Oral Microbiome Profiling in Smokers with and without Head and Neck Cancer Reveals Variations Between Health and Disease

Ashok Kumar Sharma1, William T. DeBusk2, Irina Stepanov3, Andres Gomez1, and Samir S. Khariwala2

ABSTRACT

While smoking is inextricably linked to oral/head and neck cancer (HNSCC), only a small fraction of smokers develop HNSCC. Thus, we have sought to identify other factors, which may influence the development of HNSCC in smokers including microbiology. To determine microbial associations with HNSCC among tobacco users, we characterized oral microbiome composition in smokers with and without HNSCC. 16S rRNA MiSeq sequencing was used to examine the oral mucosa microbiome of 27 smokers with (cases) and 24 without HNSCC (controls). In addition, we correlated previously reported levels of DNA damage with the microbiome data. Smokers with HNSCC showed lower microbiome richness compared with controls (q = 0.012). Beta-diversity analyses, assessed as UniFrac (weighted and unweighted) and Bray–Curtis distances, showed significant differences in oral mucosal microbiome signatures between smokers with and without HNSCC (2, 3). In this work, we identified higher rates of DNA damage in those smokers with HNSCC compared with cancer-free smokers. However, tobacco carcinogen exposure was similar in both groups such that we sought another explanation for varying levels of DNA damage in the oral cavity.

Introduction

Oral/head and neck cancer (HNSCC) represents a group of tumors strongly associated with tobacco use. While human papillomavirus (HPV) has more recently been identified as a frequent cause of oropharyngeal cancers, most head and neck tumors strongly associated with tobacco use. While smoking is inextricably linked to oral/head and neck cancer (HNSCC), only a small fraction of smokers develop HNSCC. Thus, we have sought to identify other factors, which may influence the development of HNSCC in smokers including microbiology. To determine microbial associations with HNSCC among tobacco users, we characterized oral microbiome composition in smokers with and without HNSCC. 16S rRNA MiSeq sequencing was used to examine the oral mucosa microbiome of 27 smokers with (cases) and 24 without HNSCC (controls). In addition, we correlated previously reported levels of DNA damage with the microbiome data. Smokers with HNSCC showed lower microbiome richness compared with controls (q = 0.012). Beta-diversity analyses, assessed as UniFrac (weighted and unweighted) and Bray–Curtis distances, showed significant differences in oral mucosal microbiome signatures between smokers with and without HNSCC (2, 3). In this work, we identified higher rates of DNA damage in those smokers with HNSCC compared with cancer-free smokers. However, tobacco carcinogen exposure was similar in both groups such that we sought another explanation for varying levels of DNA damage in the oral cavity.

One possible explanation for the varying levels of oral cell DNA damage found in our prior work pertains to the oral microbiome. The impact of the bacterial microbiome on human disease has been explored and discussed extensively (4). Specifically, it has been hypothesized that specific microbiome fingerprints in susceptible individuals may influence disease phenotypes in a variety of infectious, metabolic, and immune disorders as well as in cancer (5, 6). In cancer, the role of microorganisms in impacting disease onset and progression has been deemed as central (7), including such examples as the role of Helicobacter pylori, Salmonella typhi, and Chlamydia pneumoniae in gastric, gallbladder, and lung cancer, respectively.

Because of its multiple risk factors and associated disorders, multifactorial nature, and the highly diverse oral micro ecosystem (second in diversity after the colonic microbiome), studying bacterial carcinogens in the oral cavity poses...
significant challenges. For instance, culture-dependent, and narrow-range molecular techniques (e.g., checkerboard hybridization) initially showed differences in the abundance of specific oral commensals in saliva of patients affected with HNSCC compared with healthy subjects (8). In addition, cancers in HNSCC have shown greater abundance and diversity of culturable bacteria (aerobes and anaerobes) in contrast with normal tissues (9, 10). However, the use of next-generation sequencing techniques to profile microbiomes has revealed novel microbial dynamics in HNSCC. Given this information, we used 16S rRNA MiSeq to profile the oral microbiome in a group of smokers with HNSCC who have demonstrated higher levels of DNA damage than smokers without HNSCC.

