The oral microbiome as a biological matrix for assessing tobacco-related exposures and biological effects in persons who use tobacco products

A DISSERTATION SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA BY

Aleksandra Alcheva

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Irina Stepanov, Ph.D., Advisor

September, 2023

Acknowledgements

I would like to thank my academic and research advisor, Dr. Irina Stepanov. I think of her as more than just my academic advisor. She is my mentor, confidante, and constant source of inspiration. Her dedication to fostering my growth as a researcher and scholar has left a profound mark on my academic and personal development.

I would also like to thank my dissertation committee, Dr. Silvia Balbo, Dr. Dana Carroll, Dr. Andres Gomez, and Dr. Ryan Demmer. Their mentorship, constructive feedback, and dedication to my academic growth have been instrumental in shaping the quality of this thesis.

This work could not have happened without the academic and research financial support provided to me throughout my graduate studies. I was fortunate to receive funding support from the National Cancer Institute (NCI R01-CA179246) and the Cancer Prevention Fund, Masonic Cancer Center. Also, I am deeply grateful to the participants of the Tobacco Research Studies who generously donated biological samples for my thesis research.

Furthermore, I would like to acknowledge the support and encouragement I have received from my family and friends. In particular, I would like to thank Shannon, Gabriela, and Logain for their support academically and emotionally. Having their friendship made the long hours of research more bearable, and I cherish the memories we made together. The love, patience, and understanding of my husband, Aleksandar, have been pillars of strength for me. Last but not least, I am immensely grateful to my parents, Lenka and Mitko, as well as my brother Kire, for their love, sacrifice, and strong belief in my potential. It has been their belief in me that has driven me to pursue a higher education.

This thesis represents years of hard work and dedication, and it would not have been possible without the support of these amazing individuals. I am truly fortunate to have had such a strong support system, and for that, I am eternally grateful.

Dedication

To my beloved parents, Lenka and Mitko, whose unwavering love and support have shaped me into who I am today. To my cherished husband, Aleksandar, whose love is my anchor, my confidant, and my greatest blessing. And to my precious children, Mila and Marko, the beating hearts of our family, who fill my life with endless joy and purpose.

Abstract

Chemicals present in tobacco smoke and e-cigarettes may cause unique alterations to an individual's oral microbiome (microorganisms that reside in the oral cavity), which in turn is likely to have an impact on immune responses, inflammatory processes, uptake and metabolism of tobacco toxicants and carcinogens and the overall health status. This dissertation involves three complementary studies which cumulatively provide important insights into the utility of the oral microbiome as a comprehensive and robust biological matrix for studies of tobacco product use and disease susceptibility.

The first study investigated the responsiveness of the oral microbiome to changes in tobacco exposure through an 8-week trial in which persons who smoke were asked to switch to e-cigarette use. Compared to baseline smoking, notable changes in the oral microbial composition were found in those study participants who switched to exclusive use of e-cigarettes. These findings suggest that the oral microbiome could potentially serve as an indicator of changes in smoking-derived exposures and potentially for verification of e-cigarette use status in observational and interventional studies.

The second study aimed to determine whether the oral microbiome is associated with the tobacco-induced biological effects in the oral cells of individuals who smoke. In this study, specific taxa of the oral microbiome in persons who smoke were associated with high levels of HPB-releasing DNA adducts (DNA damage caused by tobacco-specific carcinogens) in their oral cells. Many of these taxa contribute to nitrate metabolism and subsequently to carcinogenic nitrosamine formation in oral cavity – a potential mechanism through which the oral microbiome may be related to oral cell DNA damage and/or contributes to tobacco-related cancer risk. The findings of this study suggest that the oral microbiome can serve as a predictor of tobacco-induced DNA damage in oral cavity, and therefore used as a tool for evaluating tobacco product harm and/or as a marker for assessing cancer susceptibility of individuals who smoke.

The third study explored the potential effect of sociodemographic factors on the oral microbiome of individuals who smoke. Answering this question helps to inform future studies that may use the oral microbiome as a matrix for assessing tobacco-derived exposures and biological effects. In this study, self-identified race emerged as a significant factor associated with the oral microbial composition. Given that race is a social construct, this finding likely reflects the cumulative effects of social and environmental stressors on the oral microbiome. Future studies of the oral microbiome as a biomarker of tobacco-derived

exposures and biological effects should adjust for such stressors. In addition, this finding suggests that the oral microbiome could be used as a tool in studies exploring factors contributing to health inequities.

In conclusion, this comprehensive investigation advances our understanding of the interplay between tobacco exposure, the oral microbiome, and biological effects. It positions the oral microbiome as a promising marker for tobacco regulatory research and in studies of disease susceptibility and prevention.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Background and significance

Even though smoking rates have decreased significantly in the U.S in recent years, from 14.0% in 2019 to 12.5% in 2020[1], cigarette smoking remains to be the leading preventable cause of cancer and cardiovascular disease, two diseases responsible for the majority of premature deaths.[2] This is mostly due to a wide range of toxicants and carcinogens present in cigarette smoke: there are more than 9,000 known chemical compounds, [5,6] out of which 250 are harmful to human health and over 70 are recognized as cancer-causing agents.[6] Continuous efforts to reduce smoking prevalence through cessation interventions, public education, and taxation are key for addressing this persistent public health priority. However, such efforts are complicated by the fact that tobacco products are highly addictive because they contain nicotine, the key psychoactive tobacco alkaloid that acts on the brain's reward system and increases dopamine levels.[7] Chronic exposure to nicotine leads to tolerance and neuroadaptation, reinforcing addiction to tobacco[8] and perpetuates the harmful health effects of cigarette smoking. To achieve meaningful decreases in smoking-related morbidity and mortality certain actions must be taken. Such actions may include (i) shifting persons addicted to nicotine towards products that result in lower toxicant and carcinogen exposures ("harm reduction") and (ii) among those who are unable or unwilling to quit smoking: identifying individuals at elevated risk for tobacco-associated diseases so that preventive and cessation interventions can be targeted to such individuals. The availability of effective tools for detecting and measuring product-associated and disease-relevant exposures and biological effects in persons who use tobacco/nicotine products is critical to pursuing these strategies. Self-reported amount of product use and biomarkers of exposures and biological effects have been widely used for such purposes in tobacco research studies. While informative, these approaches have important limitations. Self-reported product use behaviors, such as the amount, frequency, and duration are subject to misreporting and do not account for individual differences in tobacco constituent uptake and metabolism. Biomarkers represent more accurate measures of tobacco exposures and biological effects, but are typically measured one at a time, often require specific approaches to collecting and handling the required biospecimens, and the analytical assays are typically time-consuming, expensive, and require highly specialized laboratory expertise. These limitations undermine the use of biomarkers in low-resource settings and in studies with hard-to-reach

populations which are often most affected by tobacco use and its consequences. Therefore, the tobacco research field could benefit from new biomonitoring approaches that capture a broad range of tobacco-derived exposures and biological effects. ***This dissertation aims to explore the potential utility of the oral microbiome as a comprehensive and robust biological matrix for studies of tobacco product use and disease susceptibility*** by examining the responsiveness of the oral microbiome to changes in tobacco-related exposures; its association with the host biological effects in tobacco users; and how it is affected by the sociodemographic stressors linked to increased susceptibility to tobacco-associated diseases. The rationale for the focus on the oral microbiome and for the research approach is provided below, and the key research gaps that will be addressed by this dissertation are underlined.

The oral microbiome is an emerging and potentially powerful biological matrix in studies linking tobacco use to health outcomes. It is becoming increasingly evident that the oral microbial community features may be linked to a variety of cancer types. Some specific species have been found to strongly correlate with oral cancer, such as *Streptococcus sp., Peptostreptococcus sp., Prevotella sp., Fusobacterium sp., Porphyromonas gingivalis, and Capnocytophaga gingivalis*.[9,10] A study has shown that the composition of the oral microbiome plays a significant role in the prospective risk of developing esophageal cancer.[11]

In saliva samples from lung cancer patients, there was a greater abundance of *Veillonella, Neisseria, Capnocytophaga, and Selenomonas*.[12] A greater number of oral microbial species was observed in colorectal cancer tissues when compared with healthy controls,[13,14] suggesting that there was an influx of bacteria from the oral cavity that was detected in the cancer cases. The results of these studies indicate that dysbiosis of microbiome is associated with a disordered immune system, which in turn is one of the key mechanisms responsible for carcinogenesis and cardiovascular diseases. However, additional mechanisms may be involved.

Smoking has been shown to be an important determinant of the oral microbiome composition.[15,16] Oral bacterial communities are the first to come into contact with chemical constituents present in smoke and, thus, are likely to be most affected by them compared with microbiomes in other human body sites. Besides altering species profiles, exposures associated with tobacco product use might trigger the expression of different microbial genes and therefore contribute to the microbially-mediated metabolism of xenobiotics. For example, bacterial taxa such as *Corynebacterium, Kingella and Stenotrophomonas* have been functionally related to xenobiotic biodegradation and have been implied in metabolic pathways that

are capable of metabolizing phenolics, toluene and phenanthrene.[17,18] Studies indicate that functional pathways related to nitrotoluene, styrene, chlorocyclohexane and chlorobenzene (chemicals found in cigarette smoke) were depleted in individuals who smoke, as was cytochrome P450 xenobiotic metabolism.[15] Conversely, polycyclic aromatic hydrocarbon and xylene degradation were enriched in individuals who smoke.[15] Multiple analysis of inferred metagenomes also revealed depletion in aerobic metabolism pathways, including oxidative phosphorylation, and increased abundance of glycolysis and other oxygen-independent carbohydrate metabolism pathways in individuals who smoke.[15,19] Metabolic pathways involved in denitrification, sulfate reduction, the tricarboxylic acid (TCA) cycle, glyoxylate cycle, 2-methylcitrate cycle were also significantly different between individuals who smoke and those who do not. [19]

The oral microbiome may serve as a biological matrix for assessing and interpreting changes in tobacco-related exposures. In addition to studies comparing the oral microbiome between persons who do or do not smoke, important supportive evidence is emerging from studies comparing smoking to electronic cigarette (e-cigarette) use. Such comparisons are meaningful because levels of many harmful constituents that have been measured in some e-cigarette liquids and/or aerosols are generally much lower when compared to cigarette smoke [20] however, these products still expose users to certain toxicants and carcinogens such as aldehydes and other inflammatory agents.[21] Indeed, e-cigarette use has been found to influence the oral microbiome in a way that is different from what is found in individuals who smoke. For example, a large study observed depletion of Proteobacteria and enrichment for Firmicutes and Actinobacteria in oral wash samples from 1,204 US adults who smoke.[15] In addition, subgingival biofilms in individuals who smoke tended to be enriched with Gram-negative obligate anaerobes. E-cigarette users (who self-reported not smoking tobacco) had an enrichment of species within several genera including *Actinobacteria, Fusobacteria, Proteobacteria*, and *Saccharibacteria*.[22] Distinctive responsiveness of the oral microbiome to exposures from cigarettes and e-cigarettes in part might be due to important product differences. Studies have suggested that e-cigarettes may modulate the oral microbiome indirectly due to the components such as glycerol or polyethylene glycol that can be a source of nutrients for microorganisms. The clinical implications of these effects and how they are related to the risks for tobacco-associated health outcomes are still unknown. For example, an increased pathogen abundance, as occurs with cigarette and e-cigarette use, might provoke immune responses in the oral cavity, and proinflammatory mediators, secreted locally in the oral cavity and saliva, may enter the circulation and

promote systemic inflammatory processes.[23] The etiopathogenesis of many diseases where tobacco use is a risk factor such as CVD and cancer, is tightly linked to such immuno-inflammatory mechanisms.

All of the existing research that could inform about the potential utility of the oral microbiome to detect differences and/or changes in tobacco-related exposures has been based on cross-sectional studies in individuals with different tobacco/nicotine product status. There are no studies assessing the potential responsiveness of the oral microbiome to changes in exposure that occur in the typical clinical trials of short-term product switching. Addressing this research gap is important because (i) within-person changes in oral microbiome can help to identify trends that may not be detected in cross-sectional studies due to the significant inter-individual variability of oral microbiome profiles; (ii) cross-sectional studies include participants using various and often multiple e-cigarette devices and liquids, as well as dual users of cigarettes and e-cigarettes, and often rely on self-report of the product use status and history; and (iii) clinical trials of product switching allow for a relatively rapid and controlled assessment of alternative nicotine delivery products, such as e-cigarettes, and therefore represent an important asset in tobacco regulatory science.

The oral microbiome may be associated with disease-related biological effects in persons who use tobacco products. Inflammation, oxidative stress, and DNA damage are among the key biological effects mechanistically linking harmful exposures to the health effects of cigarette smoking.[24-26] These processes occur in the oral cavity of persons who use tobacco products, and the resulting macromolecular alterations are related to respiratory and systemic effects. For example, inflammation involves the infiltration of immune cells into stressed tissues, leading to lipid peroxidation and the generation of reactive oxygen species (ROS) and reactive nitrogen species. Mitochondria, which have their own DNA (mitochondrial DNA or mtDNA), are particularly susceptible to ROS attacks associated with oxidative stress. An increase in mitochondrial copy number has been observed as a potential compensatory mechanism for the decline in mitochondrial respiratory function due to oxidative damage.[27] Studies have shown that individuals who smoke have about 30% higher mtDNA content in their oral cells compared to those who don't, and this increase persists even after smoking cessation.[28,29] Importantly, higher mtDNA copy number has been positively associated with lung cancer risk.[30-32] As described above, the oral microbiomes of persons who smoke have altered functional pathways related to oxidative stress and inflammation. However, there is a limited number of studies assessing the potential association of the oral microbiome with the host inflammatory changes in the oral cavity of persons who smoke. Addressing this gap could provide important

insights for the use of the oral microbiome as a comprehensive indicator of various inflammatory and oxidative effects occurring in the oral cavity, which would be more effective than measuring individual biomarkers in oral cells.

Another important example is DNA adduct formation, which is a crucial step in cancer development and potentially provides the most direct measure of cancer risk.[33] Many tobacco carcinogens require metabolic activation through cytochrome P450 enzymes to exert their carcinogenic effects. Some metabolites produced during this activation process are electrophilic and can react with DNA, leading to DNA damage and the formation of DNA adducts. [34] DNA adduct formation can disrupt DNA replication and, if left unrepaired, can result in inheritable DNA mutations or abnormal gene expression, ultimately leading to carcinogenesis.[34] For example, the metabolic activation of the carcinogenic tobacco-specific nitrosamines NNK and NNN leads to the formation of pyridyloxobutyl (POB)-DNA adducts, most of which decompose under acid hydrolysis conditions to release a compound called 4-hydroxy-1-(3-pyridyl)-1-butanone (HPB).[35,36] Consistent with the tobacco-specificity of NNK and NNN, significantly higher levels of HPB-releasing DNA adducts have been reported in oral cells of individuals who smoke compared to those who do not.[36] Moreover, higher levels of HPB-releasing DNA adducts have been found in the oral cells of individuals who smoke and had oral/head and neck squamous cell cancer (HNSCC) compared to cancer-free individuals.[37] Recent study reported an association between microbial characteristics and HPB-releasing DNA adducts in the oral cells of individuals who smoke.[38] Key limitations of that study were the small number of participants and the presence of disease (head and neck squamous cell carcinoma, HNSCC) in individuals who displayed an association between the oral microbiome and the levels of HPB-releasing DNA adducts. There is a need to examine this exciting association in a larger number of persons who smoke and who are cancer-free, in order to better understand whether the oral microbiome may reflect the susceptibility to tobacco carcinogen-induced DNA damage.

The oral microbiome is likely to be affected by individual characteristics and stressors associated with the burden of tobacco-induced diseases. The risks for tobacco-associated diseases vary significantly among individuals and population subgroups. For example, certain races and ethnicities, individuals with lower education level or income, and those with mental health illnesses may not only have higher prevalence of smoking but also have higher susceptibility to relevant diseases.[3] On-going studies investigate inter-individual differences in the uptake and metabolism of tobacco-derived toxicants and carcinogens as the potential mechanisms; however, non-tobacco factors such as concomitant exposures to environmental

and sociodemographic stressors may play a role. Existing research suggests that such factors may also affect the oral microbiome. For example, one study found that African American showed a higher abundance of oral Bacteroidetes and a lower abundance of Actinobacteria and Firmicutes.[39] Studies have also reported differential microbiome patterning based on sociodemographic factors, regardless of smoking. For example, a recent study in adults living in New York City found that variation in taxonomic abundances could be linked to marital status, ethnicity, education, and age.[40]

A key limitation of studies that assessed the association of sociodemographic factors with the oral microbiome is that such studies included mixed populations of persons who do and do not smoke, typically with uneven distribution of smoking status (only a few people who smoke) and adjusting for smoking as a confounding factor in their analysis. To better understand how the oral microbiome can be used in tobacco research, there is a need for studies specifically focused on examining such associations among persons who smoke. Such studies could also provide key insights into the potential utility of the oral microbiome in future research of tobacco-associated health inequities.

Research objective: This dissertation contributes to the knowledge base on the potential utility of the oral microbiome as a tool for examining tobacco-related exposures and biological effects in persons who smoke and/or switch to alternative nicotine products. The conceptual framework illustrating the plausible interaction of the oral microbiome with tobacco-related exposures, host macromolecular and DNA damage, and sociodemographic factors in the context of this dissertation is presented in **Figure 1-1.** Towards achieving this dissertation's objective, the following specific aims addressed the research gaps and limitations identified above, as further presented in three separate dissertation chapters:

CHAPTER 2

Specific Aim 1: Assess the responsiveness of the oral microbiome to changes in tobacco exposures in a typical trial of short-term switching from smoking to e-cigarette use. To achieve this goal, oral samples were used from a previously completed clinical trial in which individuals who smoke were assigned to exclusive or *ad-libitum* use of e-cigarettes for 8 weeks, or to continued smoking (control group). The oral microbial composition was characterized using the available samples to assess the potential changes between baseline (smoking) and week 8 (end of study) in individuals who became exclusive e-cigarette users, as verified by a panel of smoking-related biomarkers. The oral microbiome trends were also compared

between these participants and those who became dual users (cigarettes and e-cigarettes) or were assigned to continued smoking.

Hypothesis: There are significant differences in the oral microbiome profile between baseline and week 8 in individuals who switched completely from smoking to e-cigarettes, while there are no detectable changes in the oral microbiomes between baseline and week 8 in participants who continued exclusive smoking or became dual users.

CHAPTER 3

Specific Aim 2: Investigate the association between the oral microbiome and tobacco-induced biological effects in oral cells of individuals who smoke. This Aim used biological samples and data from a previously completed study of tobacco constituent exposure and metabolism in 160 individuals who smoke. The parent study aimed to understand the mechanisms underlying the observed differences in smoking-induced cancer risk between African American (AA) and White (WH) persons who smoke. The oral microbiome was characterized in the available oral cell samples from this study. The content of mtDNA (measured as part of this dissertation work) and HPB-releasing DNA adducts (previously generated data) in these samples were used as biomarkers of tobacco-derived biological effects. A particular strength of this study is that there was substantial (56-fold) interindividual variation in oral cell HPB-releasing adduct levels across all participants. Importantly, the levels of adducts were not associated with smoking exposures (self-reported cigarettes per day or urinary biomarkers of smoking dose), potentially reflecting inter-individual variability in metabolic and/or repair enzyme activity. Therefore, this study provides an opportunity to investigate the unique, dose-independent association between the oral microbiome and the tobacco-induced DNA damage.

Hypothesis: Taxonomic characteristics of the oral microbiome are associated with mtDNA content, and the levels of HPB-releasing DNA adducts in oral cells of individuals who smoke.

CHAPTER 4

Specific Aim 3: Investigate the relationship between sociodemographic factors and the oral microbiome in individuals who smoke. In order to achieve this goal, data collected from the same study as in Aim 2 was used. The study population was well-balanced in terms of self-reported race/ethnicity (75 AA and 71 WH), biological sex (71 male and 75 female), age, and smoking history. This allowed to examine how other

sociodemographic factors that are commonly used as the surrogate indicators of psychosocial stress (education, employment, income, marital status – all available from questionnaire data collected in the study) may be associated with the oral microbiome composition.

Hypothesis: Despite the significant impact of smoking on the oral microbiome, there are detectable differences in the oral microbiome between AA and WH persons who smoke. Some of the other studied sociodemographic characteristics such as education, employment, income, marital status and living situation, may be associated with the oral microbiome in these study participants.

A NOTE ON MICROBIOME MEASURES USED IN THIS DISSERTATION

Microbiome research involves specialized methods to assess and describe specific measurements and parameters. Table 1-1 below lists key methods and measures used in this dissertation, as a reference for further reading.

