdom intermediates. *Nat Rev Microbiol*. Feb 2021;19(2):77-94. doi:10.1038/s41579-020-0438-4
33. Vogtmann E, Hua X, Zhou L, et al. Temporal Variability of Oral Microbiota Over 10 Months and the Implications for Future Epidemiologic Studies. *Cancer Epidemiol Biomarkers Prev*. Feb 23 2018;doi:10.1158/1055-9965.EPI-17-1004
34. Belstrom D, Holmstrup P, Bardow A, Kokaras A, Fiehn NE, Paster BJ. Temporal Stability of the Salivary Microbiota in Oral Health. *PLoS One*. 2016;11(1):e0147472. doi:10.1371/journal.pone.0147472
35. Cameron SJ, Huws SA, Hegarty MJ, Smith DP, Mur LA. The human salivary microbiome exhibits temporal stability in bacterial diversity. *FEMS Microbiol Ecol*. Sep 2015;91(9):fiv091. doi:10.1093/femsec/fiv091
36. Caporaso JG, Lauber CL, Costello EK, et al. Moving pictures of the human microbiome. *Genome biology*. 2011;12(5):R50. doi:10.1186/gb-2011-12-5-r50
37. Cabral DJ, Wurster JI, Flokas ME, et al. The salivary microbiome is consistent between subjects and resistant to impacts of short-term hospitalization. *Sci Rep*. Sep 8 2017;7(1):11040. doi:10.1038/s41598-017-11427-2
38. Belstrom D, Holmstrup P, Nielsen CH, et al. Bacterial profiles of saliva in relation to diet, lifestyle factors, and socioeconomic status. *J Oral Microbiol*. 2014;6doi:10.3402/jom.v6.23609
39. Kato I, Vasquez A, Moyerbrailean G, et al. Nutritional Correlates of Human Oral Microbiome. *J Am Coll Nutr*. Feb 2017;36(2):88-98. doi:10.1080/07315724.2016.1185386
40. Nearing JT, DeClercq V, Van Limbergen J, Langille MGI. Assessing the Variation within the Oral Microbiome of Healthy Adults. *mSphere*. Sep 30 2020;5(5)doi:10.1128/mSphere.00451-20
41. Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT, Quince C. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res*. Mar 31 2015;43(6):e37. doi:10.1093/nar/gku1341
42. Edgar RC. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv*. 2016:074161. doi:10.1101/074161

43. F. Escapa I, Huang Y, Chen T, et al. Construction of habitat-specific training sets to achieve species-level assignment in 16S rRNA gene datasets. *Microbiome*. 2020/05/15 2020;8(1):65. doi:10.1186/s40168-020-00841-w
44. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*. Aug 2007;73(16):5261-7. doi:10.1128/aem.00062-07
45. Baldini F, Heinken A, Heirendt L, Magnusdottir S, Fleming RMT, Thiele I. The Microbiome Modeling Toolbox: from microbial interactions to personalized microbial communities. *Bioinformatics*. Jul 1 2019;35(13):2332-2334. doi:10.1093/bioinformatics/bty941
46. Heirendt L, Arreckx S, Pfau T, et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat Protoc*. Mar 2019;14(3):639-702. doi:10.1038/s41596-018-0098-2
47. Magnusdottir S, Heinken A, Kutt L, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature biotechnology*. Jan 2017;35(1):81-89. doi:10.1038/nbt.3703
48. Baldini F, Hertel J, Sandt E, et al. Parkinson's disease-associated alterations of the gut microbiome predict disease-relevant changes in metabolic functions. *BMC biology*. Jun 9 2020;18(1):62. doi:10.1186/s12915-020-00775-7
49. Klitgord N, Segre D. Environments that induce synthetic microbial ecosystems. *PLoS computational biology*. Nov 18 2010;6(11):e1001002. doi:10.1371/journal.pcbi.1001002
50. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? *Nature biotechnology*. Mar 2010;28(3):245-8. doi:10.1038/nbt.1614
51. Noronha A, Modamio J, Jarosz Y, et al. The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic Acids Res*. Jan 8 2019;47(D1):D614-D624. doi:10.1093/nar/gky992
52. Shaheen NJ, Falk GW, Iyer PG, Gerson LB. ACG Clinical Guideline: Diagnosis and Management of Barrett's Esophagus. *Am J Gastroenterol*. Jan 2016;111(1):30-50; quiz 51. doi:10.1038/ajg.2015.322
53. Vennalaganti P, Kanakadandi V, Goldblum JR, et al. Discordance Among Pathologists in the United States and Europe in Diagnosis of Low-Grade Dysplasia for Patients With Barrett's Esophagus. *Gastroenterology*. Feb 2017;152(3):564-570.e4. doi:10.1053/j.gastro.2016.10.041
54. Wani S, Falk GW, Post J, et al. Risk factors for progression of low-grade dysplasia in patients with Barrett's esophagus. *Gastroenterology*. Oct 2011;141(4):1179-86, 1186.e1. doi:10.1053/j.gastro.2011.06.055
55. Wani S, Rubenstein JH, Vieth M, Bergman J. Diagnosis and Management of Low-Grade Dysplasia in Barrett's Esophagus: Expert Review From the Clinical Practice Updates Committee of the American Gastroenterological Association. *Gastroenterology*. Nov 2016;151(5):822-835. doi:10.1053/j.gastro.2016.09.040
56. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. *Isme j*. Feb 2011;5(2):169-72. doi:10.1038/ismej.2010.133

57. Ke G, Meng Q, Finley T, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In: Guyon I, Luxburg UV, Bengio S, et al, eds. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. 2017.
58. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. Mar 17 2011;12:77. doi:10.1186/1471-2105-12-77

DECLARATIONS

Ethical approval and consent to participate: This study was approved by the Columbia University Institutional Review Board. All patients provided written informed consent.

Consent for publication: Not applicable

Availability of data and materials: 16S rRNA gene sequencing files were uploaded to NCBI Sequence Read Archive (PRJNA785879).

Competing interests: The authors have none to disclose.

Funding: This study was supported in part by the National Cancer Institute (U54 CA163004; R01 CA238433) and the Digestive Disease Research Foundation. T.K. is a CIFAR Azrieli Global Scholar in the Humans & the Microbiome Program. DEF was supported in part by a Department of Defense Peer Reviewed Medical Research Program Clinical Trial Award (PR181960) and by a Columbia University Irving Scholar Award.

Authors' contributions: Quinn Solfisburg performed data analysis, data interpretation, and original drafting of the manuscript. Federico Baldini performed data analysis, data interpretation, and assisted with drafting the manuscript. Brittany Baldwin Hunter performed study conduct and provided critical input to the manuscript. Harry Lee performed data analysis and provided critical input to the manuscript. Daniel Freedberg performed data interpretation and provided critical input to the manuscript. Charles Lightdale performed study conduct and provided critical input to the manuscript. Tal Korem supervised data analysis, performed data interpretation, and assisted with drafting the manuscript. Julian Abrams designed the study, supervised data analysis, performed data interpretation, and assisted with drafting the manuscript. All authors approved the final version of the manuscript.

Acknowledgments: None

Authors' information: See title page.