Materials and Methods

Study subjects

This study was approved by the University of Minnesota Research Subjects’ Protection Programs Institutional Review Board: Human Subjects Committee (IRB Study # 0903M62203). Patients with HNSCC were identified and enrolled following written informed consent during outpatient visits to the University of Minnesota Otolaryngology-Head and Neck Surgery Clinic with a new diagnosis of squamous cell carcinoma of the upper aerodigestive tract. This included tumors of the oral cavity, oropharynx, larynx, and hypopharynx. In some cases, cancers were first identified in our clinic whereas in others, cancers were diagnosed at outside institutions and referred to our clinic for definitive management. Inclusion criteria included self-report of current daily smoking and having smoked at least five cigarettes per day for at least 5 years. A total of 66 cases (smokers with HNSCC) and 51 controls (smokers without HNSCC) were enrolled in the study from February 2014 to May 2017. DNA sufficient for analysis was available in 27 cases and 24 controls.

Cancer-free controls were recruited in the same outpatient clinic. The control subjects were visiting the clinic for clinical evaluation of problems other than cancer (i.e., sinusitis, hearing loss) and were approached for enrollment upon identification as daily cigarette smokers. All enrolled subjects were smoking cigarettes daily at the time of enrollment. Demographic data collected included cigarettes per day, duration of use, alcoholic drinks per day, alcoholic frequency, and tumor-related variables such as subsite and stage.

Buccal cell collection

Oral brushings from the buccal mucosa were collected from enrolled subjects through the Department of Otolaryngology-Head and Neck Surgery at the University of Minnesota (Minneapolis, MN). Oral cells were collected by brushing the oral mucosa inside one cheek with a clean toothbrush and swirling the brush in a sterile polypropylene centrifuge tube with a commercial mouthwash to transfer the collected buccal cells from the brush into the liquid. After the collection, the samples were centrifuged at 2,700 rpm to pellet cells; the pellets were washed with Tris-EDTA buffer (pH 7.4) and stored at –20°C until DNA isolation and analysis.

Microbiome analyses

Extraction of genomic DNA from each brushing was performed using the DNA Purification Kit (Qiagen). After DNA integrity was measured, high-quality DNA of 27 cases and 24 controls was used for oral bacteria community profiling through 16S rRNA amplicon sequencing, targeting the V4 hyper variable region (barcode primer pair 515f-GTGCCAGCMGCGCGGTAA and 806r- GGACTACHVGGGTWTCTAAT) on the Illumina MiSeq sequencing platform. Raw reads were trimmed to remove primers using cutadapt, and filtered to remove low-quality reads (less than Q = 30) using fastx_toolkit. High-quality reads were considered for downstream analysis using the DADA2 plugin within qiime2 (11), which performs denoising, merging of paired-end reads, and removal of chimeric sequences to produce unique amplicon sequence variants (ASV). Taxonomic assignment of these ASVs was carried out by the trained naive bayes classifier on reference sequences (clustered at 99% sequence identity) from Greengenes 13_8, using feature-classifier fit-classifier-naive-bayes, and feature-classifier classify-sklernk plugins within qiime2. ASV abundances were used for functional prediction analyses using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt, version 2; maximum nearest sequenced taxon index = 2), and predicted KEGG pathway relative abundances were used for further analysis.

DNA adduct analysis

The process for DNA adduct quantification has been described previously, and values for these samples were reported separately (2, 3). Briefly, DNA was isolated from the collected samples by using the commercial DNA Purification Kit (Qiagen). The isolated DNA samples were subjected to acid hydrolysis to release 4-hydroxy-1-(3-pyridyl)-1-butonanone (HPB) and purified on 25-mg HyperSep Hypercarb Artridges (Thermo Fisher Scientific). The analysis of HPB in the purified samples was carried out on an LTQ Orbitrap Velos Instrument (Thermo Fisher Scientific) interfaced with a Nano2D-LC HPLC (Eksigent) system with nanoelectrospray ionization.

Statistical analyses

All microbial community ecology analyses were performed within the R statistical interface (12). Relative abundances of each ASV on complete and rarefied data (depth = 1,000 sequences) along with the rooted tree (generated using the alignment mafft plugin on qiime2 and representative sequences) were used for the calculation of distance matrices (Bray–Curtis and UniFrac), ordination analysis [principal coordinates analysis (PCoA)], and alpha diversity indices (observed taxa and Shannon’s H indices) using the R phyloseq package (13). Permutational multivariate analysis of variance (PERMANOVA) was calculated using the adonis function.
within the R vegan package (14). Random forest (RF) classification models at ntree = 500 and cv = 10 folds were constructed in the random Forest R package (15) to assess classification accuracy of cases and controls based on microbial taxonomic features and predicted KEGG pathways. These RF classification models were used to select potential discriminating taxa and predicted pathways based on their respective mean decrease in accuracy observations. In addition, indicator species analyses within the labdsv package (16) and Wilcoxon rank-sum test, always correcting for FDR with an acceptable threshold of 0.05, were applied to check for discriminatory power and statistical significance of all discriminant features. Strength and direction of association between (i) significantly discriminating taxa, (ii) between HPB-releasing adducts levels and significantly discriminating taxa, and (iii) between predicted pathways and significantly discriminating taxa was measured using Spearman correlations coefficients within the R psych package (http://www.test.personality-project.org/r/psych/psych-manual.pdf; ref. 17). All graphs were plotted using the ggplots R package (18), whereas correlation networks were generated and visualized using Cytoscape v3.7.1 (19). All analysis codes along with the data files are available on https://github.com/ashoks773/Oral-microbiome-16S.