Table 1-1. Methods and measures used in microbiome research

Method or measure	Description
16S rRNA amplicon gene sequencing	Method to characterize bacterial populations by sequencing specific variable regions of the highly conserved bacterial 16S rRNA gene
Amplicon sequencing variants (ASV)	Clusters sequencing reads based on their similarity and identifying unique genetic variants within a defined region of interest
Abundance	Total proportion of a given bacterial taxon within a sample
Richness	Number of different taxa (ASVs) present in a sample
Evenness	Measure of similarity in relative abundance/frequency distribution of taxa within a community. Dominance of one taxa implies decreased evenness
Alpha diversity	Measure of diversity within a sample, taking into account richness and/or evenness. Commonly used metrics of α-diversity include Chao1, Shannon and Simpson indexes
Shannon index	Measure of diversity combining richness and evenness. Higher values indicate higher diversity
Beta diversity	Measure of dissimilarity between samples. High beta diversity implies large dissimilarities between samples
Principal coordinates analysis	Multivariate ordination method used to visualize individual or group similarities (beta diversity)
Bray-Curtis dissimilarity	A beta-diversity index that measures similarities or dissimilarities between samples based on the presence and abundance of microbes that are shared between them
UniFrac	A beta-diversity index that measures similarities or dissimilarities between samples accounting for phylogenetic relatedness between microbes; microbial communities that are more similar are composed of members closely related phylogenetically, implying a shared evolutionary history
Machine learning (Random Forest)	Identify the subset of microbial taxa whose relative abundances best predict target variable

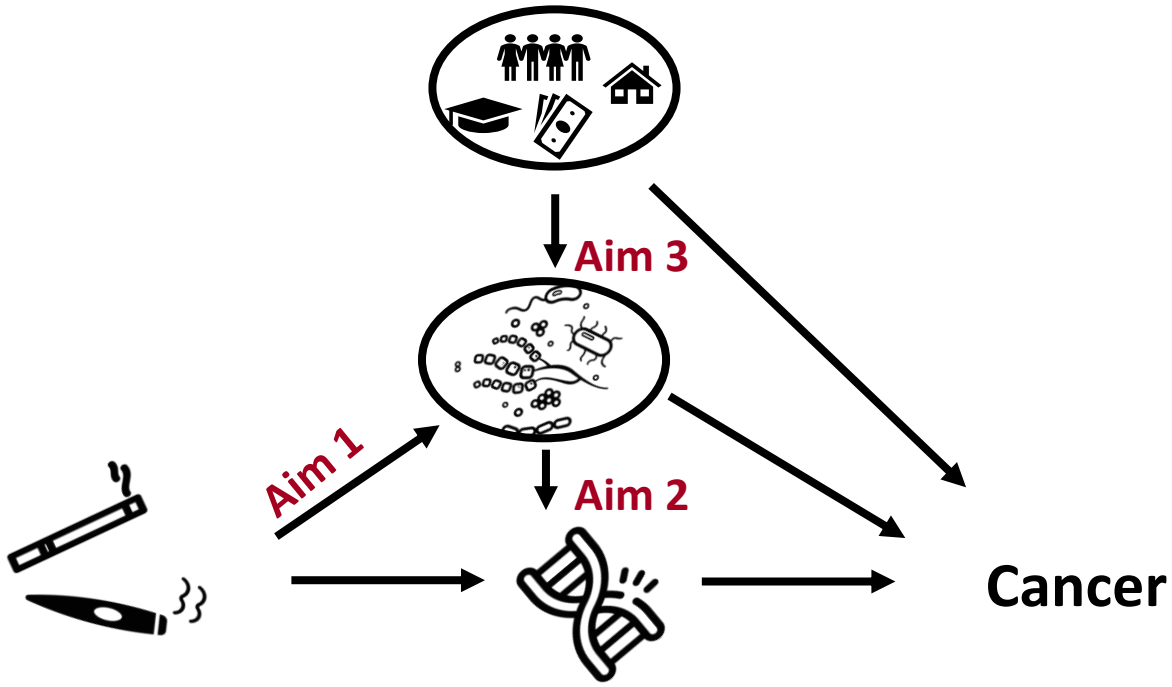


Figure 1-1. The hypothesized links between the oral microbiome composition, tobacco exposure, biological effects, and sociodemographic factors in individuals who smoke cigarettes. The associations explored in the Specific Aims of this dissertation are indicated.

CHAPTER 2

Responsiveness of the oral microbiome to changes in tobacco exposures in a typical trial of short-term switching from smoking to e-cigarette use

Introduction

The chemistry of tobacco products is highly complex, with the tobacco plant itself containing over 4,000 compounds.[4] During cultivation, processing, manufacturing, and burning, the composition of tobacco smoke undergoes further changes.[5] Cigarette smoke alone contains more than 5,000 known chemical compounds, [5,6] out of which 250 are harmful to human health and approximately 70 are considered cancer-causing agents.[6] An important group of harmful constituents in the smoke is comprised of a wide range of reactive species and free radicals, including nitric oxide (NO), superoxide anion (O_2^-), and hydroxyl radical (OH).[41] These species are generated during tobacco burning and can initiate oxidative damage and contribute to oxidative stress by promoting lipid peroxidation, protein oxidation, and DNA damage. Tobacco-specific nitrosamines (TSNA) is another group of important constituents in tobacco and cigarette smoke. These compounds are formed from nicotine by a nitrosation that mainly occurs during the curing and processing of tobacco.[42] *N*′-nitrosonornicotine (NNN) and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) are two of the most studied TSNA; their role in the induction of several malignancies, including lung, esophageal, and pancreatic cancers is well-documented.[43-45] NNN and NNK are classified as carcinogenic to humans (Group 1) by the International Agency for Research on Cancer (IARC).[6] Polycyclic aromatic hydrocarbons (PAH) are another significant group of chemicals present in tobacco smoke. They are formed as a result of incomplete combustion, among them, benzo[*a*]pyrene (BaP) is considered the most potent carcinogen and is classified as Group 1 human carcinogen by the IARC.[46] The incomplete combustion of cigarette tobacco filler during and between puffs also generates many volatile organic compounds (VOC). Certain VOC such as acrolein, crotonaldehyde, and propylene oxide are important toxicants and carcinogens in cigarette smoke.[47] Acrolein is an intense irritant, causes cilia-toxicity in the lung, and is proposed to be a lung carcinogen,[47,48] crotonaldehyde is a potent irritant and forms DNA adducts in the human lung,[49] and propylene oxide is an IARC Group 2B carcinogen. [50]

Trends in tobacco products are continually evolving, with e-cigarettes and heated tobacco products becoming increasingly popular. The relative risks and benefits of such products are not well-characterized, although temporary switching from smoking to e-cigarettes is likely to significantly reduce harmful

exposures in persons who smoke. This is because levels of many harmful constituents that have been measured in some e-cigarette liquids and/or aerosols are generally much lower when compared to cigarette smoke.[20] However, long-term use of e-cigarettes by persons who used to smoke or by those with no prior smoking history may pose health risks because e-cigarettes still expose users to toxicants and carcinogens.[51] For example, heating and vaporizing propylene glycol and glycerin in e-cigarettes produce a range of aldehydes and other inflammatory agents.[46] Furthermore, many individuals who smoke and adopt e-cigarettes become dual users of both products and depending on their use patterns may or may not decrease their smoking-related exposures.

Clinical trials contribute significantly to our understanding of tobacco use and its potential health effects. Switching studies, examining the transition from combustible cigarettes to electronic cigarettes, have been a significant area of interest in recent years. The importance of such clinical trials in tobacco research is crucial, as they provide a rigorous scientific evaluation of the effects of switching from cigarettes to e-cigarettes. By eliminating confounding factors such as self-reported smoking behaviors, as well as capturing the dynamic changes associated with switching, clinical trials provide more reliable and comprehensive data, allowing for a more accurate assessment of the potential benefits and risks associated with switching to e-cigarettes. New measures for assessing the exposures associated with changes in tobacco product use in clinical trials are of the utmost importance in tobacco regulatory science. Indeed, many regulatory agencies have issued calls for the development of robust and effective markers that can be used for the assessment of constituents at different concentrations and patterns of user behavior (such as dual or poly use and product switching).

Biomarkers of tobacco constituent exposure are important, but insufficient tools for assessing health risks. Biomarkers of tobacco constituent exposure provide objective measurements of toxicant doses in individuals, considering variations in product use, metabolism, and individual characteristics.[52] Such biomarkers have been well-established for many key tobacco toxicants and carcinogens. For example, the intake of the major addictive alkaloid in tobacco products, nicotine, can be assessed by urinary total nicotine equivalents (TNE) which accounts for about 80–90% of a daily nicotine dose.[52] Human exposure to NNK can be measured by analyzing the sum of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) and its glucuronides in urine, referred to as total NNAL.[53] Exposure to the related tobacco carcinogen NNN can be measured by analyzing the sum of unchanged NNN and its *N*-glucuronide excreted in urine (total NNN).[42,53] Urinary 1-hydroxypyrene (1-HOP) is a metabolite of the representative PAH pyrene and has been

widely used as a surrogate biomarker of exposure to PAH.[21] Exposure to volatile toxicants is commonly measured by their mercapturic acids excreted in urine.[21]

The application of tobacco constituent biomarkers in various tobacco studies shows that their levels are associated with exposure dose and the risk for cancer. Biomarkers such as TNE and total NNAL correlate with self-reported cigarettes per day (CPD), and total NNAL is a reliable marker for NNK uptake.[54] Urinary 1-HOP and mercapturic acids decrease after the reduction of the number of cigarettes smoked per day or complete cessation of smoking.[55] Several prospective epidemiological studies have demonstrated the relationship of some of these biomarkers with cancer risk in individuals who smoke. For example, levels of cotinine (nicotine metabolite) and total NNAL were predictive of lung cancer risk in individuals who smoke in the Shanghai and Singapore cohorts.[56] Urinary total NNN was strongly associated with the risk of developing esophageal cancer.[57] In all these analyses, the associations were significant even after adjustment for self-reported smoking history and other confounders. However, while biomarkers of exposure are useful in assessing exposures and in some cases have been linked to the risk of specific disease at the population level, they do not provide information on the disease pathophysiology-related biological effects that occur as the result of harmful chemical uptake. Therefore, biomarkers of exposure alone have limited capacity for predicting disease risk among tobacco users or explaining inter-individual differences in such risks. For instance, in studies comparing different ethnic groups such as Multiethnic Cohort (MEC), biomarkers did not fully explain the variations in smoking-associated lung cancer risk across various ethnic groups.[58] Significant differences were observed in levels of NNAL and TNE between African American and White individuals who smoke; and analyses of mercapturic acids in individuals who smoke revealed inconsistent trends in relation to cancer risk among different ethnic groups.[58,59]

Studies with electronic cigarette (e-cigarette) users reveal additional limitations of urinary biomarkers of exposure. Such studies typically show significant reductions in the levels of many smoking-related biomarkers of exposure.[20,21] For example, a population-based, longitudinal cohort study conducted in the United States in 2013-2014 showed that exclusive e-cigarette users had lower urinary total NNAL (98% lower), biomarkers of PAH (34%–62% lower), and some biomarkers for volatile organic compounds (VOC) (59%–97% lower) when compared with exclusive cigarette users.[21] However, most e-cigarette users also smoke regular cigarettes, and studies report that nearly all biomarkers in such dual users are either similar or higher than in exclusive cigarette users.[20,21] It is not clear how these findings can be used to assess the potential consequences of dual use. Furthermore, e-cigarettes are a known source of aldehydes and

other inflammatory agents, although at levels that generally lower than in cigarette smoke.[60] Such constituents are highly reactive with various cellular components, such as proteins and DNA, upon contact with tissues. Therefore, urinary biomarkers may not capture biologically important exposures to such constituents that are lower than in persons smoking cigarettes but higher than in persons who do not use any tobacco or nicotine product. Indeed, e-cigarette use was associated with an increased risk of respiratory diseases (COPD, emphysema, chronic bronchitis, and asthma) when compared with non-users in studies assessing health outcomes and symptoms.[61,62] Such outcomes could not have been predicted based on urinary mercapturic acid levels. Biomarker studies are important, but their ability to predict health outcomes depends on a nuanced understanding of what they represent (exposure and/or effect) as well as the factors that influence them; therefore, the optimal strategy may involve identifying and utilizing the most appropriate combination of markers.

The oral cavity and its microbiome are promising targets for assessing tobacco-related exposures, regardless of the type of tobacco product used. The oral bacterial community is the first to come into contact with smoke and/or e-liquid aerosol chemicals and, therefore, has the greatest potential to be affected by such exposures. In fact, some studies have shown that overall microbiome composition differs by smoking status.[15] E-cigarette users have also been shown to have a unique oral microbiome composition and function compared to those who smoke cigarettes or never used tobacco products.[63] It has been proposed that smoking can favor the growth of pathogenic bacteria within the oral cavity through a shift from commensal aerobic bacteria to anaerobic species. A higher percentage of potentially opportunistic oral pathogens, within the *Prevotella, Streptococcus*, and *Rothia genera*, were reported in individuals who smoke.[15] Recent studies reported modulation of bacterial profiles, with higher abundances *Porphyromonas* and *Veillonella* genera in saliva samples of e-cigarette users.[64] These findings indicate that differences in the oral microbiome composition and function among users of different tobacco or nicotine products could serve as a comprehensive and biologically reflective measure of exposure, complementing and/or replacing the traditional targeted analyses of specific constituents and adding an additional biological layer to tobacco exposure assessment. However, there is a limited ability to identify and determine the direction of the observed changes in microbial composition due to the cross-sectional study design and heterogeneity of e-cigarette use assessments, with many studies reporting concurrent use of conventional cigarettes and e-cigarettes.

In this study, the oral microbiome was examined at baseline and after switching to e-cigarettes (biochemically confirmed) for 8 weeks to explore whether the oral microbiome is responsive to changes in tobacco-associated exposures. Given the significant differences in the chemical profile of cigarette smoke and e-cigarette aerosol, and observed taxonomical changes after short-term cessation,[22] the key hypothesis in this study is that the oral microbiome will change in response to changes in exposure, resulting in different profiles between baseline and week 8 in individuals who completely switch from smoking to e-cigarettes.

Materials and Methods

In this study, we used oral cells and biospecimens collected as part of the Consortium on Methods Evaluating Tobacco (COMET) study conducted by the Tobacco Research Program at the University of Minnesota in collaboration with Ohio State University and Roswell Park Cancer Institute in Buffalo, NY.

Overview of study design and study procedures in COMET study

An overarching goal of COMET study was to determine a product's toxicity, its abuse liability (e.g., uptake and dependence potential), and how consumers perceive the product, which influences uptake and continued use (e.g., how consumers react to promotional messages, packaging, and prices). Individuals who smoke (n=245) were randomized in the clinical trial phase, following a two-week baseline smoking phase, to one of four experimental conditions for 8 weeks as described in Figure 2-1 . Participants had to be 18 years old and smoke at least 5 cigarettes daily for at least one year. They were not pregnant, had no serious quitting attempt in the last 3 months, not using any other tobacco products, were willing to quit smoking (if randomized to e-cigarettes), and were in good general health. All study participants have given informed consent. In At weeks 2, 4, and 8 of assigned product use urine sample and oral cells were collected. At every clinic visit, tobacco use patterns were assessed via the Daily Use Summary and subjective forms, vital signs, weight, oxygen saturation and expired air CO were also obtained. Questionnaires that measure factors that may influence biomarker assessments were administered at the time of biospecimen collection. Compliance with the instructions for product use was determined by: a) daily diary records; b) tobacco product logs where the number of products dispensed was recorded and unused products collected and recorded; and c) use of biomarkers to confirm if participants were abstinent from combustible products. For the purposes of this study, samples were retrieved only from participants who agreed to store and analyze their biological specimens in the future.

Collection of oral samples

At the beginning of the clinic visit, participants were instructed to brush their teeth with the provided pre-pasted toothbrush. After 20 minutes without eating, drinking, chewing gum, smoking, or using e-cigarettes, buccal cells were collected from the mouth using a cytobrush. Each participant was given two cytobrushes, one per each side of the mouth, and the cells from each cytobrush were transferred into a separate tube containing saline solution. The tubes were placed on ice, then centrifuged immediately to pellet the cells, which were then stored at -20°C.

DNA extraction from oral cells

DNA from oral swabs was isolated by using a modified Qiagen DNA isolation protocol for DNA isolation from tissues and blood.[35] Briefly, the samples were thawed at room temperature and centrifuged at 1500 × g for 15 min. The supernatant was discarded, the cell pellet were resuspended in 3 ml of cell lysis solution from Qiagen kit, homogenized, and treated with proteinase K (10 min 56˚C) and followed by treatment with RNase A, and precipitating proteins, DNA was isolated and purified according to the manufacturer protocol.

Oral bacteria community profiling

A 16S rRNA-based profiling of oral bacterial communities was completed at the at the University of Minnesota Genomics Center (UMGC) targeting the V4 hypervariable region (barcode primer pair 515 f - GTGCCAGCMGCCGCGGTA and 806r-GGACTACHVGGGTWTCTAAT) on the Illumina MiSeq sequencing platform. Sterile water was used as a control for each batch of DNA amplification. Upon completion of the analysis, the sequenced genetic data was archived at the Minnesota Supercomputing Institute (MSI).

Previously analyzed biomarkers

Carbon Monoxide (CO). CO was measured in exhaled breath (as concentration of CO in parts per million).

Cyanoethyl mercapturic acid (CEMA).

CEMA is a urinary metabolite of acrylonitrile, a toxicant found in substantial quantities in cigarette smoke, but not in non-combusted products such as e-cigarettes. CEMA was previously analyzed, in brief, [CD_3]3 HPMA and [CD_3]CEMA were added to 200 µl urine. The Oasis Max 60 mg mixed mode anion exchange 96 well plate (Waters Corp.) was preconditioned with methanol and 2% aqueous NH_4OH. The

sample was applied, and eluted with 30% methanol containing 2% formic acid, to give the fraction containing 3-HPMA and CEMA. The fraction was concentrated to dryness, and the residue dissolved in 60 µl MeOH/NH_4OAc:1/3, pH 6.8, and analyzed by atmospheric pressure chemical ionization (APCI)-LC-MS/MS SRM using an Agilent 1100 HPLC system (Agilent Technologies).The MS/MS system was run in the negative APCI mode using the following ion mass transitions for detection: 3-HPMA, *m/z* 220 → *m/z* 91; [CD_3]3-HPMA, *m/z* 223 → *m/z* 91; CEMA, *m/z* 215 → *m/z* 162, [CD_3]CEMA, *m/z* 218 → *m/z* 165. [65,66]

Urinary total nicotine equivalents (TNE). TNE is a urinary biomarker of daily nicotine intake. It is expressed as a molar sum of nicotine and its known metabolites, accounting for approximately 73-96% of the nicotine dose.[67] Urinary TNE was analyzed as previously described.[68] In brief, 10 µL urine was 10-fold diluted, the internal standards – [CD_3]nicotine, [CD_3]cotinine, [CD_3]3′-hydroxycotinine and [CD_3]nicotine N-oxide – were added to 96-well plates, treated with β-glucuronidase, and subjected to solid phase extraction. The eluents were analyzed by LC-MS/MS with transitions *m/z* 163 → *m/z* 130 and *m/z* 166 → *m/z* 130 for nicotine, [CD_3]nicotine; *m/z* 177 → *m/z* 98 and *m/z* 180 → *m/z* 101 for cotinine and [CD_3]cotinine, *m/z* 193 → m/z 134 and *m/z* 196 → *m/z* 134 for 3′-hydroxycotinine and [CD_3]3′-hydroxycotinine and *m/z* 179 → *m/z* 130 and *m/z* 182 → *m/z* 130 for nicotine N-oxide and [CD_3]nicotine N-oxide.

Total NNAL. Total NNAL was measured as previously described. Briefly, urine samples were mixed with [$^{13}C_6$]NNAL internal standard, treated with *β*-glucuronidase to release NNAL from their *N*- and *O*-glucuronides, and further purified using solid-phase extraction cartridges. The appropriate eluants were then analyzed by liquid chromatography-tandem mass-spectrometry (LC-MS/MS), monitoring transitions *m/z* 210 → 180 for NNAL, and *m/z* 216 → 186 for [$^{13}C_6$]NNAL. Unconjugated (free) NNAL was analyzed by the same method, but the urine samples were not treated with *β*-glucuronidase prior to their purification.

Selecting study participants based on biomarkers of tobacco exposure

A key aspect of clinical trials is product use compliance, since failure to fully replace cigarettes with the assigned product will impact biomarker changes. The compliance with the protocol of the assigned product use could be assessed by evaluating selected cigarette smoke and e-cigarette constituents, such as CO, which is an established biomarker for distinguishing tobacco users from nonusers, responding to cessation and reduced use.[69] NNAL is another useful biomarker for verifying product use status because of its prolonged half-life: studies indicate that about 34% of NNAL remains up to three week following smoking cessation.[69] Lastly, cyanoethyl mercapturic acid (CEMA) is a urinary metabolite of acrylonitrile, a toxicant

found in substantial quantities in cigarette smoke, but not in non-combusted products such as e-cigarettes.[66]

Participants who (i) switched to e-cigarettes completely; (ii) switched to e-cigarettes partially (dual users); or (iii) continued to smoke after the study period of 8 weeks were identified based on the available biomarker data. Participants were classified into one of the three groups based on CO, NNAL, and CEMA levels at baseline and week 8. Participants who had CO < 4 ppm, significant NNAL reduction (>90%) at week 8, and CEMA < 27 pmol/mL were considered complete e-cigarette switchers. Participants were classified as dual users based on the percent reduction of tobacco exposure biomarkers (CO, NNAL and CEMA) at the end of the study period (week 8) as described in the Results section.

Bioinformatics and statistical analyses

Sequence processing was performed using the DADA2 bioinformatics pipeline. Forward and reverse reads were trimmed to 200 nucleotides to remove low-quality reads and merged. High-quality sequences were aligned against the SILVA database and used for downstream statistical analysis. All microbial community ecology analyses were performed in the R statistical interface.[70] Relative abundances of each ASV on rarefied data (depth ¼ 1,000 sequences) were used for assessing microbial diversity. Microbiome diversity metrics such as alpha diversity indices (Observed Richness and Shannon's H indices), beta diversity (Bray–Curtis and UniFrac distance matrices), ordination analysis [principal coordinates analysis (PCoA)] were preformed using the R phyloseq package.[71] Permutational multivariate analysis of variance (PERMANOVA) was calculated using the adonis function within the R vegan package.[72] Kruskal–Wallis and Wilcoxon rank-sum tests were used to test the differences in diversity metrics between the groups. Lastly, differential abundances were analyzed using the DESeq2 package.[73] All graphs were plotted using the ggplots package.[74]

Results

A summary of participant characteristics and biomarker levels (at baseline and after 8 weeks of assigned product use) is shown in Table 2-1. Age, biological sex, race, and years of smoking were distributed equally across groups. Only 78 of 245 participants met the rigorous selection criteria and were included in the study. Individuals who smoked at baseline and were randomly assigned to different conditions did not differ based on biomarkers of smoking exposure. Across all groups, there was a statistically significant

difference in all biomarkers of exposure after eight weeks of assigned product type use. An important strength of this study is that stringent criteria were used based on available biomarker data to biochemically confirm smoking status/product type use. One such biomarker, carbon monoxide (CO), is widely recognized for its ability to distinguish tobacco users from nonusers, and it also responds to changes in smoking behavior. A cut-off ≤ 4 parts per million (ppm) was used to identify individuals who had fully switched to e-cigarettes. However, because of the short half-life of CO (4-6 hours), this measurement alone is not sufficient to verify the absence of smoking over the past several days. Therefore, urinary biomarkers of tobacco exposure were also used when selecting study participants. Previous research has shown that NNAL, which is a metabolite of NNK, remains detectable for up to three weeks after smoking cessation. Therefore, only participants with a >90% reduction in NNAL levels by week 8 were considered "complete switchers". Lastly, cyanoethyl mercapturic acid (CEMA), a urinary metabolite of acrylonitrile, was used as additional biomarker to verify complete switching from smoking to e-cigarettes. Acrylonitrile is a combustion-derived toxicant found in cigarette smoke but not in e-cigarettes, and a CEMA cut point of 27 pmol/mL urine was identified as a highly accurate differentiator between cigarette smokers and nonsmokers, with sensitivity and specificity exceeding 95%.[66] Out of 70 participants randomly assigned to switch to e-cigarettes, 12 had all these biomarkers below the established cut-off values and were included in the complete switching group. The difference in biomarkers of smoking exposure (C0, NNAL, CEMA) within the individuals who were assigned to e-cigarette use and became either dual users or completely switched to e-cigarettes is illustrated in **Figure 2-2**. Consistent with the applied selection criteria, there was a significant reduction in all three biomarkers ($p<0.001$, $p=0.029$, $p<0.001$ for CO, NNAL and CEMA, respectively) in those who completely switched to e-cigarettes. In those who became dual users at week 8, urinary NNAL decreased ($p=0.046$), while other biomarkers remained similar.

Oral microbiome profiles

A total of 4,293,246 16S rRNA sequence reads were obtained, which after quality filtering, represented 1744 unique amplicon sequence variant (ASVs). Following further filtering (e.g., removing taxa with present in fewer than three samples and showing fewer than 10 reads per taxa), 230 unique ASVs were identified in individuals who smoked at baseline and after 8 weeks of using assigned products. **Figure 2-3** shows the overall distribution of bacterial genera in all participants at baseline before randomization while smoking cigarettes (A) and after 8 weeks of using the assigned product type (B). Dominant genera

were comprised by *Actinomyces, Fusobacterium, Gemella, Haemophilus, Leptotrichia, Neisseria, Porphyromonas, Prevotela, Rothia, Streptococcus,* and *Veillonella* irrespective of product type usage.

Taxonomic composition and diversity

Alpha diversity in microbial communities, including Shannon index and Observed Richness, is a key indicator of the composition of these communities. A pairwise test was conducted to determine whether Shannon's alpha-diversity had changed within the groups over time. There was a statistically significant difference in Shannon index between baseline and week 8 for those who smoked at baseline and switched completely to e-cigarettes for 8 weeks (p=0.027, **[Fig. 2-4A]**) with lower diversity observed at the end of the study period (week 8) in those who completely switched to e-cigarettes Shannon Indexes did not change significantly for those who remained smoking or became dual users at the end of the study period **(Fig. 2-4 B,C)**. When Observed Richness was assessed across the groups with different status of tobacco use for 8 weeks, there was no significant difference (p>0.05; **[Fig. 2-5A]**). There was a significant difference in Shannon indexes across product types (p=0.021 for smoking vs. e-cigarette use and p=0.047 for dual vs. e-cigarette use **[Fig. 2-5B]**) with greater diversity observed in those who smoked or were dual users when compared to e-cigarette users. Bray–Curtis dissimilarity distances were computed to analyze compositional dissimilarities between groups with different patterns of use. Beta diversity was not different among individuals at baseline while smoking cigarettes and were later randomized to different tobacco product (p>0.05, **[Fig. 2-6A]**). However, compositional differences were found between those who continued smoking, became dual users, or completely switched to e-cigarette use for 8 weeks (r^2 =0.04; p=0.018, **[Fig. 2-6B]**).

Identifying discriminatory features of the oral microbiome by different tobacco product use

To identify discriminatory features of the oral microbiome across individuals with different patterns of tobacco product use, differential abundance analysis was performed using the DESeq2 R package. First, the differentially abundant taxa were assessed in those who smoked at baseline and after they switched to e-cigarettes for 8 weeks **(Fig. 2-7A)**. Furthermore, differentially abundant taxa were identified between those who smoked at baseline and those who became dual users at the end of the study period **(Fig. 2-7B)**. The abundance of fifteen genera increased significantly when users switched to e-cigarettes completely. These taxa belonged to the genera *Streptococcus, Rothia, Veillonella,* and *Fusobacterium*. Those who became dual users at the end of the study period had decreased abundances of *Leptotrichia, Lactobacillus,*

Limosilactobacillus, Neisseria, Streptococcus, Actinomyces, Haemophilus, and Lancefieldella. Moreover, differences in abundance between the groups (smoking, dual use, and e-cigarette use) were analyzed cross-sectionally at week 8. Differentially abundant taxa are listed in **Supplemental Table 2-1, 2-2 and 2-3.** A total of 28 genera were differentially abundant between those who smoked and those who used e-cigarettes, with *Pasteurellaceae, Leptotrichia, Actinobacillus, Amniculibacterium, Neisseria* being decreased in e-cigarette users. When compared to dual users abundance of *Limosilactobacillus, Ligilactobacillus, Streptococcus, Fretibacterium Lactobacillus*, and *Campylobacter* increased in e-cigarette users.

Discussion

The objective of this study was to fill a crucial knowledge gap by investigating the potential changes in the oral microbiome between baseline smoking and after biochemically verified switching to e-cigarettes in a typical 8-week product switching study. The goal of the research undertaken in this study aimed to determine if the oral microbiome is responsive to product change and could serve as a potential matrix for assessing the impact of changes in tobacco-related exposures in short-term clinical trials of product switching involving diverse products.

As part of this study, alpha and beta diversity measures were used to assess differences and similarities in microbial composition within individuals and across the assignment groups. Alpha diversity estimates diversity within a single participant and encompasses measures of richness (number of microbial taxa) and evenness (distribution of taxa). To understand compositional shifts in the microbiome with changes in exposures Shannon's alpha diversity was compared within each group over time using a pairwise test and found that individuals who initially smoked and completely switched to e-cigarettes for 8 weeks, had significant differences in the Shannon index between baseline and week 8 ($p=0.027$, **Fig. 2-5A**). Unlike previous reports [22,63,64] that found an increase in alpha diversity with e-cigarette use over time significant decrease in diversity according to the Shannon index in complete switchers to e-cigarettes when compared to cigarettes and dual users was observed in this cohort **(Fig.2-5A, B).** Propylene glycol, constituent of e-cigarette liquids, has been shown to have a bactericidal effect and can therefore modulate the oral microbiome[75] which may explain the findings in this study of an overall lower diversity in complete switchers to e-cigarettes. The observed increase over time in alpha diversity in previous reports may be due to most participants presenting some level of periodontitis indicating a progression in periodontitis over time.[76]

The finding of statistically significant differences across product assignment groups at week 8 (Fig 2-4B) are consistent with a previous study that assessed the oral microbiome of 18 e-cigarette users (56% of whom also smoked cigarettes) and 18 controls found that dual users exhibited higher alpha diversity (Shannon Index) than exclusive e-cigarette users.[77] A longitudinal clinical study evaluating the adverse effects of e-cigarette use on periodontal health[78] found no significant differences in Shannon Index between cohorts (cigarettes; $n=27$, e-cigarettes users; $n=28$), and controls; $n=29$). All study participants had at least mild periodontitis, a condition resulting from imbalances in the microbial community inhabiting the periodontal pocket and subsequent host immune and inflammatory responses. An earlier study from the same group has also shown an association between periodontal status and the microbiome in e-cigarette users.[64] To put these findings into perspective, oral health and other related metrics must be carefully considered as potential contributors when discussing the potential effect of cigarettes and e-cigarette exposures on the oral microbiome.

Unlike alpha diversity, which reflects microbiome diversity in an individual person, beta diversity is a measure of dissimilarity across multiple individuals. Bray-Curtis dissimilarity matrices were used to assess beta diversity . There were no difference in beta diversity among study participants at baseline while they were smoking cigarettes. However, after 8 weeks of assigned product use, those who continued smoking, became dual users or completely switched to e-cigarettes showed different oral microbiome composition (Fig.2-6). These results are in line with previously reported cross-sectional assessments of beta diversity among different patterns of tobacco product use.[22,63] More importantly, a compositional shift in individuals who smoked and were randomly assigned to specific product use for 8 weeks was found, evidence that strongly emphasizes the effect of product type used on the structure of oral microbial communities..

Differentially abundant taxa among those with different product use status were identified to further the analysis of microbial profiles. Within the oral microbiome of those who initially smoked but switched completely to e-cigarette use, certain taxa, such as *Streptococcus, Rothia, Gemella, Actinomyces, Granulicatella, Veillonella, Haemophilus,* and *Fusobacterium*, were enriched. These results are in agreement with previous studies that have also reported an increase in the presence of these taxa among e-cigarette users compared to smokers, raising an interesting question about the potential enrichment mechanisms due to altered tobacco exposures and the possible health effects. There are few potential mechanisms by which e-cigarette aerosol exposure could contribute to formation of distinctive microbial profiles. Studies have shown that cigarette smoking alters the oral microbiome through

immunosuppressive effects , biofilm formation, altered O_2 tension and pH, and changes the chemical environment.[15] E-cigarette use has also been linked to increased inflammation, impaired mucosal defense, and an increased risk of periodontal disease.[64,78] In addition, e-cigarettes can damage the oral mucosa, resulting in a decrease in salivary flow and an increase in pathogen colonization.[79] E-cigarette aerosols can promote a microenvironment on enamel that is favorable to microbial adhesion and biofilm formation, with propylene glycol and glycerol potentially being the main drivers of the observed microbial shifts in e-cigarette users.[63] Studies have demonstrated a high similarity between primary biofilms exposed to nicotine-containing or nicotine-free aerosolized e-liquids containing propylene glycol and glycerol but not to aerosolized propylene glycol-free nicotine.[80] Biofilm formation requires significant energy investment and carbon resources. As a carbon source, glycerol is used by bacteria to synthesize lipids, facilitating the production of extracellular matrix and other essential compounds, as well as contributing to the construction of bacterial cell walls. It is also known to play a role in the intracellular growth of pathogenic bacteria. Therefore, enrichment of specific taxa in the oral microbiomes of e-cigarette users compared to individuals who smoke may be influenced by the microenvironmental changes caused by e-cigarette aerosols and the presence of constituents like glycerol and propylene glycol.

Distinctive profiles in individuals who became dual users at the end of the study period, with several genera showing decreased abundances compared to baseline were found. These included *Leptotrichia, Lactobacillus, Limosilactobacillus, Neisseria, Streptococcus, Actinomyces, Haemophilus*, and *Lancefieldella.* The shift in abundance of these taxa indicates that the introduction of e-cigarettes as another tobacco product, in addition to cigarettes, prompted shifts in the oral microbiome composition. Furthermore, in a cross-sectional analysis at week 8 the three groups: individuals who smoke, dual users, and exclusive e-cigarette users were compared. A total of 28 genera were found to be differentially abundant between individuals who smoke and e-cigarette users. *Pasteurellaceae, Leptotrichia, Actinobacillus, Amniculibacterium*, and *Neisseria* were significantly decreased in e-cigarette users. *Limosilactobacillus, Ligilactobacillus, Streptococcus, Fretibacterium, Lactobacillus*, and *Campylobacter* were more abundant in e-cigarette users compared to dual users. Individuals who smoke and e-cigarette users have distinct microbiomes that are influenced by specific exposures related to these products, resulting in unique microbial compositions. This indicates that the oral microbiome has the potential to serve as a tool for assessing different tobacco exposures.

This study has several limitations. First, the study population did not include individuals who did not use any form of tobacco. While the the primary objective of this study, which was to understand the potential sensitivity of the oral microbiome to changes in tobacco product exposure, was addressed by the current approach, future research should compare switching to e-cigarettes vs. complete cessation of any product. Such studies will help to interpret the potential implications of the oral microbiome changes after switching from smoking to e-cigarette use.

Also, certain factors that may contribute to the oral microbiome were not included in this study, such as diet, alcohol consumption, oral hygiene, and even the host's genetics. Additionally, because of the limitations of self-reporting, the duration and intensity of e-cigarette use were not accounted for in the analyses in this study. With the advent of more advanced versions of e-cigarettes, users have greater control over the quantity of e-liquid usage, power, and airflow settings. As a result of this customization, the oral microenvironment can be altered further, and the microbiome can be affected as well. Lastly, the study did not use metagenomic sequencing, which could provide further insight into the responsiveness of the microbial composition and function with increased resolution. Metagenomics can also provide information on microbial interactions and metabolic pathways subsequently contributing to differences in tobacco-related exposures and health outcomes.

Although the number of study participants who completely switched to e-cigarette use is limited, significant strength in determining the direction of changes in microbial composition has been harnessed by utilizing biochemically confirmed switchers. By using biomarkers, such as CO, CEMA, and NNAL, individuals were reliably categorized based on the product that they have used. The validity and robustness of these findings was strengthen by focusing on confirmed switchers and eliminating confounding factors such as misreported smoking behavior. Consequently, any changes in the oral microbiome can be attributed to the specific shift in tobacco-related exposures. This approach enabled to provide insights into the specific effects of e-cigarette use on the oral microbiome, leading to a more comprehensive understanding of the microbial changes associated with different tobacco-related exposures. Therefore, these results indicate that switching from smoking to exclusive e-cigarette use led to a notable shift in oral microbial composition and provides insights into the potential usefulness of the microbiome as a receptor matrix for testing tobacco-associated exposures.

TABLES AND FIGURES FOR CHAPTER 2

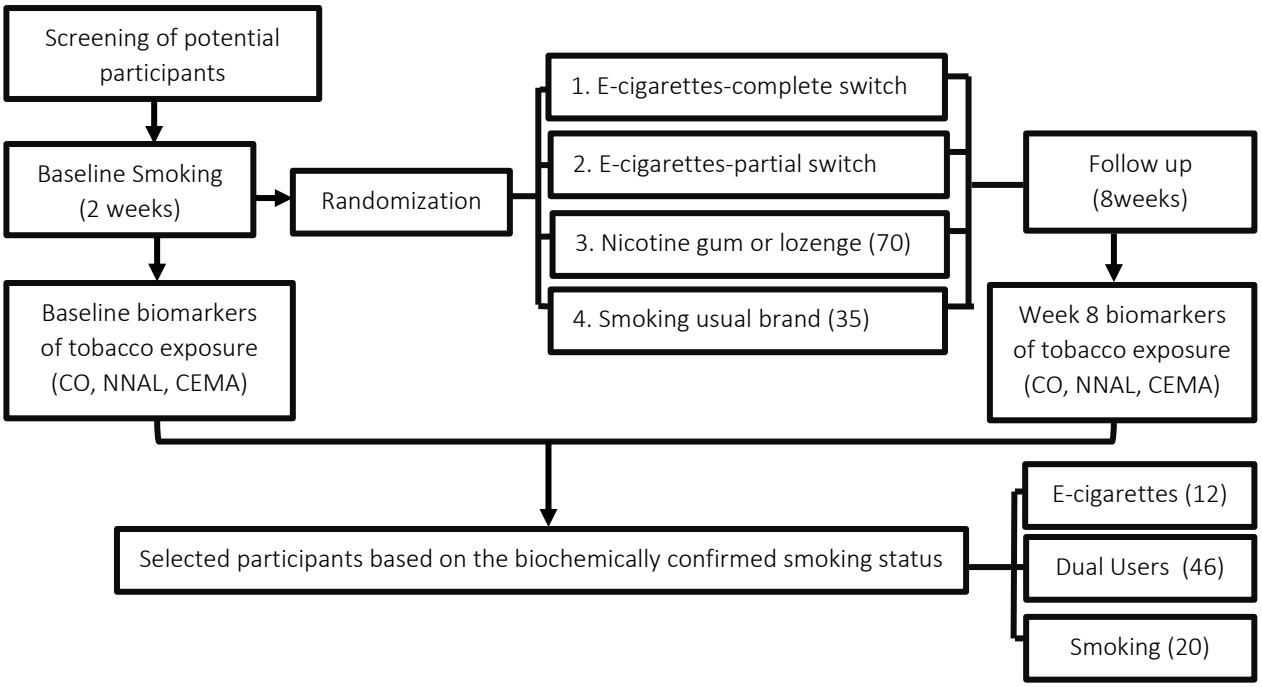


Figure 2-1. Schematic representation of the overall COMET study design and selected subset of participants with biochemically confirmed smoking status by week 8 included in current study

Table 2-1. Study participant's characteristics and biomarkers – summary statistics

	Total (N=78)	Smoking (N=20)	E-cigs (N=12)	Dual (N=46)	p-value
Race (AA)	42 (53.8%)	10 (50.0%)	4 (33.3%)	28 (60.9%)	0.141
Age (Years)	47.5 (37.3-55.8)	46.5 (37.8-56.3)	58 (42.8-62.8)	47.5 (34-54)	0.176
Sex (Female)	41 (52.6%)	8 (40.0%)	8 (66.7%)	25 (54.3%)	0.319
Years of smoking	28 (19-38)	30 (18.5-39.5)	38 (28-45.5)	23.5 (18.3-37.8)	0.095
Biomarkers at baseline					
Cigarettes per day (CPD)	12.07 (9-16.7)	12.07 (9.3-21.2)	14.59 (10.6-17.8)	11.54 (8.2-16.6)	0.622
CO	16 (11-23.3)	14 (11-21)	20 (14.5-28.3)	15.5 (11.3-19)	0.232
CEMA (pmol/mL urine)	490.23 (277.7-769.1)	351.155(193.4-603.7)	597.035 (364-990.2)	576.64 (265.1-778.9)	0.097
NNAL (pmol/mL urine)	1.27 (0.81-1.65)	1.08 (0.8-1.4)	1.85 (1.23-2.72)	1.06 (0.75-1.65)	0.067
TNE (pmol/mL urine)	62.17 (39.86-86.37)	60.13 (39.37-80.88)	70.43 (31.39-102.31)	61.56 (40.44-85.25)	0.931
Biomarkers at week 8					
Cigarettes per day (CPD)	8.81 (3.31-12.18)	11.105 (9.52-19.63)	0	8.72(5.23-13)	**< 0.001**
CO	12.5 (5.25-18)	14 (10.75-19.5)	3 (2-3)	14 (8.25-18.75)	**< 0.001**
CEMA (pmol/mL urine)	387.65 (152.35-685.17)	428.33 (320.91-686.11)	28.42 (12.56-45.61)	461.39 (222.77-734.7)	**< 0.001**
NNAL (pmol/mL urine)	0.81 (0.39-1.44)	1.15 (0.76-1.52)	0.18 (0.14-0.29)	0.89 (0.52-1.6)	**< 0.001**
TNE (pmol/mL urine)	66.18 (37.85-93.92)	55.94 (40.18-90.23)	67.33(17.24-81.47)	71.39 (42.37-94.36)	**< 0.001**

*Summaries shown are median (1st quartile, 3rd quartile) or N (percent)

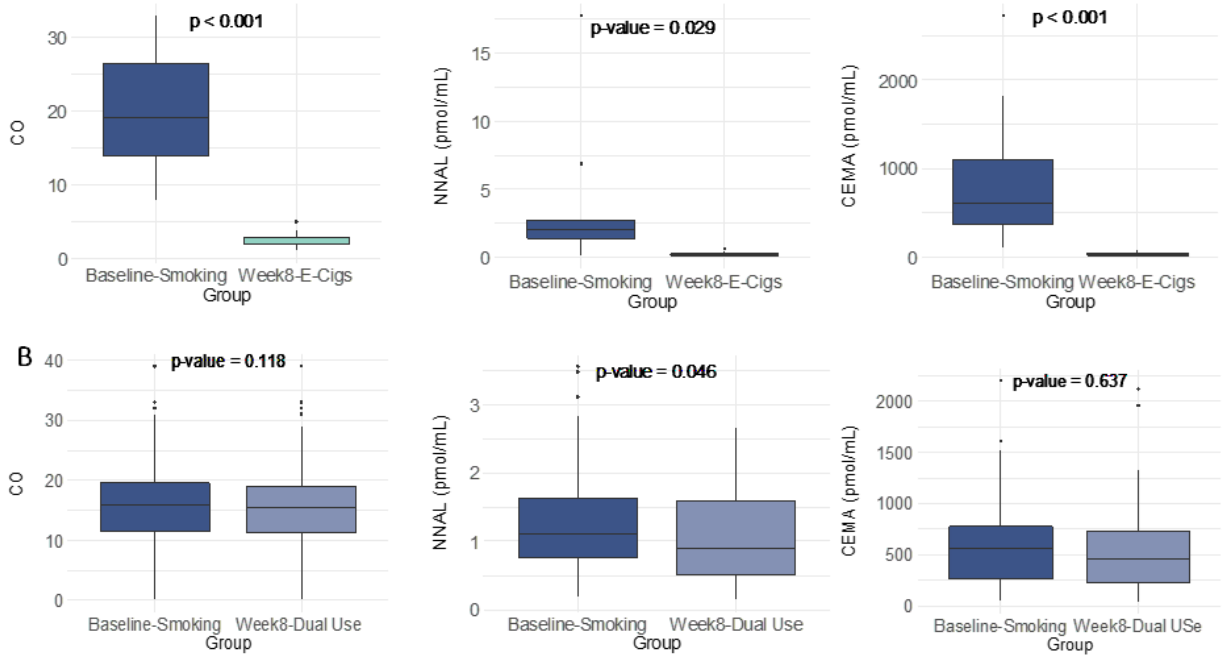


Figure 2-2. Differences in biomarkers of smoking exposure (CO, NNAL, CEMA) between baseline and end of study period, week 8 within those who became A) exclusive e-cigarette users; B) dual users of cigarettes and e-cigarettes.