Results

Demographic data and tobacco use

Our study population consisted of 24 cigarette smokers (controls) and 27 smokers who had developed tumors of the upper aerodigestive tract (cases). Demographics data are summarized in Table 1. Mean age of cases was 58 years, and the mean age of controls was 48 years. Eighty-three percent of cases were male, and 57% of controls were male. Tumor site distribution was oropharynx (12/27, 44%), larynx (6/27, 22%), oral cavity (6/27, 22%), hypopharynx (2/27, 7%), and other (1/27, 3%). Univariate analysis revealed both groups reported the same level of cigarette exposure (14 cigarettes/day). Mean duration of cigarette use was 22.3 years for cases and 23.6 years for controls.

Oral microbiome profiles in smokers with and without head and neck cancer

A total of 1,544,643 16S rRNA sequence reads were obtained, 1,317,658 of which remained after quality filtering, reflecting 1,463 unique ASVs, each representing a unique taxon. After taxonomic assignment, we obtained an average sequencing depth of 22,082 reads per sample (range = 537–62532, SD = 15482; Supplementary Table S1). Lower microbiome richness (number of observed ASVs) was observed in smokers with head and neck cancer, compared with controls (Wilcoxon rank–sum tests, q = 0.012), and using both all reads (Fig. 1A) or a set of 1,000 reads randomly selected in each sample to control for sequencing depth differences (Supplementary Fig. S1). Although these trends were maintained, no differences were observed in terms of the Shannon’s H diversity index (q > 0.05; Fig. 1A).

Beta-diversity analyses assessed as UniFrac (weighted and unweighted) and Bray–Curtis distances, showed significant differences in oral mucosal microbiome signatures between cases and controls (PERMANOVA, weighted UniFrac: r² = 0.04, P = 0.05; unweighted UniFrac: r² = 0.05, P = 0.001; and Bray–Curtis distances: r² = 0.03, P = 0.03), along axis 1 or 2 of a PCoA ordination. Ordination scores differed significantly depending on distance chosen (Fig. 1B; q < 0.01, Wilcoxon rank–sum test), with unweighted UniFrac and Bray–Curtis distances showing the most discriminant patterns. In addition, the potential effect of alcohol consumption on oral mucosal microbiome signatures was explored, which was found not significant (Supplementary Table S2). To overcome potential sequencing depth biases across samples, we rarefied the data at 1,000 reads per sample, observing the same beta diversity patterns as when using all reads generated (Supplementary Fig. S2A–S2C). Beta diversity analyses also showed significantly higher interindividual microbiome variation among smokers with head and neck cancer compared with controls, demonstrating high oral mucosal microbiome heterogeneity under disease conditions (Fig. 1C; q ≤ 0.01, Wilcoxon rank–sum test).

A RF classification model was used to identify the discriminatory taxonomic features (using mean decrease accuracy power and statistical significance of all discriminant features. Strength and direction of association between (i) significantly discriminating taxa, (ii) between HPB-releasing adducts levels and significantly discriminating taxa, and (iii) between predicted pathways and significantly discriminating taxa was measured using Spearman correlations coefficients within the R psych package (http://www.test.personality-project.org/r/psych/psych-manual.pdf; ref. 17). All graphs were plotted using the ggplots R package (18), whereas correlation networks were generated and visualized using Cytoscape v3.7.1 (19). All analysis codes along with the data files are available on https://github.com/ashoks773/Oral-microbiome-16S.

Table 1. Demographics and characteristics of subjects classified as case and control.