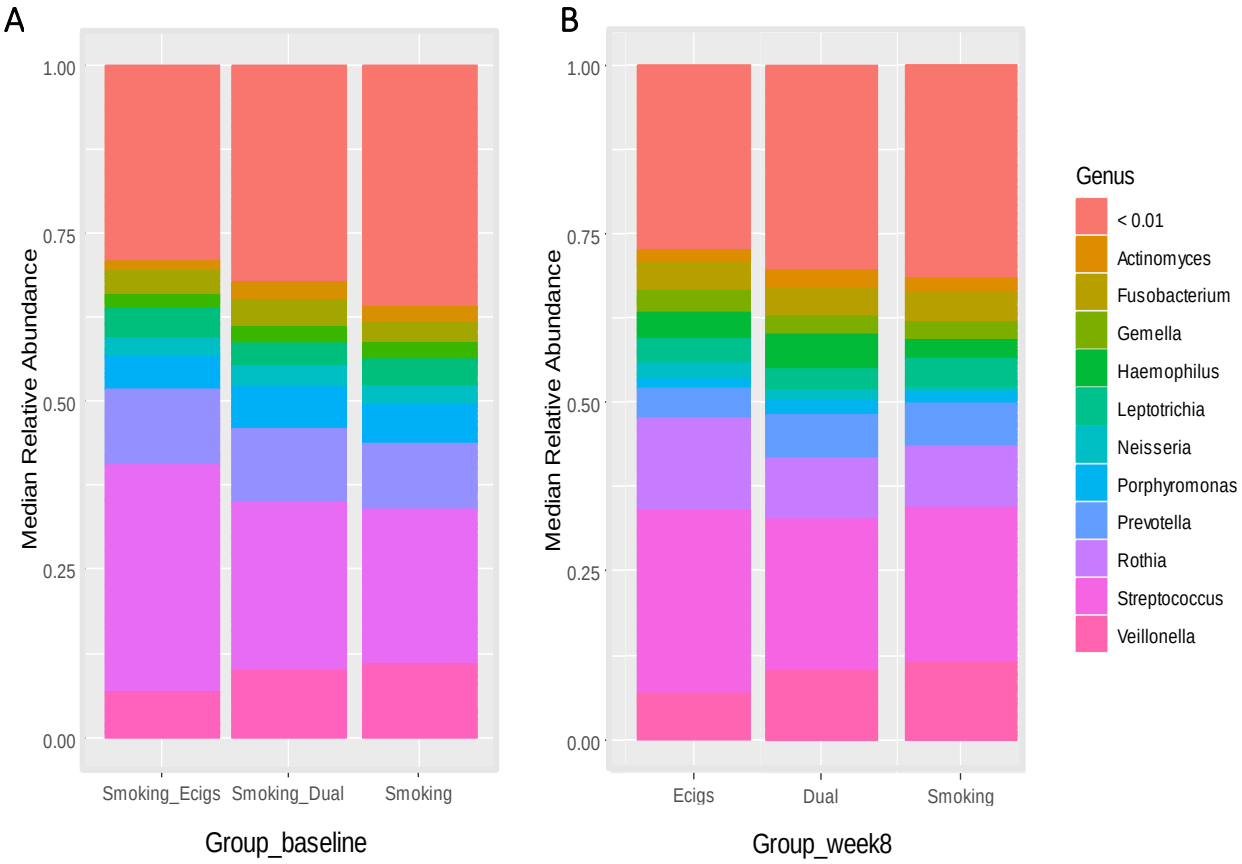


Figure 2-3. Relative abundance of bacterial genera across groups randomized to different product use: A) at baseline, when all participants were smoking; and B) at week 8 (end of study).

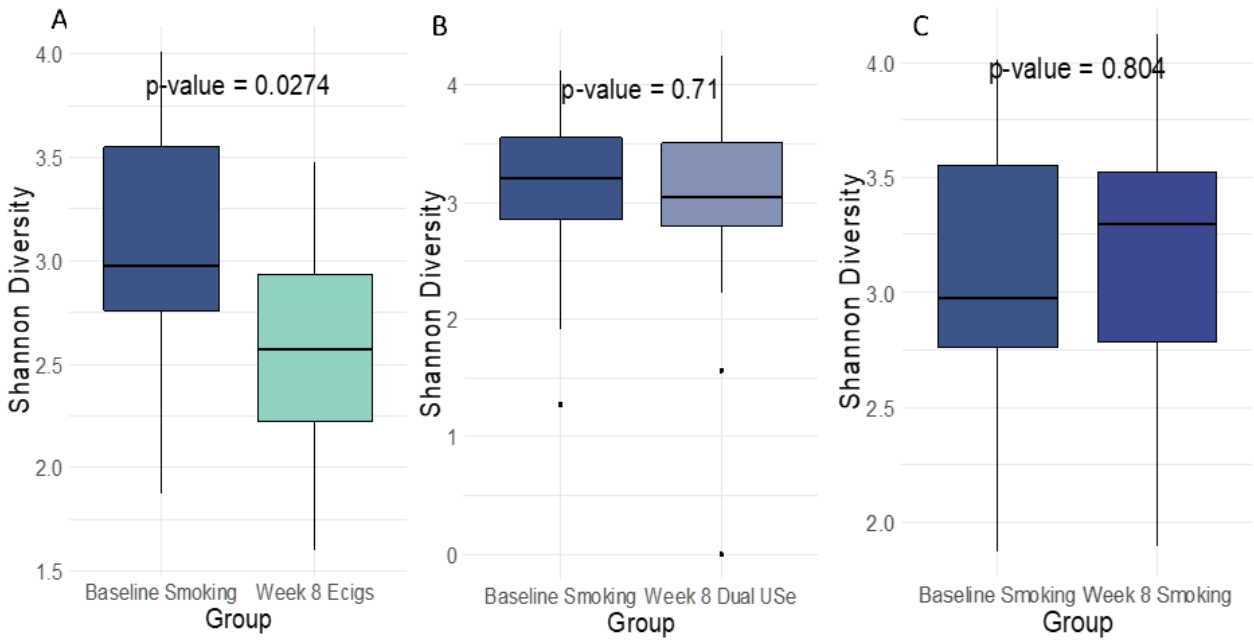


Figure 2-4. Differences in alpha diversity measure (Shannon Index) between baseline and week 8 within groups of participants who became: A) exclusive e-cigarette users; B) dual users; C) continued smoking

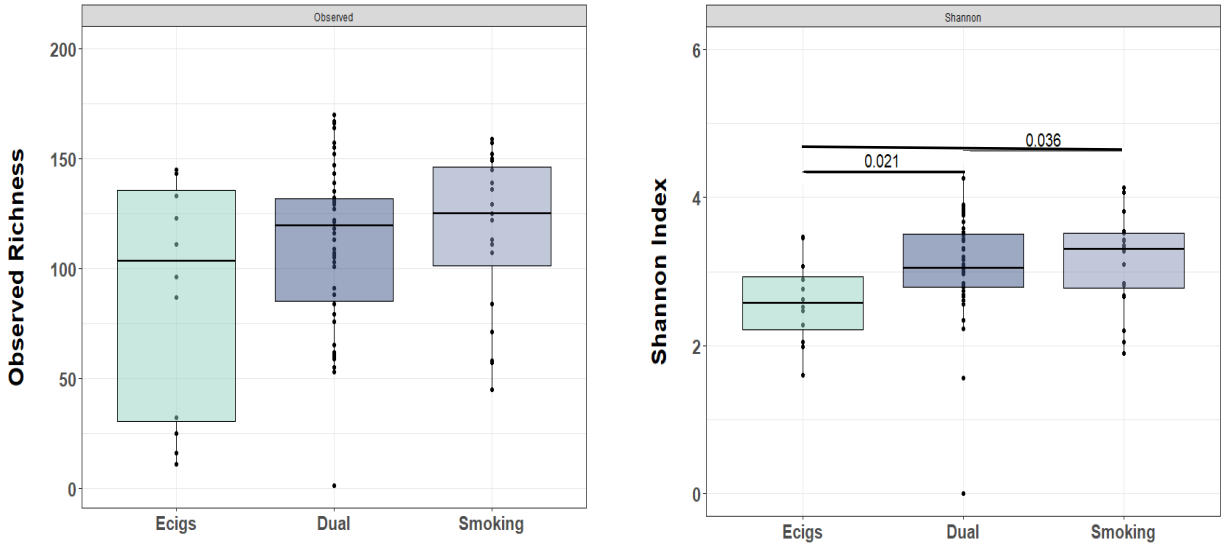


Figure 2-5. Alpha diversity measures A) Observed Richness and B) Shannon Index across participants using different products at the end of the study period, week 8

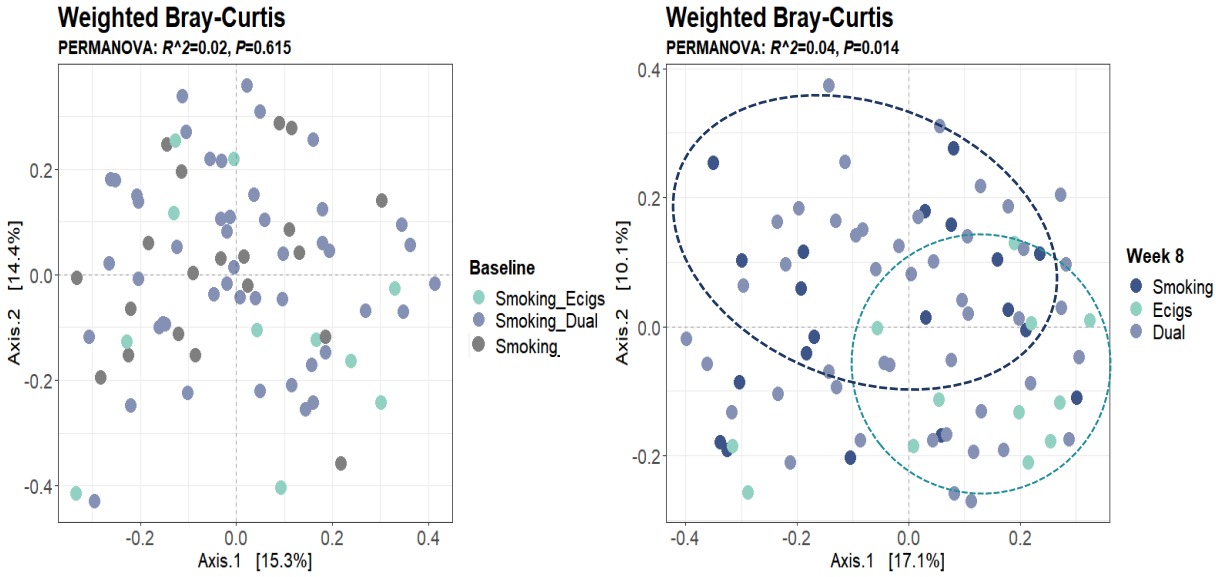


Figure 2-6. Beta diversity (Weighted Bray Curtis) in all participants at A) baseline while smoking and B) at the end of the study period, week 8. Each point on the plots represents microbial composition of a sample, and the distances between points reflect their Bray-Curtis dissimilarity. Samples that cluster

closely together share similar microbiome profiles, while those farther apart exhibit greater dissimilarity. The scatter plots reveal clustering patterns of samples based on their microbial composition between those who continued smoking, became dual users, or completely switched to e-cigarette use at week 8 (Panel B: r^2 =0.04; p=0.014), but no clustering at baseline (Panel A).

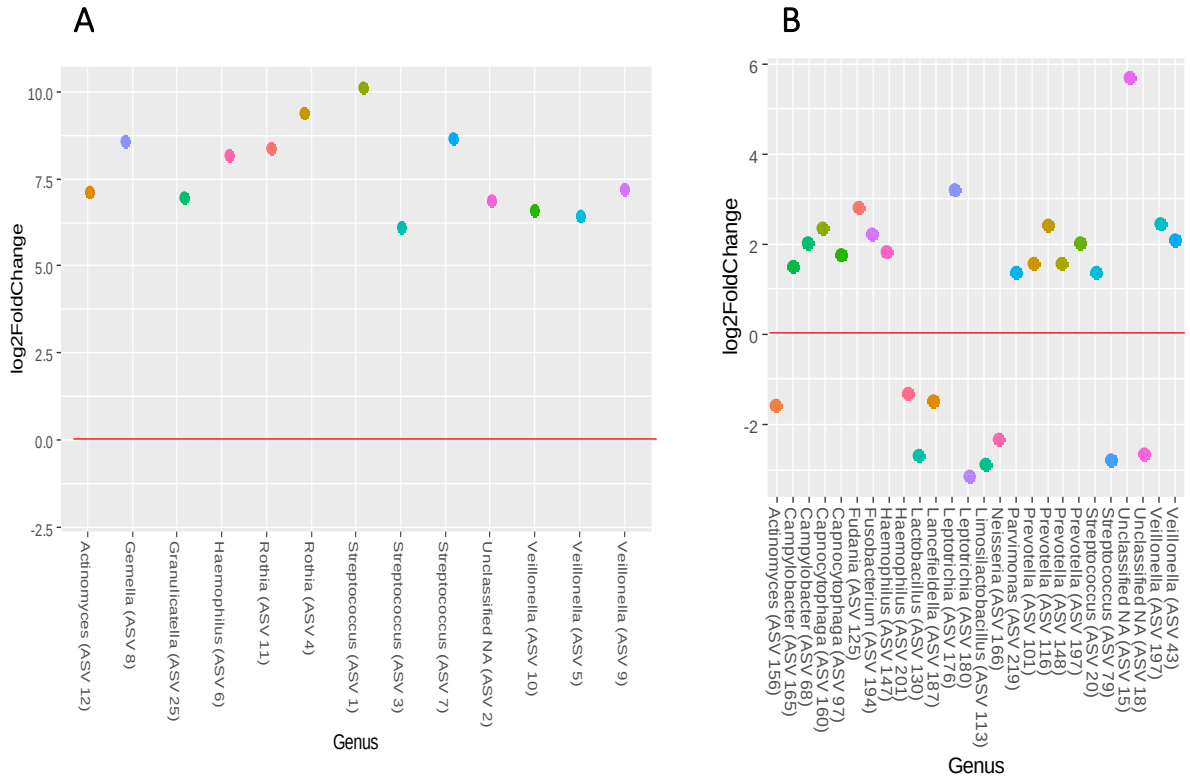


Figure 2-7. Differential abundance (ASV level agglomeration) between the end of the study period (week 8) and baseline within participants that became A) exclusive e-cigarette users; B) dual users. Each data point represents a genus-level ASVs (x-axis) identified as significantly different along with the log2 fold change (y-axis).

*Positive log2 fold change values indicate that the abundance has increased while negative values indicate a decrease at the end of the study period (week 8)

** padj or q-value – FDR corrected significance level significance level (alpha) set at 0.01

SUPPLEMENTAL DATA FOR CHAPTER 2

Supplemental Table 2-1. Differently abundant taxa between Smoking and E-cigarettes use groups – cross sectional comparison at week 8

ASV level agglomeration	baseMean	log2FoldChange*	lfcSE	stat	pvalue	Padj**
Neisseria (ASV 44)	163.0714311	-3.451424302	1.040533416	-3.316975938	0.000909975	**0.008371766**
Unclassified Pasteurellaceae (ASV 160)	11.93997746	-2.986871203	0.766039086	-3.899110708	9.65E-05	**0.001586123**
Amniculibacterium (ASV 81)	61.36862637	-2.825820074	0.781373659	-3.616477268	0.00029864	**0.003122141**
Unclassified Pasteurellaceae (ASV 227)	3.550594306	-2.804302875	0.681826536	-4.112927158	3.91E-05	**0.000991612**
Actinobacillus (ASV 203)	4.879069859	-2.513314915	0.678751241	-3.7028513	0.00021319	**0.002724093**
Leptotrichia (ASV 195)	3.801787545	-2.421601662	0.609978041	-3.969981703	7.19E-05	**0.001377665**
Campylobacter (ASV 103)	29.43077858	2.312762881	0.708902503	3.262455516	0.001104515	**0.008526463**
Dialister (ASV 89)	36.9421093	2.438909996	0.748016453	3.260503145	0.001112147	**0.008526463**
Fretibacterium (ASV 172)	9.951682931	2.445307909	0.737171703	3.317148364	0.000909413	**0.008371766**
Stomatobaculum (ASV 220)	9.577200764	2.504091648	0.74847232	3.345603548	0.000821037	**0.00821037**
Actinomyces (ASV 64)	20.64547909	2.571680925	0.784665422	3.277423539	0.001047591	**0.008526463**
Fretibacterium (ASV 100)	22.2087138	2.597002561	0.704773598	3.684874928	0.000228815	**0.002769865**
Veillonella (ASV 5)	1310.187023	2.634139726	0.704305308	3.740053774	0.000183981	**0.002489153**
Bifidobacterium (ASV 86)	17.38409852	2.653056028	0.808618176	3.28097501	0.001034489	**0.008526463**
Actinomyces (ASV 156)	9.176752895	2.724817475	0.691461186	3.940665838	8.13E-05	**0.001437602**
Lactobacillus (ASV 102)	11.72946237	2.830641897	0.861892015	3.284218729	0.001022655	**0.008526463**
Candidatus Saccharibacteria (ASV 52)	35.12029189	2.845192966	0.7463657	3.812062861	0.000137812	**0.002113114**
Streptococcus (ASV 93)	23.99327992	3.04137379	0.747639594	4.067967796	4.74E-05	**0.000991612**
Lactobacillus (ASV 130)	5.907752388	3.267927267	0.773432389	4.225226812	2.39E-05	**0.000686263**
Prevotella (ASV 96)	20.44350246	3.367635701	0.923534234	3.646465475	0.000265872	**0.002911933**
Campylobacter (ASV 106)	11.32521787	3.395303178	0.790116544	4.29721818	1.73E-05	**0.000568281**
Streptococcus (ASV 35)	85.68269137	3.662925406	1.000505928	3.661073165	0.000251161	**0.002888352**
Veillonella (ASV 137)	8.350999916	3.705002275	0.80911256	4.579093761	4.67E-06	**0.000214818**
Limosilactobacillus (ASV 63)	25.37226934	3.908469632	0.957255526	4.082995109	4.45E-05	**0.000991612**
Limosilactobacillus (ASV 70)	35.71492544	4.0465784	0.911185487	4.44100401	8.95E-06	**0.000343237**
Scardovia (ASV 30)	79.83176212	4.335308507	0.931228813	4.655470756	3.23E-06	**0.000185864**
Lactobacillus (ASV 32)	63.75585095	4.48401433	0.929034132	4.826533469	1.39E-06	**0.000106513**
Limosilactobacillus (ASV 113)	15.66071001	4.79136706	0.810948019	5.908352876	3.46E-09	**7.95E-07**
Ligilactobacillus (ASV 58)	38.39193411	4.825545697	0.970912584	4.970113456	6.69E-07	**7.70E-05**

*positive log2FoldChange indicates an increase in abundance, while a negative log2FoldChange indicates a decrease in abundance in the Smoking group compared to the E-cigarettes use group;

** padj or q-value – FDR corrected significance level significance level (alpha) set at 0.01

Supplemental Table 2-2. Differently abundant taxa between Dual and E-cigarettes use groups – cross sectional comparison at week 8

ASV level agglomeration	baseMean	log2FoldChange*	lfcSE	stat	pvalue	Padj**
Leptotrichia (ASV 72)	27.16780357	-2.407984363	0.744508517	-3.234327492	0.001219296	0.008970808
Leptotrichia (ASV 180)	7.150232009	-2.3264842	0.613926957	-3.789513025	0.000150943	0.002812425
Leptotrichia (ASV 195)	3.801787545	-1.852897107	0.521719481	-3.551519879	0.000383013	0.004894056
Veillonella (ASV 10)	597.8220493	2.029101033	0.584326678	3.472545594	0.000515547	0.006240837
Stomatobaculum (ASV 78)	32.86400533	2.105463525	0.613594828	3.431358007	0.000600567	0.006906526
Selenomonas (ASV 80)	23.06826503	2.280382286	0.689190341	3.308784451	0.000937019	0.008289018
Selenomonas (ASV 182)	11.56733141	2.347638135	0.718386691	3.267930995	0.001083368	0.008899092
Streptococcus (ASV 20)	150.698757	2.384822974	0.618651134	3.854875296	0.000115789	0.002421035
Campylobacter (ASV 103)	29.43077858	2.393247826	0.636819582	3.758125366	0.000171191	0.002812425
Candidatus Saccharibacteria (ASV 52)	35.12029189	2.421933848	0.670116612	3.614197594	0.000301279	0.004330889
Prevotella (ASV 204)	9.535389746	2.431472673	0.720692663	3.373799679	0.000741383	0.0076067
Lancefieldella (ASV 88)	55.06273473	2.45006799	0.744299864	3.291775409	0.000995571	0.008480789
Capnocytophaga (ASV 82)	26.28116746	2.481958126	0.65808365	3.771493376	0.000162273	0.002812425
Streptococcus (ASV 59)	92.43434158	2.487824977	0.750836254	3.313405504	0.000921672	0.008289018
Prevotella (ASV 177)	11.31423723	2.509223423	0.735551656	3.411349021	0.000646423	0.00707987
Campylobacter (ASV 68)	49.39376026	2.528566742	0.70962046	3.5632664	0.000366269	0.004894056
Porphyromonas (ASV 95)	7.588210175	2.533901601	0.783999797	3.232018186	0.001229192	0.008970808
Prevotella (ASV 22)	207.3864323	2.696767044	0.835520078	3.227650793	0.001248112	0.008970808
Dialister (ASV 89)	36.9421093	2.72648337	0.670801098	4.06451835	4.81E-05	0.001970089
Prevotella (ASV 96)	20.44350246	2.808056924	0.834062308	3.366723203	0.00076067	0.0076067
Prevotella (ASV 147)	9.312898465	2.834537831	0.732761541	3.868295033	0.000109599	0.002421035
Fusobacterium (ASV 120)	28.92961297	2.844023013	0.728272238	3.905164669	9.42E-05	0.002406343
Ligilactobacillus (ASV 58)	38.39193411	2.920982476	0.875471597	3.33646744	0.000848504	0.008131494
Veillonella (ASV 43)	134.2477979	2.945721127	0.803072368	3.668064352	0.000244394	0.00374737
Veillonella (ASV 21)	137.9707837	3.032223909	0.934051354	3.246313916	0.001169098	0.008970808
Fusobacterium (ASV 192)	11.35212442	3.06121797	0.758705259	4.034792083	5.47E-05	0.001970089
Lancefieldella (ASV 115)	21.49471958	3.080925173	0.767741656	4.012971222	6.00E-05	0.001970089
Fudania (ASV 125)	14.87518879	3.14319671	0.766243661	4.102085105	4.09E-05	0.001970089
Streptococcus (ASV 35)	85.68269137	3.536216926	0.892835684	3.960658146	7.47E-05	0.002148875
Scardovia (ASV 30)	79.83176212	3.806893587	0.833959528	4.564842127	5.00E-06	0.000574851
Unclassified Prevotellaceae (ASV 47)	64.05030907	3.809060874	0.873163484	4.362368493	1.29E-05	0.000986408

*positive log2FoldChange indicates an increase in abundance, while a negative log2FoldChange indicates a decrease in abundance in the Dual use group compared to the E-cigarettes use group;

** padj or q-value – FDR corrected significance level (alpha) set at 0.01

Supplemental Table 2-3. Differently abundant taxa between Smoking and Dual use groups – cross sectional comparison at week 8

ASV level agglomeration	baseMean	log2FoldChange*	lfcSE	stat	pvalue	Padj**
Neisseria (ASV 44)	163.0714311	-2.983815008	0.764751663	-3.90167835	9.55E-05	**0.001373425**
Veillonella (ASV 196)	8.133762044	-2.589769558	0.606154388	-4.272458652	1.93E-05	**0.000342044**
Prevotella (ASV 92)	8.058093638	-2.586557135	0.554615474	-4.663694497	3.11E-06	**9.74E-05**
Prevotella (ASV 175)	7.534136392	-2.561233647	0.551312207	-4.645704586	3.39E-06	**9.74E-05**
Unclassified Prevotellaceae (ASV 47)	64.05030907	-2.557583929	0.704542793	-3.630132839	0.000283275	**0.003212966**
Campylobacter (ASV 68)	49.39376026	-2.487534563	0.580178839	-4.287530667	1.81E-05	**0.000342044**
Prevotella (ASV 197)	8.51924716	-2.42035584	0.55343167	-4.373359838	1.22E-05	**0.000281402**
Abiotrophia (ASV 129)	33.6122918	-2.396376213	0.631578459	-3.794265271	0.000148081	**0.002003453**
Fusobacterium (ASV 105)	31.67737838	-2.365433692	0.647218601	-3.654767786	0.000257415	**0.003116075**
Porphyromonas (ASV 37)	49.62675683	-2.341216311	0.723571572	-3.23563888	0.001213708	**0.009035043**
Fudania (ASV 125)	14.87518879	-2.289545406	0.586815953	-3.901641383	9.55E-05	**0.001373425**
Leptotrichia (ASV 122)	27.70144441	-2.151939367	0.631705333	-3.406555642	0.000657881	**0.006304697**
Kingella (ASV 51)	103.388733	-1.958467612	0.605458381	-3.234685776	0.001217767	**0.009035043**
Mannheimia (ASV 90)	6.173928892	-1.895975466	0.588544408	-3.221465434	0.001275368	**0.009166711**
Prevotella (ASV 181)	4.210641988	-1.86965924	0.552690972	-3.38282935	0.000717432	**0.006461193**
Prevotella (ASV 53)	28.32471028	-1.762922973	0.552867475	-3.188689972	0.001429191	**0.009961026**
Capnocytophaga (ASV 97)	33.02086905	-1.753988004	0.521360424	-3.364252296	0.000767513	**0.006538078**
Streptococcus (ASV 225)	2.245100095	1.478993042	0.408437965	3.621095902	0.000293358	**0.003212966**
Porphyromonas (ASV 206)	3.642905522	1.511240735	0.428372976	3.527861978	0.000418931	**0.004379728**
Campylobacter (ASV 106)	11.32521787	1.717978989	0.524684655	3.274307665	0.001059212	**0.008700671**
Streptococcus (ASV 76)	6.173394301	1.837400521	0.538831528	3.409972179	0.000649695	**0.006304697**
Haemophilus (ASV 201)	2.549168208	1.899852498	0.423912674	4.481707239	7.40E-06	**0.000189234**
Veillonella (ASV 5)	1310.187023	1.92155359	0.516354068	3.721387531	0.000198131	**0.002531675**
Bifidobacterium (ASV 86)	17.38409852	1.945728477	0.57601587	3.37790776	0.000730396	**0.006461193**
Lactobacillus (ASV 102)	11.72946237	1.958322483	0.604589019	3.239097009	0.001199088	**0.009035043**
Lactobacillus (ASV 32)	63.75585095	2.671922883	0.666642391	4.008030269	6.12E-05	**0.001005877**
Limosilactobacillus (ASV 70)	35.71492544	2.824089449	0.650275751	4.342910592	1.41E-05	**0.000293997**
Veillonella (ASV 137)	8.350999916	2.918752404	0.532191584	5.484401656	4.15E-08	**1.59E-06**
Leptotrichia (ASV 180)	7.150232009	2.967103684	0.50414483	5.885419248	3.97E-09	**1.83E-07**
Lactobacillus (ASV 130)	5.907752388	3.260864941	0.518313581	6.291297509	3.15E-10	**2.41E-08**
Limosilactobacillus (ASV 113)	15.66071001	3.722189539	0.528818243	7.038693514	1.94E-12	**2.23E-10**

*positive log2FoldChange indicates an increase in abundance, while a negative log2FoldChange indicates a decrease in abundance in e Smoking group compared to the Dual use group;

** padj or q-value – FDR corrected significance level significance level (alpha) set at 0.01

CHAPTER 3

Relationship between the oral microbiome tobacco-induced biological effects in oral cells of individuals who smoke

Introduction

While smoking is the major preventable cause of cancer and cardiovascular diseases, not all people who smoke develop these diseases. For example, only 24% of all people who smoke will develop lung cancer.[81] The underlying mechanisms of the increased susceptibility to smoking-associated diseases in some individuals are not fully understood. The effects of smoking history and differences in the uptake of tobacco toxicants and carcinogens have been explored [58,59]; however, such factors do not fully account for the observed inter-individual variations in disease risks. Biomarkers of biological effects that are triggered by such exposures and are part of pathophysiology of the relevant diseases could potentially serve as a more direct and informative measure of the tobacco-induced harm. Macromolecular damage caused by inflammation, oxidative stress, and chemical carcinogens present in cigarette smoke is a promising category of biomarkers of biological effects that have been explored in some studies.

Inflammation response, oxidative stress, cell death, and genomic instability have all been implicated as toxicity mechanisms associated with cigarette smoke. Oxidative stress occurs when reactive oxygen species (ROS) and reactive nitrogen species (RNS) exceed the neutralizing capacity of the cell. These species at higher concentrations could cause oxidative and nitrosative damage to most cellular components, including membrane lipids, enzymes, and DNA, which plays an important role in the pathogenesis of cigarette-smoke-induced diseases. Biomarkers of inflammation and oxidative stress, such as isoprostanes and other products of protein and DNA oxidation, have been used extensively in studies of smoking, showing higher levels of such biomarkers in people who smoke compared to those who do not.[82] Mitochondrial DNA (mtDNA) copy number offers a potentially more effective alternative to measuring individual biomarkers because it represents a measure of cumulative effect of inflammation and oxidative stress. The mitochondrial respiratory chain generates the majority of ROS by incomplete reduction of molecular O_2 to H_2O during oxidative phosphorylation and during microsomal and peroxisomal oxidations. A number of tobacco toxicants specifically affect mitochondria because of their inherent characteristics, either by inhibiting mitochondrial enzymes or accumulating within them.[83] In mitochondria, lipophilic compounds (such as PAH) accumulate in the membranes, whereas xenobiotics accumulate in the organelle

as a result of their negative charge and alkaline pH. These characteristics, along with the association of mitochondrial DNA with the inner membrane (lipophilic environment), the absence of histones, relatively low levels of protective proteins (nucleoids), and the presence of cytochrome P450 enzymes that are capable of converting compounds into active toxicants, make mtDNA a prime target for tobacco toxicants. Indeed, as outlines in Chapter 1, persons who smoke have elevated mtDNA copy number, which in turn is associated with the risk for lung cancer.[28-32]

Measurement of tobacco constituent-derived DNA adducts in oral cells has been expanding in recent years, and represents an exciting direction in biomarker research. The levels of such DNA adducts in oral cells far exceed those measured in blood or tissues, potentially because of the direct exposure to, and local metabolism of, the high levels of reactive constituents present in tobacco product emissions. For example, in previous studies, HPB-releasing DNA adduct in oral cells of people who smoke averaged 6.2 pmol/mg DNA in oral cells[84] while levels of these adducts in blood DNA or in the lung tissues are challenging to measure without high-resolution equipment. Similarly, substantial levels of N^2-ethyl-dG, an acetaldehyde-derived DNA adduct were detected in oral cells, while levels of the same adduct in blood were generally detected in lower yield.[85] The reasons for examining the effects of tobacco exposure in the oral cavity extend beyond the mere quantification of adducts; it is rather linked to direct exposures and the unique local microenvironment of the oral cavity, further supporting the importance of the oral cavity as the easily accessible and relevant compartment for assessing macromolecular damage and potentially other biological effects caused by tobacco-related exposures.

Despite their promise, biomarkers such as mtDNA content and individual DNA adducts have potential limitations that may limit their applicability to studies of tobacco product use. For example, mtDNA measurement in studies of tobacco product use is that it is not likely to be very sensitive to changes in product use status or patterns; a study of mtDNA in oral cells showed that the smoking-associated increases in mtDNA copy number persist for over 20 years after smoking cessation.[28] An important limitation of carcinogen-specific DNA adducts is that they are usually measured one at a time, with the assays being expensive, time-consuming, and requiring specialized expertise. Therefore, the field would benefit from innovative tools, such as the oral microbiome, for assessing a broad spectrum of biological effects resulting from tobacco or other nicotine-containing product use. Mechanistically, the oral microbiome may be involved in chemical carcinogenesis through immune-inflammatory responses and carcinogen metabolism and production, among other potential mechanisms.[86] Therefore, the composition

and functional pathways of the oral microbiome, combined with an appropriate set of biomarkers of effect, can potentially be used to predict the overall landscape of inflammatory and chemical carcinogen-induced DNA damage on the oral cavity of persons who use tobacco products. Indeed, as outlined in Chapter 1, a study recently reported an association between microbial characteristics and HPB-releasing DNA adducts in the oral cells of individuals who smoke and had HNSCC.[38] This could be driven, at least in part, by nitrate-reducing oral bacteria. These microorganisms convert dietary nitrate to nitrite and other nitrosating species,[87] which can further react with tobacco alkaloids and their metabolites and form NNN and NNK (the sources of HPB-releasing DNA adducts). Studies have demonstrated such endogenous formation of NNN in the oral cavity of persons consuming nicotine, including e-cigarette users.[88-90]

The goal of the research in this chapter is to further investigate the association between the oral microbiome composition and biomarkers of tobacco-derived macromolecular damage in oral cells. The biomarkers were selected to represent damage caused by oxidative and inflammatory agents (mtDNA content) and the tobacco-specific carcinogen DNA modifications (HPB-releasing DNA adducts). Available biological samples and data from a study that recruited 160 healthy persons who smoke and analyzed a wide range of tobacco-related biomarkers, including HPB-releasing DNA adducts in oral cells was used. The levels of HPB-releasing adducts in that study varied 56-fold across all study participants, which offered an opportunity to investigate the potential association of oral microbiome with these variations. In this study, the available oral DNA samples from the same individuals was used to measure mtDNA content and characterize their oral microbial communities, and to investigate the relationship between the oral microbiome and (i) mtDNA content and (ii) HPB-releasing DNA adducts in oral cells. It is expected that oral microbiome composition will be associated with mtDNA content and the levels of HPB-releasing DNA adducts in the oral cells of these individuals.

Materials and Methods

Overview of study design and study procedures in the Mechanisms of Ethnic/Racial Differences in Lung Cancer Due to Cigarette Smoking study

Oral cells collected from participants that took part in the Mechanisms of Ethnic/Racial Differences in Lung Cancer Due to Cigarette Smoking Study, Masonic Cancer Center, University of Minnesota were analyzed as part of the this study. The parent study was approved by the University of Minnesota Institutional Review Board (Study No. 1007M85757). The main study goal was to investigate differences in NNK metabolic activation and detoxification and the formation of NNK-derived DNA adducts and related

DNA repair capacity in African American and White individuals who smoke. Participants were recruited through campus and metropolitan newspapers, radio, and television advertisements. A smoking history of at least ten cigarettes daily for at least one year and good physical and mental health were required. The study excluded participants who were: (i) suffering from unstable medical conditions such as cancer, coronary heart disease, or arrhythmia; (ii) pregnant or breastfeeding; (iii) using other tobacco or nicotine-containing products; (iv) taking medications that affected metabolic enzymes; and (v) unable or unwilling to identify their ethnic/racial heritage. Participants who met the study criteria attended an orientation in the research clinic.

Those who were eligible for the study signed written consent and completed two structured questionnaires, a tobacco history questionnaire and a medical history questionnaire. Each participant was asked to record each cigarette smoked on a daily diary card. Participants were asked to refrain from using any other nicotine-containing products; however, if any are used, they should record them. During the second visit, urine, exfoliated oral mucosa, and blood samples were collected. The participants' vital signs (heart rate, blood pressure) and carbon monoxide levels (CO) were also measured. After this visit, participants switched to specially prepared study cigarettes containing deuterium labeled NNK and returned to the clinic for subsequent visits. This study used oral cell samples collected at baseline while participants smoked their usual cigarette brand. Only samples from participants who consented to future storage and analysis of their biospecimens were included in this study.

Collection of oral samples

Participants were asked to brush their teeth one hour before they visited the Tobacco Research Programs Clinic. They were also asked not to smoke, drink, or eat during that time. During the clinic visit, they received two sterile cytobrushes and were instructed to brush the inside of their left and right cheeks with a new brush each. Oral samples were collected under the supervision of a trained study coordinator. The brushes were placed in polypropylene tubes containing 10 ml of Scope mouthwash (14.3 wt% alcohol). The cells were pelleted by centrifugation at 1500 g for 15 minutes, the supernatant was discarded, and the pellets were rinsed with saline and stored at -20 °C for DNA isolation. Other biological samples, such as urine and blood, were collected for biomarker analysis. In addition, alveolar carbon monoxide (CO) levels were measured, along with vital signs (heart rate, blood pressure), daily cigarette diary cards were taken, and the time since smoking the last cigarette was recorded. Participants were compensated for their time, effort, and transportation.

DNA isolation from oral mucosa cells and oral bacteria community profiling

The oral cells were processed and analyzed in the same way as previously described. [35] In brief, cells were suspended in 100 mM phosphate buffer (pH 6.8), treated with RNase A and proteinase K, and DNA was extracted using the QIAGEN DNA Mini Kit. Most of the isolated DNA was previously used for the analysis of mtDNA content and HPB-releasing DNA adducts in these study participants. The remaining available DNA was used for microbiome analyses. Oral bacteria community profiling was conducted at the University of Minnesota Genomics Center (UMGC). To assess the quality of the DNA, DNA purity was measured by the qubit dsDNA HiSensitivity Kit yield. An Illumina MiSeq Sequencing platform was used to amplify and sequence the V3-V4 region of the 16S rRNA gene using barcode primer pairs (515f-GTGCCAGCMGCCGGTAA and 806r-GGACTACHVGGGTWTCTAAT). Sterile water was used as a control for each batch of DNA amplification. Upon completion of the analysis, the sequenced genetic data was archived at the Minnesota Supercomputing Institute.

mtDNA content. The content of mtDNA was determined using quantitative real-time PCR (iQ5 Multicolor Real-Time PCR Detection System; Bio-Rad) as previously described.[91] Briefly, three sets of primers were used in the RT-qPCR analysis: mtDNA Set 1 (Forward: CCCCACAAACCCCATTACTAAACCCA; Reverse: TTTCATCATGCGGAGATGTTGGATGG); mtDNA Set 2 (Forward: CAGTGAAATTGACCTGCCCGTGAA; Reverse: TCTTAGCATGTACTGCTCGGAGGT); β-globin (Forward: CTTGGGTTTCTGATAGGCAC; Reverse: CTTAGGGTTGCCCATAACAG).Total mtDNA copies were quantified by amplifying two regions of the mitochondrial genome, Cytochrome b (Cyt b) (222 bp) and the 16 S rRNA (197 bp) normalized to the invariable β-globin nuclear gene (147 bp). Results are expressed as mtDNA to nDNA ratio or mtDNA content.

Previously analyzed biomarkers

HPB-releasing DNA adducts. The HPB releasing DNA adducts were analyzed by previously developed LC-MS/MS method. [84] The isolated DNA samples were mixed with 0.12 pmol [3,3,4,4-D]HPB and 1000 pg [$^{13}C_2{}^{15}N$]guanine internal standards and subjected to acid hydrolysis at a final concentration of 0.8N HCl at 80 °C for 3 hrs. Samples were subsequently purified on activated HyperSep Hypercarb cartridges. HPB was eluted in 1ml 65% methanol in H_2O, dried, and stored at -20°C until analysis. The analysis of HPB in the purified samples was carried out on an LTQ Orbitrap Velos instrument (ThermoScientific) interfaced with a Nano2D-LC HPLC (Eksigent) with nanoelectrospray ionization, using accurate mass extracted ion chromatograms of *m/z* 106.0287 (parent ion *m/z* 166.1) for HPB and

corresponding fragment (*m/z* 110.0538, parent ion *m/z* 170.1) for [pyridine-D_4]HPB with a mass tolerance of 5 ppm.

Urinary total nicotine equivalents (TNE). TNE is a urinary biomarker of daily nicotine intake. It is expressed as a molar sum of nicotine and its known metabolites, accounting for approximately 73-96% of the nicotine dose.[67,68] Urinary TNE was analyzed as previously described in Chapter 2.

Total NNAL. Total NNAL was measured as previously described in Chapter 2.[34] Briefly, urine samples were mixed with [$^{13}C_6$]NNAL internal standard, treated with *β*-glucuronidase to release NNAL from their *N*- and *O*-glucuronides, and further purified using solid-phase extraction cartridges. The appropriate eluants were then analyzed by liquid chromatography-tandem mass-spectrometry (LC-MS/MS), monitoring transitions *m/z* 210 → 180 for NNAL, and *m/z* 216 → 186 for [$^{13}C_6$]NNAL. Unconjugated (free) NNAL was analyzed by the same method, but the urine samples were not treated with *β*-glucuronidase prior to their purification.

Nicotine metabolism ratio (NMR) in plasma. The NMR, a measure of CYP2A6 activity, is the ratio of 3′-hydroxycotinine to cotinine. These nicotine metabolites were quantified as previously described.[92] Briefly, plasma was mixed with deuterium-labeled internal standards, purified on Oasis MCX 96-well plates and analyzed by LC-MS/MS monitoring transitions for cotinine, [CD_3]cotinine, 3′-hydroxycotinine, and [CD_3]3′-hydroxycotinine.

Bioinformatics and statistical analyses

All quality control, trimming, and merging were completed using QIIME2's DADA2 plugin. Reads were trimmed for primers with cutadapt and filtered for low-quality reads (less than Q 1/4) with fastx_toolkit. For downstream analysis, only high-quality reads are considered. Bacterial taxonomy was assigned using a pre-built classifier on Greengenes 16SrRNA sequences.

R statistical interface was used for microbial community ecology analyses. R phyloseq[71] was used to analyze microbiome diversity characteristics: alpha diversity indices (Observed Richness and Shannon's indices), distance matrices (Bray-Curtis and UniFrac), ordination analyses (principal coordinates analysis (PCoA)). Adonis and anosim functions within the R vegan package[72] were used to calculate permutational multivariate analysis of variance (PERMANOVA). ASV abundances were used for functional prediction analyses using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt)[93] and predicted KEGG pathway relative abundances were used for further analysis. Continuous variables, the mtDNA content and HPB-released DNA adducts, were discretized into four distinct categories

based on specific cutoff values. This discretization process enabled to analyze the variations within biologically meaningful intervals. To assess differences in alpha diversity features between groups with different levels of DNA damage, Kruskal-Wallis and Wilcoxon rank-sum tests were used. The association between the oral microbiome and mtDNA content and HPB-releasing DNA adducts was tested in a multivariable regression model, adjusted for the potential confounders of age, sex, and smoking status. The differential abundances were analyzed using the DESeq2 package.[73] Lastly, Random Forest (RF) classification[94] was used to assess the classification accuracy of mt DNA content and HPB releasing adducts levels based on taxonomic features and predicted KEGG pathways. All graphs were plotted using the ggplots R package.[74]

Results

This study included 146 of the 162 participants who completed the parent study. Sixteen participants were excluded due to missing data (relevant biomarker and questionnaire data) or poor sequencing read quality. The summary of participants' characteristics and biomarker levels by race is presented in **Table 3-1.** The age of all participants ranged from 37 to 52, and there was even distribution by biological sex (71 male and 75 female) and race (75 AA and 71 WH). All study participants' median self-reported cigarettes smoked per day (CPD) were 15 (12-20). Consistent with the literature, AA participants smoked fewer CPD and had lower NMR when compared to WH ($p < 0.001$ for both). However, there was no difference in urinary biomarkers of tobacco smoke exposure, the median TNE was 54 (36.03-78.06) in AA and 54.36 (35.41-81.87) in WH (p=0.789). For statistical analysis purposes, study participants were categorized into four groups of mtDNA and HPB-releasing adduct levels according to the distribution among all study participants, ranges are presented in **Table 3-1**.

Oral microbiome profiles

A total of 13,078,884 16S rRNA sequence reads were obtained, 11,405,326 of which remained after quality filtering, reflecting 2477 unique amplicon sequence variant (ASVs), each representing a unique taxon. After further filtering (e.x., dropping taxa present in less than 3 samples), 595 unique ASVs were included in the analysis. The overall distribution of the median relative abundance of bacterial genera by biological sex and race is presented in **Figure 3-1**. The predominant genera *Actinomyces, Fusobacterium, Leptotrichia, Prevotella, Rothia, Fusobacterium, Streptococcus, Unclasified Gemellacea* and *Veillonella* were commonly found with similar representation regardless of sex and race.

Taxonomic composition and diversity

In this study, the relationship between microbial composition and oxidative stress and tobacco-carcinogen-derived damage was tested. There was no significant difference as assessed by Observed Richness and Shannon indexes across the groups with different levels of mtDNA content ($p>0.05$; **[Fig. 3-2A]**). However, significant difference in alpha-diversity metrics: Observed Richness and Shannon indexes (Kruskal-Wallis test, $H = 8.42$, $p <0.01$ and $H = 6.85$, $p <0.03$, respectively) was found across levels of HPB-releasing DNA adducts (very low, low, moderate, and high). FDR adjusted pairwise comparisons, showed that, for both indexes, individuals with high levels of HPB-releasing DNA adducts had significantly lower alpha diversity than those with low levels ($p <0.05$); **(Fig. 3-2B).** To further understand the variation of the microbiome across the groups with different levels of acquired DNA damage Bray–Curtis dissimilarity and UniFrac distances were computed. There were no differences in beta diversity across individuals with different levels of mtDNA content nor HPB-releasing DNA adducts as measured by weighted Bray-Curtis and weighted UniFrac ($p> 0.05$ for both metrics, **[Fig. 3-3]**). Specific quantitative metric for alpha diversity (Observed Richness and Shannon Index) were further used to assess the relationship between microbial composition and DNA damage. Significant association between Shannon Index and levels of HPB-releasing adducts was found across all three multivariate models **(Table.3-3)**. Observed Richness was also associated with levels of HPB-releasing adducts, however after adjusting for smoking (CPD, TNE, NNAL) did not reach statistical significance ($p=0.58$, **Table 3-3**).