<table>
<thead>
<tr>
<th>Case</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>27</td>
</tr>
<tr>
<td>Male/Female</td>
<td>22 (81%)/5 (19%)</td>
</tr>
<tr>
<td>Age (mean)</td>
<td>58.1 years</td>
</tr>
<tr>
<td>Tumor site</td>
<td></td>
</tr>
<tr>
<td>Oral cavity</td>
<td>6 (22%)</td>
</tr>
<tr>
<td>Oropharynx</td>
<td>12 (44%)</td>
</tr>
<tr>
<td>Larynx</td>
<td>6 (22%)</td>
</tr>
<tr>
<td>Hypopharynx</td>
<td>2 (7%)</td>
</tr>
<tr>
<td>N/A</td>
<td>1 (4%)</td>
</tr>
</tbody>
</table>

**AJCCstage**

<table>
<thead>
<tr>
<th>Case</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>3</td>
</tr>
<tr>
<td>Stage II</td>
<td>5</td>
</tr>
<tr>
<td>Stage III</td>
<td>4</td>
</tr>
<tr>
<td>Stage IV</td>
<td>15</td>
</tr>
</tbody>
</table>

**Alcohol use**

<table>
<thead>
<tr>
<th>Case</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>7</td>
</tr>
<tr>
<td>Monthly</td>
<td>2</td>
</tr>
<tr>
<td>2-4 ×/month</td>
<td>3</td>
</tr>
<tr>
<td>&gt;4 ×/month</td>
<td>13</td>
</tr>
<tr>
<td>N/A</td>
<td>2</td>
</tr>
</tbody>
</table>

**Cigarettes per day (mean)**

<table>
<thead>
<tr>
<th>Case</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.9</td>
<td>13.8</td>
</tr>
</tbody>
</table>

**Smoking duration years (mean)**

<table>
<thead>
<tr>
<th>Case</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>22.3</td>
<td>23.6</td>
</tr>
</tbody>
</table>

**Mean total urinary cotinine (ng/mL)**

<table>
<thead>
<tr>
<th>Case</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,609.8</td>
<td>1,817.4</td>
</tr>
</tbody>
</table>

Beta-diversity analyses assessed as UniFrac (weighted and unweighted) and Bray–Curtis distances, showed significant differences in oral mucosal microbiome signatures between cases and controls (PERMANOVA, weighted UniFrac: r² = 0.04, P = 0.05; unweighted UniFrac: r² = 0.05, P = 0.001; and Bray–Curtis distances: r² = 0.03, P = 0.03), along axis 1 or 2 of a PCoA ordination. Ordination scores differed significantly depending on distance chosen (Fig. 1B; q < 0.01, Wilcoxon rank–sum test), with unweighted UniFrac and Bray–Curtis distances showing the most discriminant patterns. In addition, the potential effect of alcohol consumption on oral mucosal microbiome signatures was explored, which was found not significant (Supplementary Table S2). To overcome potential sequencing depth biases across samples, we rarefied the data at 1,000 reads per sample, observing the same beta diversity patterns as when using all reads generated (Supplementary Fig. S2A–S2C). Beta diversity analyses also showed significantly higher interindividual microbiome variation among smokers with head and neck cancer compared with controls, demonstrating high oral mucosal microbiome heterogeneity under disease conditions (Fig. 1C; q ≤ 0.01, Wilcoxon rank–sum test).

A RF classification model was used to identify the discriminatory taxonomic features (using mean decrease accuracy
index) between the oral microbiome of head and neck cancer smokers and healthy controls. The classification performance of the RF model constructed on the training set (n = 36, 70% of data) was accessed on a test set (n = 16, 30% of the data). The results showed that cases and controls could be moderately classified on the basis of their oral mucosal microbiome taxonomic features (with error rate of 23.52% for cases and 27.77% for controls and AUC of 0.75 indicating moderate accuracy and sensitivity of the model, at 500 ntrees; Fig. 2A).

The top 60 taxonomic features, according to mean decrease in accuracy (>1.14), were selected and then checked using indicator species analyses (indicator value > 0.5) and statistical significance according to FDR-adjusted Wilcoxon rank-sum test, q < 0.05. These analyses revealed 11 significant taxonomic markers distinguishing cases from controls, with relative abundances of *Stenotrophomonas* and *Ruminococcus*, and the family *Comamonadaceae* higher in smokers with head and neck cancer (Fig. 2B; Supplementary Table S3). In contrast, the remaining eight taxonomic markers were highly abundant in controls; among them, common oral commensals such as *Tannerella*, *Capnocytophaga*, *Selenomonas*, *Veillonella*, and *Kingella* were highlighted.