Association of microbial composition with mtDNA content and HPB- releasing adducts

To address the skewed distribution of proportional data, a variance-stabilizing arcsin-square root transformation was applied to the relative abundance of each taxa. The arcsin-square root transformed relative abundance of each taxa and was then standardized by dividing by its standard deviation (SD). Based on the standardized values of each taxa, a summary score, Taxonomic Summary Score (TSS), was calculated, which represents the total exposure of the microbial communities in the sample. A multivariable regression model was used to relate the continuous variables (mtDNA content and HPB-releasing adducts, response variable) to TSS. Multivariable analysis results are summarized in **Table 3-2.** The regression model showed that the TSS had a non-significant effect on mtDNA levels ($\beta = 0.8473$, $p = 0.214$). However, there was a statistically significant negative association between TSS and HPB-releasing adducts ($\beta = -0.24$, $p = 0.03$), suggesting that for each unit increase in TSS there is an expected decrease of 0.24 in the value of HPB-releasing adducts, after accounting for the effects of smoking (TNE, NNAL and years of smoking).

Identifying discriminatory features of the oral microbiome by levels of DNA damage

To identify discriminatory features of the oral microbiome across individuals with different levels of DNA damage differential abundance analysis using the DESeq2 package was completed. Over a hundred taxa(agglomerated at genus-level ASV) were identified to be differentially abundant between those with very low and those with high levels of HPB-releasing adducts, the top 50 differentially abundant taxa are shown in Fig. 3-4 and the full list is included in Supplementary Table 3-2.

To further identify discriminatory features across the level groups for HPB-releasing adducts machine learning classification approach Random Forest (RF) classification model (using mean decrease accuracy index) was used. For the RF model, the dataset was split 70% for training and 30% retained for testing, the model was cross-validated 10-fold and repeated 5 times. The results showed that based on the taxonomic features the RF model was most efficient in classifying individuals with high levels of HPB-releasing adducts (AUCs of 0.69) [Fig. 3-5A]). In addition, the results of functional prediction analysis by PICRUSt were used and applied a similar RF model to uncover predicted discriminatory pathways in the oral microbiome of individuals with different levels of DNA damage, the model performed poorly when classifying individuals with different levels of sustained DNA damage (with an AUC of 0.64 [Fig. 3-5B]). Intriguingly predicted pathways that were discriminatory for very low and low levels of HPB-releasing DNA adducts were involved in biosynthesis of vital cellular components such as peptidoglycans, amino acids and nucleotides whereas predicted pathways associated with high levels of HPB-releasing adducts were involved in degradation, assimilation and utilization processes. RF model performed poorly when classifying individuals with different levels of mtDNA content (Supplemental Figure 3-1).

Discussion

This study aimed to better characterize the relationship between the oral microbiome and smoking induced DNA damage (mtDNA content and HPB-releasing DNA adducts), and to potentially provide insights into the potential contribution of microbial communities in tobacco-induced DNA damage. In this cohort of 146 healthy individuals who smoke, it was found that many of their oral microbiomes comprised eight genera (*Actinomyces, Fusobacterium, Leptotrichia, Prevotella, Rothia, Streptococcus, Unclassified Gemellaceae* and *Veillonella*) belonging to three distinct phyla (Actinobacteria, Fusobacteria, Bacteroidetes [Fig.3-1]). A total of 128 ASVs were identified in 99% of all samples contributing to 46.5% of the total oral microbiome composition indicating that individuals tend to share similar genera within the oral cavity. These findings are in line with previous work, the Human Oral Microbiome Project showed that the oral

microbiome composition is dominated by *Streptococcus*, followed in abundance by *Haemophilus*, *Actinomyces* and *Prevotella*. [94]

Although it is biologically plausible that mitochondrial–microbiome functional interaction will contribute to the mtDNA content, this study found no significant relationship between the oral microbiome and mtDNA content among individuals who smoke. It has been established that smoking impacts mitochondrial function and structure through a variety of mechanisms, including the formation of mtDNA adducts, increased mtDNA damage, and changes in mtDNA content. Factors that affect the mtDNA content may also affect the oral microbiome, impacting its relations. For example, a variety of harmful chemicals and toxicants are introduced into the oral cavity by cigarette smoking, as well as a specific microenvironment is created by increased temperature, decreased oxygen levels, and changes in pH[15], which can adversely affect the growth and survival of specific microbial species while promoting the growth of others. Smoking also has systemic effects on the body beyond the oral cavity. It affects various physiological processes and can lead to systemic inflammation, oxidative stress, and impaired immune function. These systemic effects can indirectly influence the oral microbiome by affecting the host's immune response and inflammatory status. Beyond the immunity-inducing mechanism, there are some other metabolites of the microbiome that have been reported to affect mitochondrial function. Bacterially produced propionate has been shown to induce mitochondrial biogenesis and play a role in fatty acid metabolism. These fatty acids are reported to increase mitochondrial mass, mtDNA copy number, and mitochondrial transcription factor activity.[95] Finally, mtDNA content and the oral microbiome can be affected by other lifestyle and behavioral factors associated with cigarette smoking, those who smoke, for example, may have different alcohol consumption patterns, eating habits and oral hygiene practices. Although the oral microbiome can potentially contribute to the levels of mtDNA content, direct cigarette smoke exposure, altered oral environment, systemic effects of smoking, and other lifestyle factors associated with smoking may mask the specific influence of the oral microbiome on mtDNA content among individuals who smoke. While a relationship between the oral microbiome and mtDNA content was not detected in this particular study, potentially due to a small sample size, these relationships should be further explored in future studies. Moreover, the interaction of the three genomes (bacterial, mitochondrial, and nuclear) remains a relatively unexplored area of study and might provide meaningful insights into the individual response to toxicants and carcinogens from tobacco smoke.

Alpha diversity measures differed significantly across different levels of DNA damage caused by tobacco-specific constituents **(Fig.3-2B).** Individuals with higher levels of HPB-releasing DNA adducts exhibited significantly lower microbial diversity than those with lower levels of adducts. These findings are in the same lines with prior study, which reported that increase in HPB-releasing DNA adduct levels in oral cells was negatively associated with the number of observed bacterial taxa on the same mucosal site in individuals who smoke and had HNSCC.[38] Similar trend was observed in healthy controls; the cumulative abundance of taxonomic signatures characterizing healthy controls tended to decrease with increasing DNA adduct levels. It is likely that the oral microbiome may contribute to variations in tobacco-specific derived DNA damage. The fact that urinary biomarkers for NNN and NNK exposure, which are causally associated with tobacco-specific derived DNA damage, do not consistently correlate with HPB-releasing DNA adduct levels provides further support for the potential contribution of the oral microbiome to variations in HPB-releasing DNA adducts.

To assess the specific associations between microbial composition and HPB-releasing adducts a multivariate analysis controlling for smoking-related confounding was performed. These results revealed a statistically significant negative association between TSS and HPB-releasing adducts. For each unit increase in TSS, HPB-releasing adducts were expected to decrease by 0.24 units **(Table 3-2).** Higher TSS scores, reflecting greater taxonomic diversity within the oral microbiome, were associated with lower levels of HPB-releasing adducts further supporting the notion that there may be increased DNA damage due to a decrease in microbial diversity within the oral cavity. Quantitative metrics for alpha diversity such as Observed Richness and Shannon Index were used in a multivariate model and found that Shannon Index was significant predictor for HPB-releasing adduct. A unit increase in Shannon Index corresponds to approximately 18 units of decrease in HPB releasing adducts, a similar trend as with the TSS. It appears that there is complex interaction between the smoking-induced oral microenvironment and DNA damage, further studies need to explore the molecular mechanisms that that might be protective against smoke carcinogens.

Beta diversity using Bray-Curtis dissimilarity and UniFrac distances was analyzed to further explore variation in microbiome composition. Interestingly, differences in beta diversity between individuals with different levels of HPB-releasing DNA adducts were not observed **(Fig.3-3A).** All individuals may share a core set of microbial species in their oral microbiome, regardless of the extent of DNA damage.[94,96] Consequently, even though within participant diversity decreases with increasing DNA damage, the core

species could maintain similar patterns across participants, resulting in comparable beta diversity. It is also plausible that certain species may promote DNA damage while others may have protective effects, leading to a dynamic balance.

There were significant differences in taxonomic abundance between individuals with very low and high levels of HPB-releasing adducts. Using these two contrasting levels, very low and high DNA adduct levels, were used to examine differences in abundance of taxa across the oral microbiome and elucidate likely correlations between microbial composition and DNA adduct formation. Taxa belonging to *Fusobacterium, Desulfobulbus, Bacillus, Actinomyces, Burkholderia, Schwartzia, Eikenella, Prevotella, Corynebacterium, Lachnoanaerobaculum, Parvimonas, Lachnoanaerobaculum, Buchnera, Shuttleworthia, Johnsonella, Treponema, Leptotrichia, Atopobium* were increased in those with high levels of HPB-releasing adducts while the abundance of other taxa were decreased (Supplemental Table 3-1). Some of these taxa (*Fusobacterium, Prevotella, Corynebacterium, Leptotrichia, Actinomyces)* have been implicated in nitrate metabolism. These taxa could potentially contribute to endogenous formation of NNN from nornicotine, a tobacco constituent and a nicotine metabolite, which is a potential mechanism by which microbial composition may contribute to formation of HPB-releasing adducts. Machine learning classification models (RF) identified discriminatory features most efficiently (AUCs of 0.69, [Fig.3-6A]), further reinforces the evidence for distinctive microbial composition based of the levels HPB-releasing adducts. Machine learning classification models (RF) could be useful for identifying individuals at high risk of developing tobacco-related cancers.

Apart from altering species profiles, tobacco exposures could also trigger the expression of different microbial genes, contributing to the microbially-mediated metabolism of tobacco chemicals. In previous study increased abundance of pathways involved in degradation of chemicals (e.g., toluene, phenyl compounds) and amines (e.g., aromatic biogenic amines) was observed in those with HNSCC, in contrast healthy controls exhibited greater abundance of predicted carbohydrate metabolism pathways.[38] Among participants in the current study, predicted pathways involved in carbohydrate metabolism, such as sucrose degradation III (sucrose invertase) and pentose phosphate pathways, pathways involved in the synthesis of various compounds, such as fatty acids, vitamins, modified nucleosides, coenzymes, and other molecules that are essential for cell function and contribute to a healthy microbiome, were conserved in individuals who have sustained high DNA damage. However, these individuals also had higher predicted

abundances of pathways involved in degrading aromatic compounds, amino acids, and nucleosides (Fig. 3-6).

There are several limitations to this study. The temporal dynamics of the oral microbiome and its relationship with DNA damage were not explored. A better understanding of the oral microbiome and its association with DNA damage over time might provide important insights into the causal role of the oral microbiome in tobacco-induced DNA damage. Longitudinal studies also offer an opportunity to collect detailed information of other factors that could have an impact on the oral microbiome such as dietary habits, alcohol consumption, and oral hygiene practices. In addition, the functional analysis of microbial pathways involved in tobacco chemical metabolism relied on predicted abundances rather than functional metagenomic analysis. While predictive analysis provides insights into potential metabolic pathways, future studies should focus on the actual function analysis to provide a more robust understanding of microbial functions and their role in tobacco chemical metabolism. Another potential limitation is the absence of non-tobacco users in the study cohort. However, primary objective of this study was to investigate the relationship between oral microbial composition and the tobacco-induced macromolecular damage.

In summary, the research in this chapter provided insights into the relationship between oral microbiome and smoking-induced DNA damage, as well as potential implications for tobacco-related cancers. These results suggest that tobacco-induced DNA damage may be influenced by the composition and function of the oral microbiome. In addition, specific taxa are associated with high levels of DNA adducts, many of which contribute to nitrate metabolism and subsequently to endogenous nitrosamine formation. This study also emphasizes the importance of considering the complex interaction between the smoking-induced oral microenvironment, DNA damage, and microbial diversity when conducting studies aiming to understand interindividual variation in tobacco related outcomes. Ultimately, by identifying specific composites of taxa associated with high DNA damage we might be able to identify individuals at high risk of developing tobacco-related cancers as well as inform targeted interventions to mitigate the adverse effects of smoking on the oral microbiome and DNA damage.

TABLES AND FIGURES FOR CHAPTER 3

Table 3-1. Study participant characteristics and biomarkers – summary statistics

	Total (N=146)	Female (N=75)	Male (N=71)	p-value
Age (years)	45.5 (37-52)	44 (37-52)	47 (39-52)	0.55
Self-reported race (African American)	75 (51.4%)	36 (48.0%)	39 (54.9%)	0.40
Cigarettes per day (CPD)	15 (11-20)	15 (12-20)	15 (12-20)	0.36
Nicotine Metabolite Ratio (NMR)	0.29 (0.02-0.42)	0.31 (0.23-0.46)	0.28 (0.19-0.38)	0.07
TNE (pmol/mL urine)	53.11 (35.91-81.11)	55.82 (38.65-83.11)	46.61(31.60-77.49)	0.21
Total NNAL (pmol/mL)	1.33 (0.72-1.99)	1.45 (0.92-1.92)	1.22 (0.67-2.16)	0.45
Alcohol use (drinks per day)	3 (2-3)	3 (2-3)	3 (2-4)	0.07
mtDNA content (nDNA/mtDNA)	119.43 (35.33-351.53)	123.64 (28.57-370.08)	118.6 (41.12-330.27)	0.96
Levels of mtDNA content				0.74
Very Low <35	21 (28.0%)	16 (22.5%)	37 (25.3%)	
Low 36 – 119	16 (21.3%)	20 (28.2%)	36 (24.7%)	
Moderate 120 – 350	18 (24.0%)	18 (25.4%)	36 (24.7%)	
High >350	20 (26.7%)	17 (23.9%)	37 (25.3%)	
HPB (pmol/mg DNA)	0.88 (0.01-10.84)	1.44 (0.02-5.63)	1.21 (0.01-6.48)	0.94
Levels of HPB				0.23
Very Low <0.036	43 (29.5%)	24 (32.0%)	19 (26.8%)	
Low 0.037-1.5	35 (24.0%)	17 (22.7%)	18 (25.4%)	
Moderate 1.51-10	38 (26.0%)	15 (20.0%)	23 (32.4%)	
High >10	30 (20.5%)	19 (25.3%)	11 (15.5%)	

*Summaries shown are median (1st quartile, 3rd quartile) or N (percent)

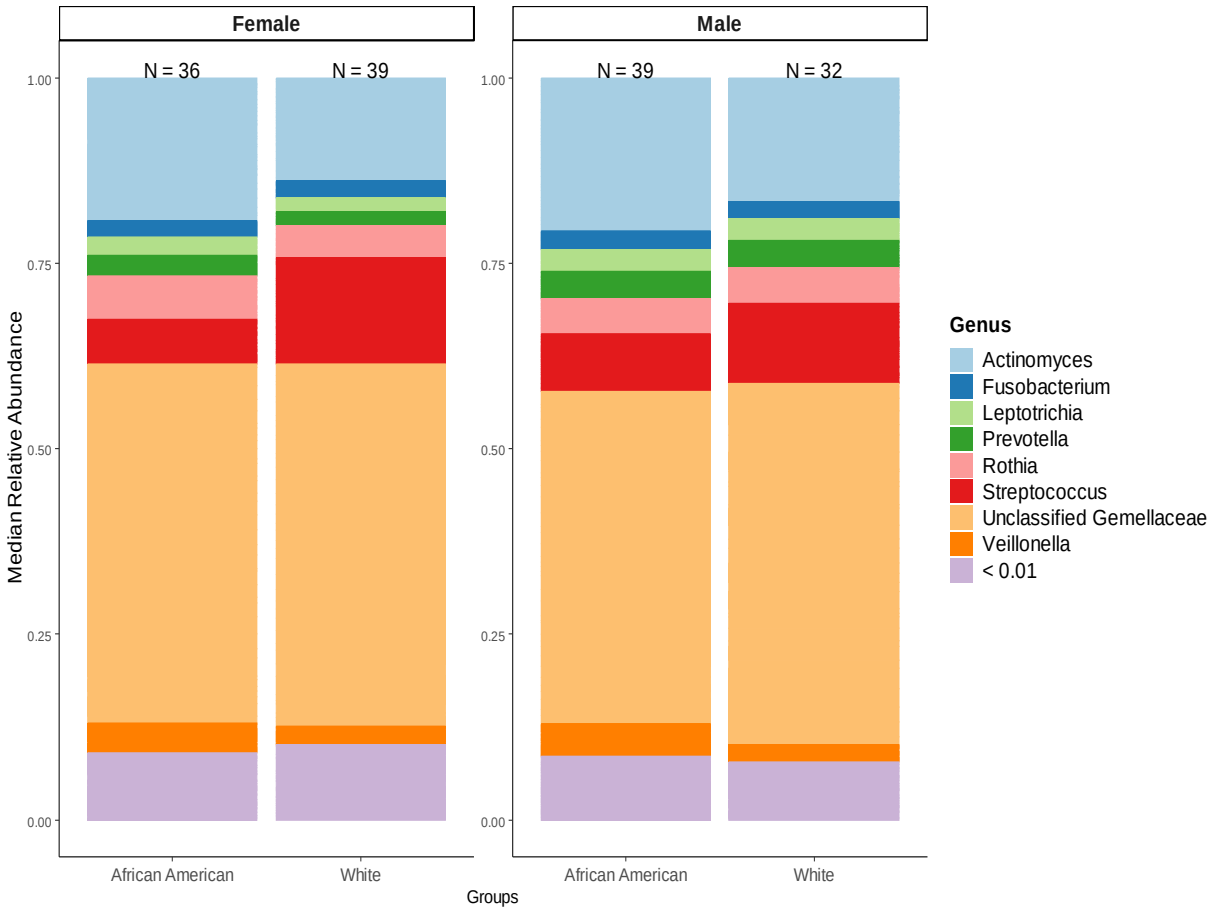


Figure 3-1. Median relative abundance of bacterial genera by biological sex and race

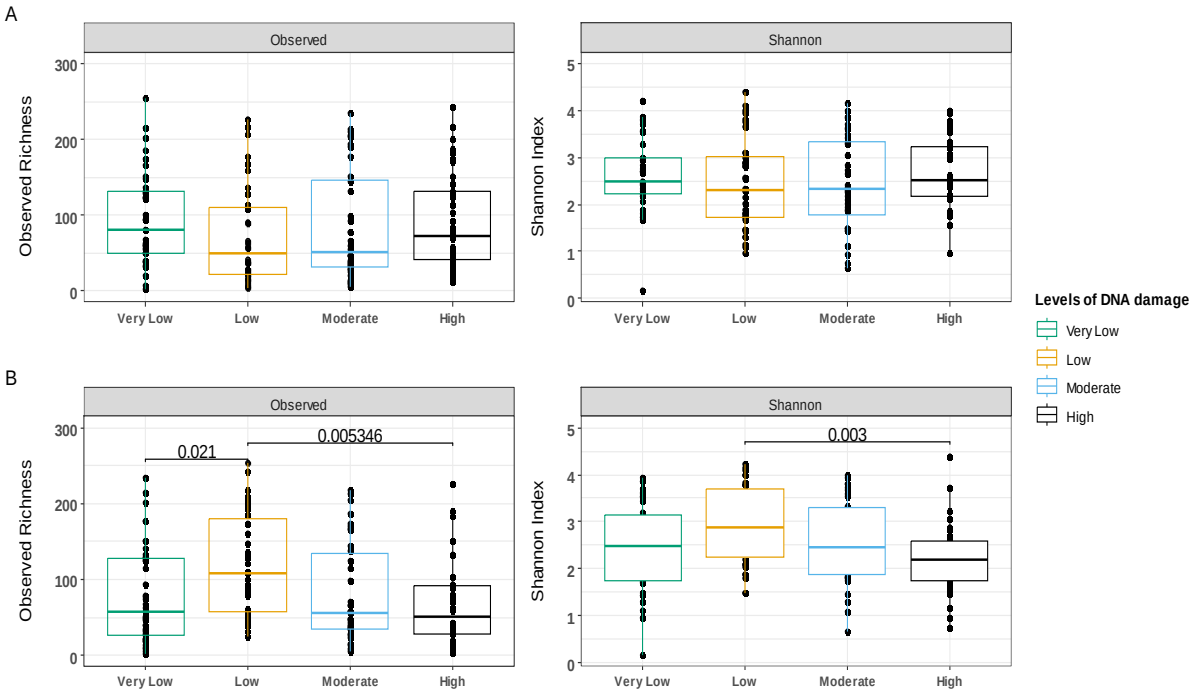


Figure 3-2. Alpha diversity measures (Observed Richness and Shannon Index). The boxplots show each alpha diversity measure by groups of: A) mtDNA content B) HPB-releasing DNA adduct levels

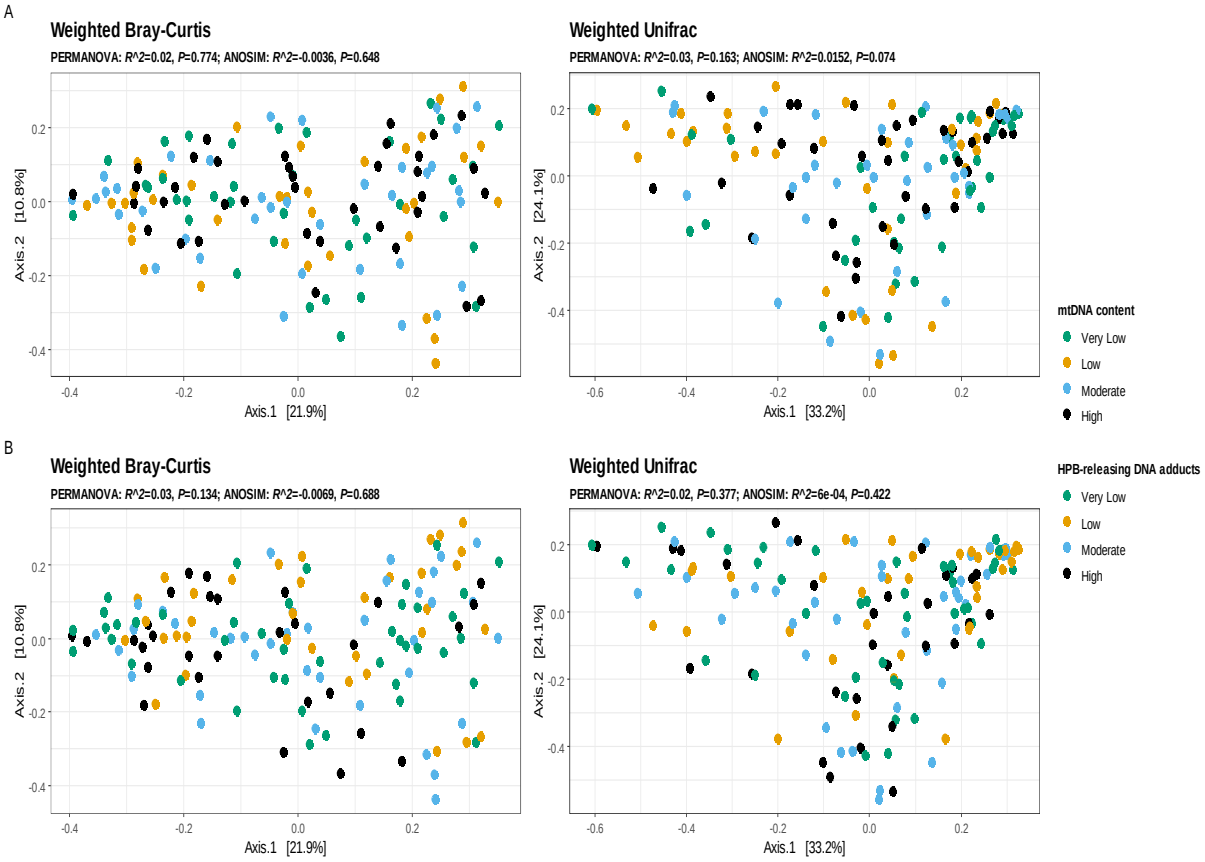


Figure 3-3. Beta diversity based on Weighted Bray-Curtis and Unifrac distances by A) mtDNA content; B) HPB-releasing DNA adduct levels. The scatter plots do not reveal clustering patterns of the microbial composition of samples by levels of DNA damage