Spearman correlation analyses were used to explore coabundance and exclusion patterns between discriminating taxa. This analysis revealed that the markers that characterized cases — *Ruminococcus*, *Stenotrophomonas*, and *Comamonadaceae*, tended to coabund together (Fig. 2C), while showing exclusion patterns with *Weeksellaceae* (r = −0.55, and q = 0.0001) and *Capnocytophaga* (r = −0.51, and q = 0.0002), markers characterizing controls. These two control-specific taxa showed significant coabundance with other oral commensals more abundant in controls such as *Haemophilus*, *Kingella*, *Tannerella*, *Selenomonas*, and *Veillonella* (Fig. 2B; Supplementary Table S4).

**Correlation of HPB-releasing DNA adducts with microbiome profiles**

High levels of HPB-releasing DNA adducts were observed in buccal cells of cases compared with the controls (Fig. 3A); therefore, we explored whether these adducts were associated with distinct microbiome markers. For instance, an increase in HPB-releasing DNA adduct levels in buccal cells was negatively associated with the number of observed bacterial taxa on the same mucosal site (Fig. 3B; Spearman correlation, ρ = −0.38,
In addition, it was observed that the adduct levels tended to increase along with the cumulative abundance (sum of relative abundances) of *Stenotrophomonas*, *Ruminococcus*, and *Comamonadaceae*, the three taxonomic markers distinguishing cases from controls (Spearman correlation, $r = 0.32$, $q = 0.02$; Fig. 3C). The cumulative abundance of taxonomic signatures characterizing controls, and that are mostly common oral commensals, tended to decrease with increasing DNA adduct levels; indicating that increased tobacco-induced DNA damage in buccal cells may be associated with lower abundance of this commensals; although this relationship did not show significance at alpha 0.05 (Spearman correlation, $r = 0.34$, $q = 0.08$).
correlation, $\rho = -0.24$, $q = 0.08$; Fig. 3C). However, when analyzing correlations between the abundance of individual taxa and HPB-releasing adduct levels, *Stenotrophomonas* (enriched in cases) showed positive correlations (Spearman $r = 0.36$, $q = 0.008$), whereas *Capnocytophaga*, *Kingella*, *Veillonella*, and *Weeksellaceae*, markers of healthy controls, showed a negative relationship with the levels of these DNA adducts (Spearman $r = -0.27$ to $-0.33$, and $q < 0.05$; Fig. 3D). Remaining taxa enriched in cases and controls also showed nonsignificant ($q > 0.05$) positive and negative correlations with HPB-releasing DNA adduct levels and hence were not included in the network (Supplementary Table S5).

**Correlation between predicted pathways and microbiome profiles**

Results from the functional prediction analysis via PICRUSt were used to detect predicted discriminatory pathways between the oral microbiome of head and neck cancer smokers and healthy controls through an RF classification model (ntrees = 500). The top 100 pathway features, according to mean decrease in accuracy ($>1.08$), were selected and then checked using statistical significance according to the Wilcoxon rank-sum test, $q < 0.05$. This analysis revealed a higher number of predicted functional pathways enriched in cases (16 pathways out of total 26), mainly involved in the degradation of...
xenobiotics (e.g., vanillin, hydroxyacetophenone, phenolics, and toluene) amino acids and amines (arginine, histidine, polygenic amines, and putrescine), antibiotic resistance (polymyxin), and other biosynthetic and metabolic pathways (Supplementary Fig. S3; Supplementary Table S6). These selected pathways showed positive correlations (Spearman $r > 0.5$, $q < 0.05$) with all three taxa enriched in cases (Fig. 4; Supplementary Table S7). In contrast, fewer pathways were enriched in controls (10 pathways out of total 26), and most of them were associated with the biosynthesis of sugars, lipids and amino acids and degradation of sugars such as xylose, galacturonate, and arabinose (Supplementary Fig. S3; Supplementary Table S6). These predicted pathways showed positive correlations (Spearman $r > 0.5$, $q < 0.05$) with Tannerella, Veillonella, Weksellaceae, Heamophilus parainfluenzae, and Lachnospiraceae (Fig. 4; Supplementary Table S7); all taxonomic markers characterizing controls.

**Discussion**

In this study, we examined the oral microbiome profiles of cigarette smokers, evaluating for a propensity toward development of HNSCC. Broadly, microbiome profiling of 27 smokers affected with HNSCC and 24 controls, via MiSeq amplicon sequencing of the bacterial 16S rRNA gene, indicated significant differences between the cohorts based on bacterial
community composition. Likewise, we report higher interindividual variability and lower bacterial richness in the oral mucosa of patients with HNSCC, as well as specific taxonomic and predicted functional markers distinguishing diseased from healthy states, likely correlated with DNA damage levels. These results add to growing evidence linking oral microbial communities to HNSCC status, highlighting the value of microbial markers for disease diagnosis and shedding light on the microbial mechanisms likely associated with disease risk.

A handful of reports have previously described associations between the oral microbiome and HNSCC also using next-generation sequencing techniques. These reports have focused on multiple traits such as tumor stage, different sources in the oral cavity (e.g., tissue, saliva, and oral rinses), healthy versus tumor tissues in the same subjects and comparisons of patients with cancer with healthy controls. For instance, reports focusing only on a limited number of patients with HNSCC and disease stage show variation in bacterial diversity and taxonomic profiles depending on mutational signatures (20), and site sampled within the oral cavity, with significant differences between saliva and tumor tissue samples (21) However, focusing only on patients with HNSCC did not show clear stage-dependent differences based on community composition (beta diversity) analysis.

In line with our findings, studies in larger cohorts including oral rinses in cancer-free subjects or healthy tissue swabs in the same affected individuals have shown significant compositional differences between healthy and HNSCC subjects (22, 23). In these studies, it is also shown that bacterial richness and diversity tend to increase in oral rinses and tumor tissues of affected individuals, which is not concordant with our observations of decreased alpha diversity in HNSCC cases. Nonetheless, the aforementioned studies, and others showing the same diversity trends in the saliva of a few HNSCC and healthy subjects (24) did not stratify the microbial signals obtained on the basis of smoking status, or other lifestyle risk factors, which can add major confounders. Indeed, smoking is reported to have significant effects on oral microbiome composition (25, 26) and possibly explain the discrepancies in alpha diversity found between ours and those previous studies.

Thus, the results presented herein should be first considered in the context of a tissue microenvironment already predisposed by smoking and the harmful constituents present in cigarette smoke, characteristics that make our study unique in comparison to all the aforementioned reports. For example, it is well known that tobacco smoke results in exposure to a myriad of toxic and carcinogenic constituents, resulting in DNA damage and subsequent development of several malignancies, including oral cancer. The tobacco-specific nitrosamines N' nitrosonornicotine (NNN) and 4-(methylNitrosamino)-1-(-3-pyridyl)-1-butanol (NNK) are two carcinogens that have specifically been shown to have high malignant potential. Exposure to NNN and NNK triggers macromolecular alterations, which interfere with replication, transcription, and normal DNA repair mechanisms. Primary among the effects of nitrosamines is the generation of HPB-releasing DNA adducts following metabolic activation. Measurement of DNA adducts offers a direct assessment of DNA damage. Recent studies have shown that the amount of NNN/NNK exposure, as assessed by urinary biomarkers, does not necessarily correlate to the level of HPB-releasing DNA adducts; therefore, there are likely to be other factors, besides the carcinogen doses, contributing to the variations in the levels of DNA damage among smokers (27–32). The oral microbiome can disrupt host's defense mechanisms, inducing chronic inflammatory changes resulting in a cascade of events capable of causing extensive DNA damage (32–35). It has recently been shown that NNK carcinogenicity, and likely that of NNN, can be substantially increased in the presence of inflammatory agents (36). It is also plausible to expect that certain bacteria are capable of metabolically activating tobacco carcinogens. This raises the question whether specific smokers are at a higher risk of developing oral cancer given a higher baseline level of inflammation or a higher level of carcinogen activation in the oral cavity, thus sparking our interest in investigating the bacterial microbiome of the oral cavity.

Along these lines, we report that predicted bacterial functions associated with the degradation of xenobiotics (e.g., toluene, phenyl compounds) and amines (e.g., aromatic biogenic amines) characterize HNSCC, in contrast with healthy smoker tissues, which exhibited greater abundance of predicted carbohydrate metabolism pathways. Increased abundance of predicted genes involved in xenobiotic degradation pathways, at the expense of carbohydrate metabolism, have also been reported in oral rinses of smokers versus never-smokers in a large cohort (26), which is in line with our comparisons of cases versus healthy controls. As such, one important question focuses on investigating whether exposure to cigarette smoke modulates the oral microbiome to increase the presence of xenobiotox-metabolizing bacteria, or whether the individuals at risk for HNSCC inherently have higher abundance of such bacteria; and on determining why this is not the case in cancer-free tobacco users. Moreover, as patterns of increased HPB-releasing DNA adducts were correlated with specific bacterial signatures in HNSCC smokers, whether such bacteria contribute to metabolic activation of NNN and/or NNK should be investigated.