Table 3-2. Association between TSS and DNA damage (mtDNA content and HPB-releasing adducts): summary of multivariable analyses

Model	mtDNA content			HPB-releasing DNA adducts		
	Coefficient (β)	95% CI	p-value	Coefficient (β)	95% CI	p-value
Model 1 (Crude)	0.85	0.62-8/0	0.2	-0.24	0.65-0.94	0.01
Model 2 (Adjusted for age, sex, race)	0.28	0.35-5.1	0.7	-0.23	0.65-0.96	0.02
Model 3 (model 2+ TNE, NNAL, years of smoking)	0.09	0.31-4.1	0.9	-0.24	0.65-0.98	0.03

Table 3-3. Assosiation between alpha diversity metrics and HPB-releasing DNA adducts: summary of multivariable analyses

	HPB-releasing DNA adducts								
Alpha	Model 1			Model 2			Model 3		
Diversity	Coefficient (β)	95% CI	p- value	Coefficient (β)	95% CI	p- value	Coefficient (β)	95% CI	p- value
Observed Richness	-0.175	0.71-0.98	**<0.001**	-0.165	0.171-1.00	**0.05**	-0.1739	0.70-1.00	0.058
Shannon Index	-17.381	0.0001-0.004	**0.005**	-16.764	0.0002-0.01	**0.008**	-17.5224	0.0005-0.01	**0.009**

Model 1: crude; Model 2: adjusted for age, sex, race; Model 3: Model 2 + smoking (TNE, NNAL, years of smoking)

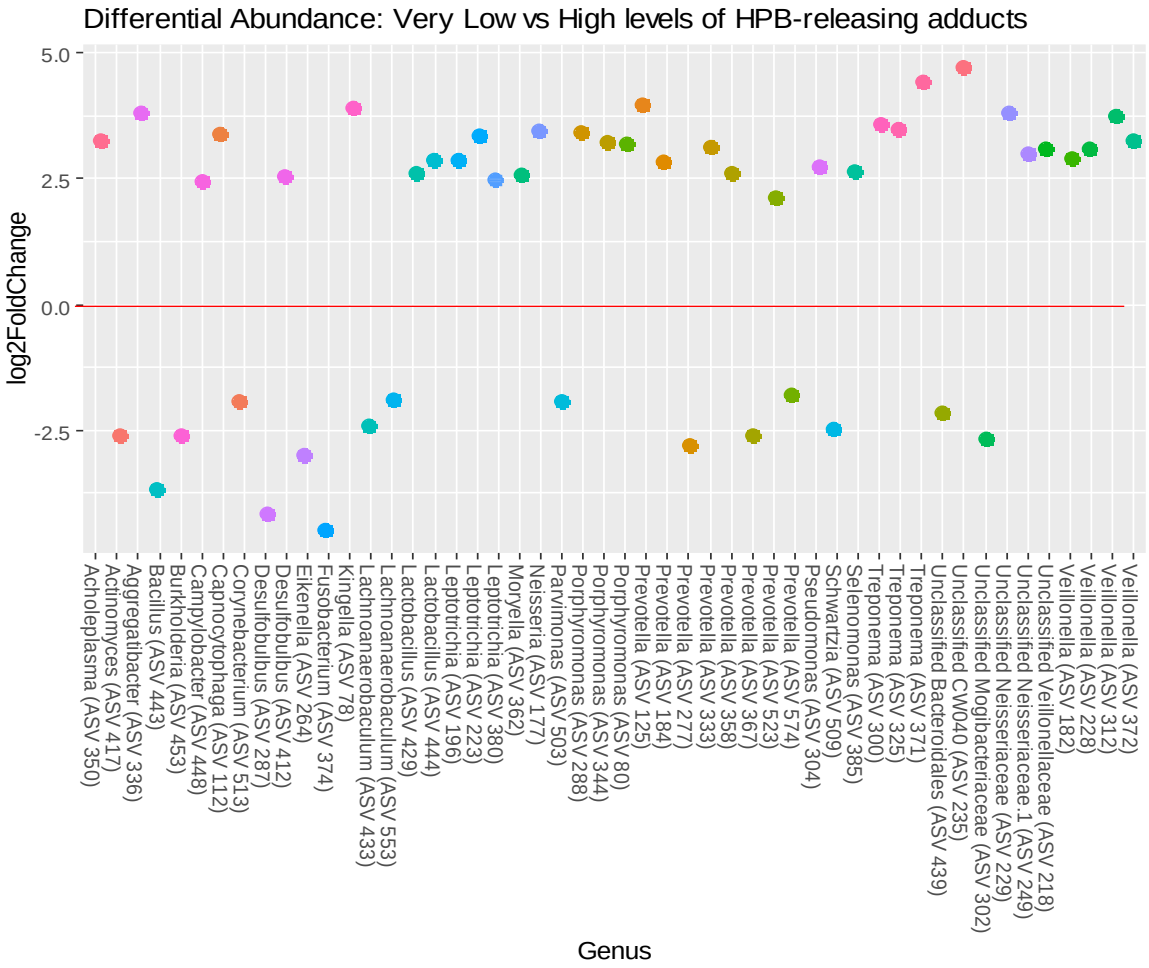
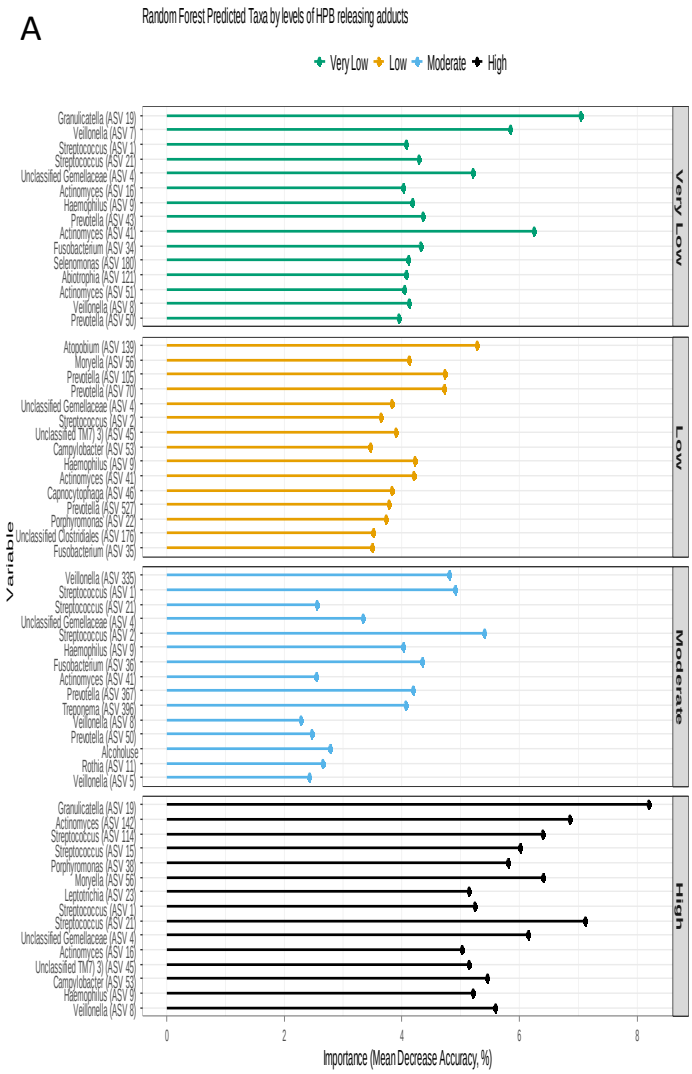
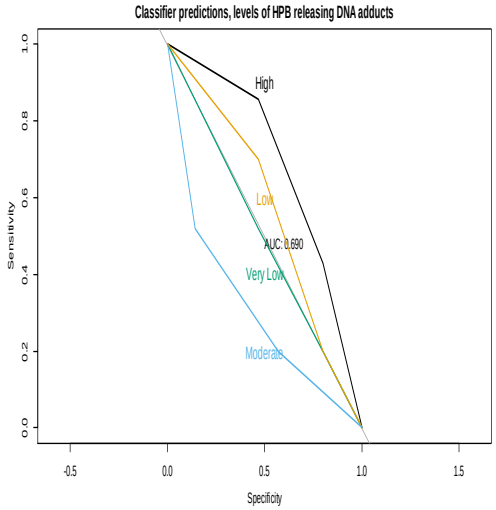
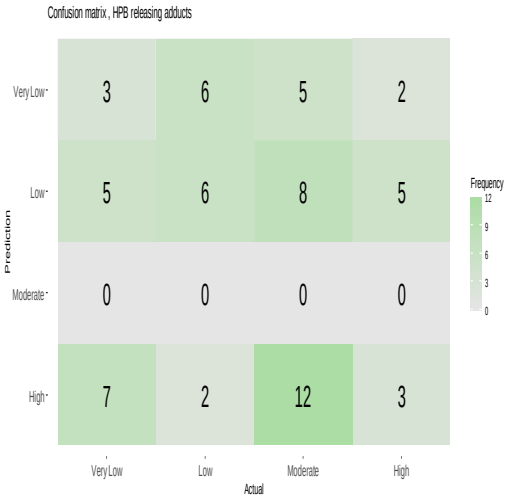


Figure 3-4. Differential abundance (genus -level ASV agglomeration) between the Very Low and High levels of HPB-releasing DNA adducts. Each data point represents a genus-level ASVs (x-axis) identified as significantly different along with the log2 fold change (y-axis); positive log2FoldChange indicates an increase in abundance, while a negative log2FoldChange indicates a decrease in abundance in those with high levels of HPB-releasing adducts compared to those with very low levels;

A


Random Forest Predicted Taxa by levels of HPB releasing adducts
Very Low
Low
Moderate
High
Granulicatella (ASV 19)
Veillonella (ASV 7)
Streptococcus (ASV 1)
Streptococcus (ASV 21)
Unclassified Gemellaceae (ASV 4)
Actinomyces (ASV 16)
Haemophilus (ASV 9)
Prevotella (ASV 43)
Actinomyces (ASV 41)
Fusobacterium (ASV 34)
Selenomonas (ASV 180)
Abiotrophia (ASV 121)
Actinomyces (ASV 51)
Veillonella (ASV 8)
Prevotella (ASV 50)
Atopobium (ASV 139)
Moryella (ASV 56)
Prevotella (ASV 105)
Prevotella (ASV 70)
Unclassified Gemellaceae (ASV 4)
Streptococcus (ASV 2)
Unclassified TM7 3) (ASV 45)
Campylobacter (ASV 53)
Haemophilus (ASV 9)
Actinomyces (ASV 41)
Capnocytophaga (ASV 46)
Prevotella (ASV 527)
Porphyromonas (ASV 22)
Unclassified Clostridiales (ASV 176)
Fusobacterium (ASV 35)
Veillonella (ASV 335)
Streptococcus (ASV 1)
Streptococcus (ASV 21)
Unclassified Gemellaceae (ASV 4)
Streptococcus (ASV 2)
Haemophilus (ASV 9)
Fusobacterium (ASV 36)
Actinomyces (ASV 41)
Prevotella (ASV 367)
Treponema (ASV 398)
Veillonella (ASV 8)
Prevotella (ASV 50)
Alcaniivorax
Rothia (ASV 11)
Veillonella (ASV 5)
Granulicatella (ASV 19)
Actinomyces (ASV 142)
Streptococcus (ASV 114)
Streptococcus (ASV 15)
Porphyromonas (ASV 38)
Moryella (ASV 56)
Leptotrichia (ASV 23)
Streptococcus (ASV 1)
Streptococcus (ASV 21)
Unclassified Gemellaceae (ASV 4)
Actinomyces (ASV 16)
Unclassified TM7) 3) (ASV 45)
Campylobacter (ASV 53)
Haemophilus (ASV 9)
Veillonella (ASV 8)
Variable
0
2
4
6
8
Importance (Mean Decrease Accuracy, %)


Classifier predictions, levels of HPB releasing DNA adducts
High
Low
AUC: 0.690
Very Low
Moderate
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
-0.5
0.0
0.5
1.0
1.5
Specificity


Confusion matrix , HPB releasing adducts
Very Low
Low
Moderate
High
Prediction
3
6
5
2
5
6
8
5
0
0
0
0
7
2
12
3
Very Low
Low
Moderate
High
Actual
Frequency
12
9
6
3
0

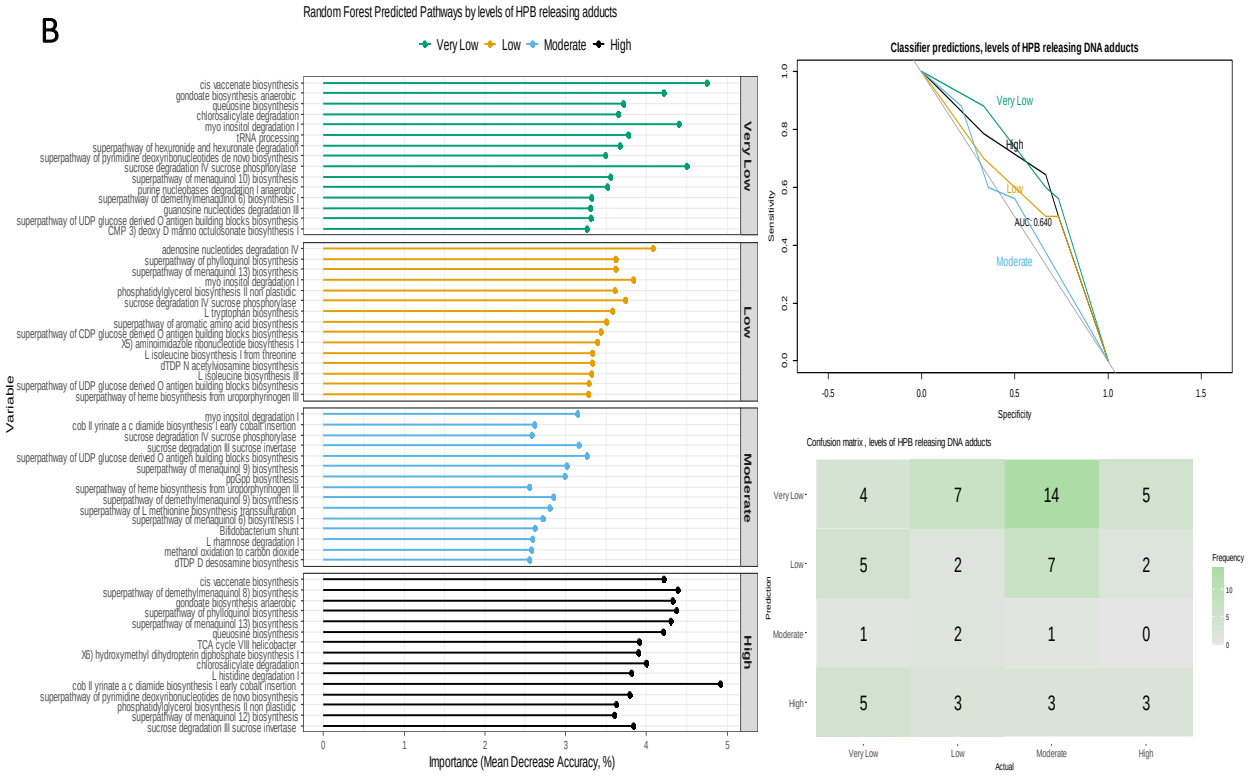
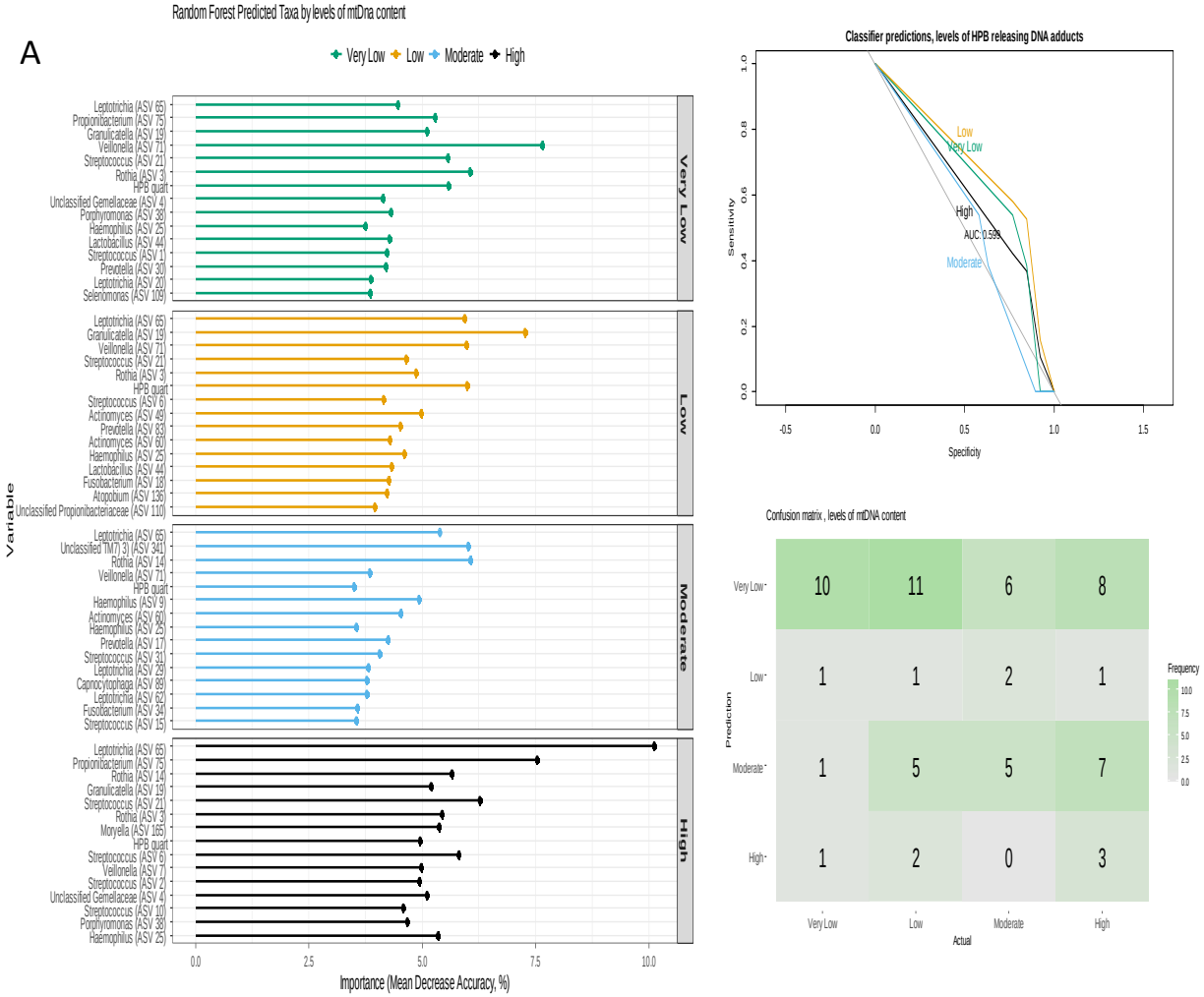
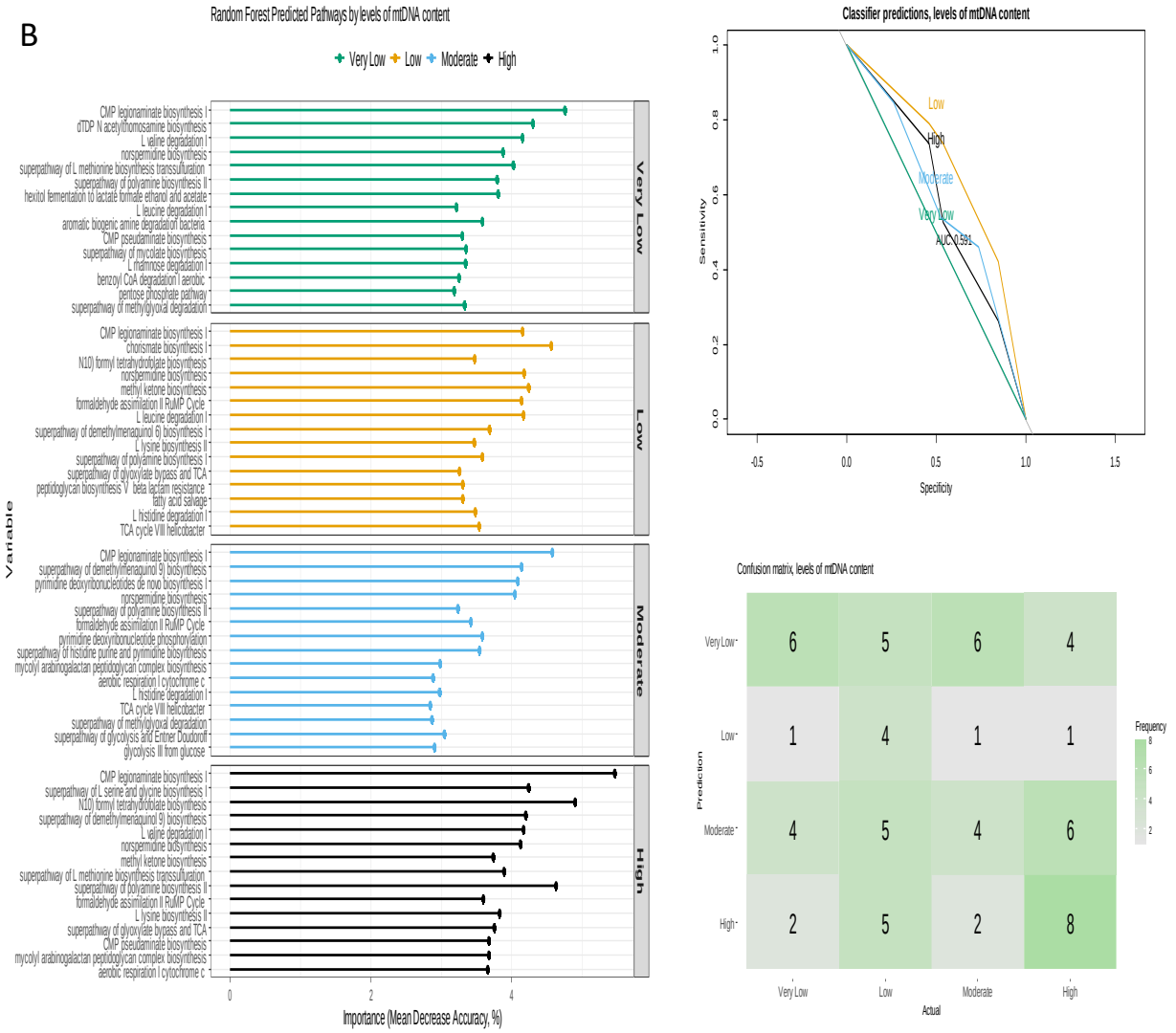