Notably, some microbial markers in patients with HNSCC that positively correlated with HPB-releasing adduct levels, have been previously characterized as pathogenic, xenobiotic degrader, and/or multidrug-resistant taxa. For instance, Stenotrophomonas is reported to be a nosocomial pathogen, particularly associated with immunosuppressed individuals, and resistance to a broad range of antibiotics (37–39). This observation is in line with our findings of increased abundance of predicted polymyxin resistance in cases, correlated with abundances of this taxon. Stenotrophomonas has been associated with a variety of infections in respiratory tract, blood, bones and joints, urinary tract, soft tissues, and lung cancer (reviewed in ref. 40.) Cytotoxic and protease activities by...
Stenotrophomonas may also degrade tissue and cause cell death (41). Furthermore, in concordance with an increase in predicted aromatic xenobiotic and amine degradation in cases, this taxon has the capacity to process a wide range of toxicants, including phenolics (42), toluene (43), and phenanthrene (44); with metabolites from the latter commonly reported in smokers at lung cancer risk (45).

Although not commonly abundant in the human oral cavity, unclassified Comamonadaceae have been detected before in the oral washes of immunocompromised (HIV+) smokers (46), in high mean abundance (25.5%) on the laryngeal tissues of smokers (47) and in the saliva of obese individuals (48). However, the influence of Stenotrophomonas and Comamonadaceae in HNSCC progression, their ecological interactions, and their potential role in metabolic activation of tobacco-related carcinogens or DNA damage is still unclear and warrants further investigation.

This study also showed that the abundance of common oral commensals (e.g., Tannerella, Capnocytophaga, Selenomonas, Veillonella, and Kingella) that are usually associated with carbohydrate metabolism in the oral cavity (49–52) (as confirmed by the functional predictions presented) are altered or diminished in smokers with HNSCC. Thus, our hypothesis that healthy smokers still exhibit markers of a homeostatic mucosal microbiome, which exerts protective effects against carcinogen activators (e.g., through competitive exclusion), should be explored further. Indeed, partially in line with our results, a recent study analyzing oral washes of a larger cohort (n = 129 cases and n = 245 controls) shows no bacterial taxa associated with HNSCC risk, while identifying higher abundance of some common oral commensals (e.g., Kingella) in controls compared with disease subjects (53). These observations are also connected with our finding that patients with HNSCC showed increased interindividual variation and heterogeneity in the abundance of common oral commensals (higher beta diversity), and lower alpha diversity (reduced richness), inversely correlated with higher HPB-releasing DNA adduct levels. As such, further studies should focus on elucidating the ecological and molecular underpinnings of microbiome-mediated protection against smoke carcinogens in healthy smokers, and microbiome-mediated toxicity in affected smokers, in the context of taxonomic and functional diversity of the oral mucosal microbiome.

Limitations

The main limitations of this study are the high microheterogeneity of the oral cavity (51) and the difficulty in obtaining markers of disease that coincide with previous studies. However, previous studies have focused on other specimen collection techniques (e.g., saliva, tissue/liquid type, etc.), varying proximity to tumors and other underlying risk factors (e.g., HPV, diet), hence, showing different disease-associated taxa [e.g., Fusobacterium and Lactobacillus, Prevotella, Streptococcus (8, 54, 55)]. Furthermore, while our sample is made up primarily of oral and oropharyngeal tumors, our cohort also included some tumors in the larynx/hypopharynx, which also potentially differentiates our dataset from those already present in the literature.

In addition, interpretation of our results requires the recognition that we have investigated an association between disease status (oral/head and neck cancer) and the oral microbiome. Therefore, we cannot make inferences on the direction of causality, as the microbiome findings may be the result of tumor development rather than a precursor to tumor formation. However, given our prior data that DNA adduct formation is greater in those with oral/head and neck cancer, we are working toward elucidating the nature of several potential mutagenic variables in this patient population, which will eventually contribute to an understanding of factors influencing tobacco-induced carcinogenesis. Another limitation of our study is the low sample size as a total of 51 subjects were examined. Although it represents the first and largest sample of patients studied for both DNA adduct formation and oral microbiome of smokers, we nonetheless desire an analysis in a larger cohort as soon as possible. However, the fact that we focus exclusively on smokers is a potential and novel strength of this report compared with other previous efforts. A slight difference in the average age between cases and controls could be perceived as another limitation; however, there is no data in the literature on the potential effect of age on oral microbiome, while research on gut microbiome suggests no effect in the relevant age range (between 40 and 60 years; refs. 56, 57).

In addition, our cohort did not contain specific oral health data, was studied via buccal cells for characterization of the oral microbiome and included a small percentage of nonoral/oropharyngeal cancers. Given that the role of the oral microbiome and cancer is still in the early stages, it is difficult to know if this is impactful in our dataset. Oral health practices may certainly impact our findings but given the dichotomous results between cases and controls relating to species variability and interindividual variation, oral health habits are less likely to be an explanation for our results. While there is likely to be variability in the nature of the microbiome at different sites in the upper aerodigestive tract, the same can be said of different sites in the oral cavity. That is, the microbial makeup of the gingiva, palate, and buccal mucosa are known to be distinct (50) and the impact of this on oral microbiome research, in the context of disease, to date is unknown. Thus, while it would be ideal to adequately capture the microbiology of the entire oral cavity, or even the upper aerodigestive tract, this would be highly cumbersome and represents a hurdle that has not yet been overcome in the medical literature.

Finally, although we aimed for increased accuracy of our functional predictions (maximum nearest taxon index = 2), our functional analyses are based on comparing the 16S rRNA sequences obtained with those of bacterial genomes reported in databases (58). Therefore, these predicted functional analyses should not be considered as substitutes of metagenomic approaches and actual functional analyses must be conducted to confirm the trends reported in a larger cohort, especially to
Cancer Prev Res; 13(5) May 2020

CANCER PREVENTION RESEARCH

Sharma et al.

determine potential HNSCC associations with the metabolism of xenobiotics by the mucosal oral microbiome.

Conclusions and future directions

With this current report, we are now able to examine our DNA adduct results in the context of information pertaining to the oral microbiome in the same patient sample. This analysis revealed that the increase in DNA adducts seen among cases corresponded to, and is associated with a reduction in bacterial richness present in the microbiome. In addition, we found concordance between increasing adduct levels and the number of taxa present among cases. The association between specific taxa of the oral microbiome, specifically the genera Stenotrophomonas, and the presence of DNA damage in mucosal tissue is readily apparent in our analysis and should be further investigated focusing on direct functional approaches (metagenomics, metabolomics, transcriptomics) to shed more light on the functions of HNSCC-associated taxa and the potential role of the microbiome on metabolic activation of smoke carcinogens among affected versus healthy smokers. Furthermore, we cannot assign cause or effect modification of cancer risk to the microbial characteristics identified here, and this association requires further study to better understand the role of bacterial richness, commensal homeostasis, and specific bacterial profiles in tobacco-induced oral/head and neck cancer.

In summary, we have examined and characterized the oral microbiome and quantified the presence of DNA adduct levels in smokers with oral/head and neck cancer compared with cancer-free smokers. Not only have we identified specific bacterial taxa characterizing both cancer status and those who are cancer free, but we found that there are specific bacterial taxa (e.g., Stenotrophomonas) directly correlated to increasing levels of DNA adducts and the increased abundance of xenobiotic-processing functions of the oral microbiome in affected versus healthy smokers. We therefore conclude that our data suggest an association between bacterial richness, bacterial diversity, and specific bacterial taxa and tobacco-induced oral/head and neck cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: I. Stepanov, A. Gomez, S.S. Khariwala

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): I. Stepanov, S.S. Khariwala

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A. Kumar-Sharma, I. Stepanov, A. Gomez, S.S. Khariwala

Writing, review, and/or revision of the manuscript: A. Kumar-Sharma, W.T. DeBusk, I. Stepanov, A. Gomez, S.S. Khariwala

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A. Kumar-Sharma, A. Gomez

Acknowledgments

This work is supported by NIDCR/NIH K23DE023572 (to S.S. Khariwala) and NCRI/NIH R01CA180880 (to I. Stepanov).

The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper. URL: http://www.msi.umn.edu.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 1, 2019; revised January 8, 2020; accepted February 14, 2020; published first February 18, 2020.

References


Cancer Prevention Research

Oral Microbiome Profiling in Smokers with and without Head and Neck Cancer Reveals Variations Between Health and Disease

Ashok Kumar Sharma, William T. DeBusk, Irina Stepanov, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1940-6207.CAPR-19-0459

Supplementary Material
Access the most recent supplemental material at:
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2020/05/06/1940-6207.CAPR-19-0459.DC1

Cited articles
This article cites 50 articles, 7 of which you can access for free at:
http://cancerpreventionresearch.aacrjournals.org/content/13/5/463.full#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cancerpreventionresearch.aacrjournals.org/content/13/5/463.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.