Figure 3-5. Performance of RF model classification used for identification of discriminant A) Taxa and B) pathways among individuals with different levels of HPB-releasing DNA adducts

SUPPLEMENTAL DATA FOR CHAPTER 3





Supplemental Figure 3-1. Performance of RF model classification used for identification of A) discriminant taxa and B) pathways among individuals with different levels of mtDNA content

Supplemental Table 3-1. Differential abundance (ASV level agglomeration) between study participants with Very Low and High levels of HPB-releasing DNA adducts

ASV level	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
(ASV 235) TM7.3	18.4952	-4.703229023	0.727557104	-6.464412207	1.02E-10	6.84E-09
(ASV 371) Treponema	9.366853	-4.432534663	0.572782389	-7.738601516	1.01E-14	1.18E-12
(ASV 125) Prevotella	49.02636	-3.958188789	0.682989468	-5.795387739	6.82E-09	3.57E-07
(ASV 78) Kingella	115.901	-3.896988095	0.680722581	-5.724781584	1.04E-08	3.58E-07
(ASV 336) Aggregatibacter	10.6826	-3.814531473	0.572728775	-6.660275584	2.73E-11	2.15E-09
(ASV 229) Neisseriaceae	19.41303	-3.791938845	0.554343545	-6.840413092	7.90E-12	7.44E-10
(ASV 312) Veillonella	12.53258	-3.743747185	0.613888749	-6.098413093	1.07E-09	6.31E-08
(ASV 300) Treponema	12.56191	-3.568365161	0.640648089	-5.569930237	2.55E-08	8.00E-07
(ASV 325) Treponema	11.70439	-3.474699738	0.605401856	-5.739493036	9.50E-09	3.58E-07
(ASV 177) Neisseria	29.64625	-3.464708991	0.662002836	-5.233676958	1.66E-07	3.73E-06
(ASV 288) Porphyromonas	13.79823	-3.410281737	0.593739274	-5.743736162	9.26E-09	3.58E-07
(ASV 112) Capnocytophaga	58.08573	-3.370785415	0.680397284	-4.954142958	7.26E-07	1.22E-05
(ASV 223) Leptotrichia	20.10838	-3.352279631	0.658648871	-5.089630877	3.59E-07	7.04E-06
(ASV 350) Acholeplasma	10.62495	-3.262184091	0.56565734	-5.767067555	8.07E-09	3.58E-07
(ASV 372) Veillonella	9.334654	-3.24877126	0.567933967	-5.720332735	1.06E-08	3.58E-07
(ASV 344) Porphyromonas	10.76451	-3.208300957	0.594198477	-5.399375936	6.69E-08	1.75E-06
(ASV 80) Porphyromonas	110.6247	-3.178676382	0.718641878	-4.423171652	9.73E-06	9.54E-05
(ASV 333) Prevotella	10.94087	-3.135647378	0.578107486	-5.423986805	5.83E-08	1.61E-06
(ASV 218) Veillonellaceae	21.27647	-3.100930655	0.595816163	-5.20450912	1.95E-07	4.16E-06
(ASV 228) Veillonella	20.69654	-3.081560557	0.693003045	-4.446676791	8.72E-06	8.80E-05
(ASV 249) Neisseriaceae	16.7207	-2.994904857	0.594089568	-5.041167219	4.63E-07	8.72E-06
(ASV 182) Veillonella	29.95314	-2.913214585	0.589989496	-4.937739747	7.90E-07	1.28E-05
(ASV 90) Streptococcus	98.21769	-2.891808652	0.681366266	-4.24413241	2.19E-05	0.0001879
(ASV 444) Lactobacillus	5.953893	-2.875802761	0.553603816	-5.194694615	2.05E-07	4.20E-06
(ASV 196) Leptotrichia	26.54532	-2.861949421	0.635089802	-4.506369669	6.59E-06	6.93E-05
(ASV 184) Prevotella	28.26556	-2.829005838	0.648122017	-4.364927845	1.27E-05	0.0001198
(ASV 152) Prevotella	38.02054	-2.812766434	0.674908168	-4.167628376	3.08E-05	0.0002457
(ASV 304) Pseudomonas	13.80278	-2.73054774	0.592607187	-4.607685833	4.07E-06	4.92E-05
(ASV 94) Campylobacter	89.43534	-2.687153368	0.637306366	-4.216423234	2.48E-05	0.0002088
(ASV 17) Prevotella	586.7441	-2.664039189	0.711772625	-3.742823331	0.000181964	0.0010713
(ASV 113) Rothia	59.60073	-2.659082818	0.704207429	-3.775993706	0.000159371	0.0009753
(ASV 385) Selenomonas	7.880615	-2.624521868	0.559215727	-4.693218989	2.69E-06	3.62E-05
(ASV 139) Atopobium	42.41943	-2.612883438	0.667508174	-3.914384182	9.06E-05	0.0006098
(ASV 358) Prevotella	9.389559	-2.609119475	0.561579428	-4.6460382	3.38E-06	4.19E-05
(ASV 335) Veillonella	11.80566	-2.604702701	0.624726317	-4.169350048	3.05E-05	0.0002457
(ASV 126) Prevotella	48.42676	-2.60440487	0.640625529	-4.06540912	4.79E-05	0.0003643

(ASV 429) Lactobacillus	6.015045	-2.593885158	0.556838461	-4.658236344	3.19E-06	4.06E-05
(ASV 362) Moryella	9.182389	-2.583906617	0.549049013	-4.706149276	2.52E-06	3.50E-05
(ASV 88) Leptotrichia	93.54326	-2.570226865	0.711086028	-3.614509025	0.000300917	0.0016502
(ASV 211) Selenomonas	21.60042	-2.560697102	0.591365238	-4.330144782	1.49E-05	0.000135
(ASV 412) Desulfobulbus	6.358227	-2.553160927	0.537290089	-4.75192262	2.01E-06	2.88E-05
(ASV 380) Leptotrichia	8.396538	-2.479283056	0.557736803	-4.445256337	8.78E-06	8.80E-05
(ASV 271) Eikenella	15.24427	-2.457169881	0.572217719	-4.29411708	1.75E-05	0.000153
(ASV 120) Prevotella	53.98056	-2.456505599	0.679687583	-3.614168714	0.000301313	0.0016502
(ASV 448) Campylobacter	5.45524	-2.450992752	0.492454295	-4.977096911	6.45E-07	1.13E-05
(ASV 267) Capnocytophaga	15.33505	-2.440402832	0.603244441	-4.045462612	5.22E-05	0.0003904
(ASV 422) Atopobium	6.059491	-2.383637905	0.570393812	-4.178933669	2.93E-05	0.000242
(ASV 185) Selenomonas	28.09163	-2.37770338	0.630418972	-3.771624085	0.000162188	0.0009794
(ASV 192) Capnocytophaga	26.65261	-2.369554556	0.633060313	-3.74301549	0.000181825	0.0010713
(ASV 381) Fusobacterium	7.959828	-2.367576196	0.543925621	-4.352757257	1.34E-05	0.0001242
(ASV 281) Moryella	14.81686	-2.33676988	0.580231917	-4.027303242	5.64E-05	0.0004152
Lachnoanaerobaculum	14.86483	-2.330876851	0.588239498	-3.962462327	7.42E-05	0.0005369
(ASV 331) Porphyromonas	11.29968	-2.317915865	0.585468002	-3.959082062	7.52E-05	0.0005369
(ASV 140) Prevotella	42.25128	-2.264183755	0.722364514	-3.134406121	0.001722023	0.0077245
(ASV 270) Capnocytophaga	14.63638	-2.262337743	0.591094307	-3.827371907	0.000129519	0.0008134
(ASV 161) Prevotella	33.38869	-2.258648895	0.613457152	-3.681836441	0.00023156	0.0013301
(ASV 320) Prevotella	11.53094	-2.238310982	0.575838751	-3.88704473	0.000101472	0.0006638
(ASV 339) Veillonella	11.55797	-2.180413461	0.553078869	-3.942319229	8.07E-05	0.0005673
(ASV 222) Veillonellaceae	20.46064	-2.176140152	0.617113227	-3.52632233	0.000421374	0.0022052
(ASV 435) Prevotella	5.800666	-2.131096963	0.543361998	-3.922057432	8.78E-05	0.0006081
Lachnoanaerobaculum	66.51124	-2.129094986	0.663585807	-3.208469747	0.001334433	0.0062852
(ASV 523) Prevotella	3.163331	-2.129014344	0.465561219	-4.57300621	4.81E-06	5.66E-05
(ASV 109) Selenomonas	61.9543	-2.124400803	0.697899209	-3.043993711	0.0023346	0.0097309
(ASV 239) Prevotella	17.81267	-2.068735846	0.643089869	-3.216868975	0.001295978	0.006282
(ASV 148) Fusobacterium	39.29945	-2.057876104	0.671506456	-3.064566371	0.002179858	0.0092435
(ASV 421) Prevotella	5.910739	-1.984039986	0.525452452	-3.775869687	0.00015945	0.0009753
(ASV 257) Eikenella	16.05499	-1.907094935	0.607809151	-3.137654197	0.001703057	0.0077129
(ASV 410) Fusobacterium	6.769996	-1.873019886	0.503758161	-3.718093386	0.000200732	0.0011672
(ASV 507) Catonella	3.648056	-1.843811549	0.472664103	-3.900891853	9.58E-05	0.0006358
(ASV 296) Prevotella	13.48483	-1.834775608	0.576353733	-3.18341932	0.001455467	0.0067208
(ASV 342) Prevotella	10.33081	-1.829670674	0.591917698	-3.091089655	0.001994234	0.0088088
(ASV 457) Prevotella	5.125522	-1.805792604	0.536658621	-3.364881387	0.000765766	0.0038782
(ASV 373) TM7.3	8.605892	-1.787890868	0.581827386	-3.072888821	0.002119975	0.0090864
(ASV 297) Schwartzia	13.01976	-1.769749017	0.573006845	-3.088530327	0.002011492	0.0088088
(ASV 445) Capnocytophaga	5.397313	-1.722324144	0.470536432	-3.660341746	0.000251879	0.0014293
(ASV 497) Neisseriaceae	3.989261	-1.706012669	0.474717519	-3.59374281	0.000325962	0.0017647

(ASV 589) Shuttleworthia	1.791147	-1.145161628	0.35658084	-3.211506336	0.001320411	**0.006282**
(ASV 496) Treponema	3.847143	1.475300596	0.45899153	3.214221831	0.001307986	**0.006282**
(ASV 579) Lactococcus	2.050847	1.484812756	0.379081458	3.916869909	8.97E-05	**0.0006098**
(ASV 572) Treponema	2.274543	1.507357485	0.392294733	3.842410712	0.000121832	**0.0007754**
(ASV 431) Johnsonella	5.890249	1.639351042	0.499135669	3.284379668	0.001022071	**0.0050145**
(ASV 346) Treponema	9.946681	1.658975432	0.53992706	3.072591757	0.002122086	**0.0090864**
(ASV 486) Tissierellaceae	4.319474	1.676734523	0.468588365	3.578267509	0.000345879	**0.0018304**
(ASV 558) Treponema	2.47331	1.708222166	0.411802165	4.148162179	3.35E-05	**0.0002631**
(ASV 405) Treponema	6.836153	1.726097104	0.49410201	3.493402313	0.000476907	**0.0024684**
(ASV 379) Shuttleworthia	8.872646	1.740210699	0.518580501	3.355719496	0.000791588	**0.0039664**
(ASV 574) Prevotella	2.284938	1.783800686	0.381750278	4.672689942	2.97E-06	**3.89E-05**
(ASV 353) Leptotrichia	9.676786	1.818259105	0.569177012	3.194540655	0.001400536	**0.0065312**
(ASV 534) Prevotella	3.050445	1.825544486	0.444524296	4.106737252	4.01E-05	**0.0003098**
Lachnoanaerobaculum	2.595109	1.876055693	0.416386563	4.505562518	6.62E-06	**6.93E-05**
(ASV 455) Bacillaceae	5.742184	1.896202756	0.490205893	3.868176173	0.000109652	**0.0007075**
(ASV 513) Corynebacterium	3.554549	1.926481018	0.424901272	4.53394976	5.79E-06	**6.65E-05**
(ASV 503) Parvimonas	3.849036	1.938093865	0.428655593	4.521331106	6.15E-06	**6.75E-05**
Lachnoanaerobaculum	9.034562	1.964939428	0.538893666	3.646247026	0.000266098	**0.0014921**
Enterobacteriaceae	14.43915	1.968447581	0.578920992	3.400200732	0.000673364	**0.0034473**
(ASV 248) Actinomyces	17.33176	2.007705087	0.603660078	3.325886805	0.000881377	**0.0043698**
(ASV 511) Buchnera	3.801426	2.033405071	0.473436115	4.294993572	1.75E-05	**0.000153**
(ASV 199) Prevotella	25.24773	2.1126958	0.684319156	3.087296012	0.002019863	**0.0088088**
(ASV 262) Corynebacterium	17.14933	2.160163891	0.60315091	3.581465029	0.000341673	**0.0018287**
(ASV 439) Bacteroidales	5.636316	2.162568322	0.493375685	4.383208151	1.17E-05	**0.0001124**
(ASV 261) Atopobium	16.03729	2.304553195	0.731790395	3.149198473	0.00163719	**0.0074866**
(ASV 433) Lachnoanaerobaculum	5.799905	2.409547313	0.532994539	4.520772985	6.16E-06	**6.75E-05**
(ASV 509) Schwartzia	3.510893	2.481982452	0.469377068	5.287822138	1.24E-07	**2.92E-06**
(ASV 367) Prevotella	9.090898	2.589051799	0.543694597	4.761959777	1.92E-06	**2.83E-05**
(ASV 453) Burkholderia	5.693623	2.59088875	0.484167871	5.351219904	8.74E-08	**2.17E-06**
(ASV 417) Actinomyces	5.935608	2.605344967	0.468866376	5.556689705	2.75E-08	**8.09E-07**
(ASV 302) Mogibacteriaceae	12.50191	2.658400812	0.558356571	4.761116731	1.93E-06	**2.83E-05**
(ASV 277) Prevotella	14.46803	2.795148406	0.583457924	4.790659774	1.66E-06	**2.61E-05**
(ASV 264) Eikenella	15.66601	2.996447324	0.599860031	4.995244176	5.88E-07	**1.06E-05**
(ASV 443) Bacillus	5.945412	3.668012258	0.465517269	7.879433269	3.29E-15	**5.16E-13**
(ASV 287) Desulfobulbus	13.47114	4.163647288	0.521145003	7.989421875	1.36E-15	**3.19E-13**
(ASV 374) Fusobacterium	8.298181	4.474907355	0.455959568	9.814263511	9.77E-23	**4.60E-20**

*positive log2FoldChange indicates an increase in abundance, while a negative log2FoldChange indicates a decrease in abundance in those with high levels of HPB-releasing adducts compared to those with very low levels;

** padj or q-value – FDR corrected significance level (alpha) set at 0.01

CHAPTER 4

Relationship between sociodemographic factors and the oral microbiome in individuals who smoke

Introduction

Cigarette smoking continues to be a major risk factor for many diseases, including cancer and cardiovascular disease (CVD), disproportionately affecting certain subgroups of the population,[97] such as those with a lower level of education, lower income, a specific race/ethnicity, and/or sexual orientation, and those who have mental illnesses. An individual's sociodemographic background plays a significant role in shaping their overall health and has been shown to contribute to a range of risk factors for tobacco use, such as initiation, current use, intention to quit, relapse, and tobacco-related mortality and morbidity.[98] A person's socioeconomic status (SES) may be influenced in a variety of ways by their racial, gender, marital, living, educational, income, or employment statuses, all of which may have varying effects on their health. Among the strongest and most consistent epidemiological relationships researchers have observed is the link between SES and health, where higher SES leads to improved health status.[99] Low socioeconomic status can have a negative impact on one's health in many different ways. For instance, poor housing conditions, poor educational opportunities, inadequate health care, and hazardous working environments can all increase the risk of certain diseases. Additionally, research has demonstrated that these factors act as surrogate measures of psychosocial stress, which in turn are closely linked to health disparities.[100] Racially and ethnically minoritized populations often experience higher rates of chronic diseases, poorer health outcomes, and limited access to healthcare services.[101] Education, income, and employment are all strong determinants of health, influencing access to resources, health services, and living conditions. It has also been reported that individuals who have lower levels of education, lower incomes, and unstable careers often experience higher levels of stress and are less likely to engage in activities that promote their health.[102-104] The use of tobacco and its associated health harms provides a good example of how these factors interact to perpetuate inequalities among certain subpopulations. Smoking prevalence and health outcomes are disparate across socio-demographic groups. Using the Multi-Ethnic Cohort Study (MEC), an observational study involving African Americans, Native Hawaiians, Whites, Latinos, and Japanese participants, the relative risks of 1,749 lung cancer cases were calculated. The risk of lung cancer in Latinos, Whites, and Japanese participants was 30% to 75% lower than among African Americans and Native Hawaiians at similar levels of smoking. As the smoking intensity increased to 30 cigarettes per day, the

differences in relative risk became nonsignificant.[105] It is however very important to understand how sociodemographic factors relate with tobacco-associated health outcomes. According to the same study (MEC), both vocational training and some college attendance were associated with a lower risk of lung cancer than completing less than eight years of education.[106] It was also reported in another study that African American men with less than 12 years of education had the highest death rates from lung cancer.[107] A study pooling data from 37 studies examined the association between education and oral cancer, and found that low education was associated with two times greater risk for oral cancer than those with higher education. [108] Those with lower levels of education and income are more likely to experience stress.[102-104] Study that examined the relationships between psychosocial stress and smoking among a national cohort of 4,938 U.S. adults at baseline and at follow-up, 9 to 10 years later found that high levels of psychosocial stress (relating to relationships, finances, employment, perceived inequality, past-year family problems, and overall stress) at both time points, baseline and follow-up significantly increased the odds of persistent smoking.[109]

Recent research has shown that the oral microbiome is likely to be affected by individual characteristics and stressors associated with the burden of tobacco-induced diseases. Some of the socio-demographic factors mentioned above have been implied as important determinants of the oral microbiome. Microbiome composition has been shown to be influenced by human host genetic variations.[110] Studies conducted on monozygotic twins have shown that their microbiomes are more similar than those of dizygotic twins.[111,112] The oral microbiome, however, may be influenced more by lifestyle, social structures, and shared environments than by intrinsic factors like genetics.[113,114] A constant exposure to chemicals and toxicants, such as that caused by cigarette smoking, disrupts the bacterial environment and results in distinctive microbial profiles between those who smoke and those who do not. [15,115] Certain microbial profiles are also linked to several tobacco-related diseases, including CVD and cancer.[116-118] In addition, some studies have reported differences in microbial compositions between racial/ethnic groups.[39,40,115] Only a small number of studies have examined how each of these sociodemographic factors affects oral microbiome variability. Importantly, these studies did not consider smoking status. On one hand, this could have made it more difficult to detect differences in the oral microbiome across sociodemographic factors, since smoking has been shown to alter the oral microbiome composition significantly. On the other hand, such approach does not allow to explore the potential contribution of sociodemographic factors to the variations in the oral microbiome among persons who smoke. In this study,

this gap was addressed by exploring the potential effects of self-identified race and other sociodemographic factors such as marital status, education, annual income, and employment on the microbiome characteristics in individuals who smoke. Due to the complexity of racial and ethnic identities, hereafter referred to as race, including aspects of culture, economics, geography, and history, it is important to note that in this study, race is not considered a biological category but rather a social construct that describes differences based on multiple intersecting social, environmental, and cultural factors.

Materials and Methods

Participant recruitment and study design

For this study biological samples and questionnaire information that were collected as part of the Mechanisms of Ethnic/Racial Differences in Lung Cancer Due to Cigarette Smoking study were used, detailed design of which was described in Chapter 3. A major objective of this study was to explore differences in metabolic activation and detoxification of tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), along with DNA adduct formation and DNA repair capacities in African American (AA) and White (WH) individuals who smoke. Newspapers, radio, and television advertisements were placed to recruit study participants from the local metropolitan area. Inclusion criteria included participants with a smoking history of at least 10 cigarettes per day for at least a year who were in good physical and mental health. Participants were classified as African American if they identified themselves, their biological parents, and both sets of their biological grandparents as African American. The same criteria were applied to Whites (e.g., self, both parents, and four grandparents).

Questionnaire information

At the screening visit (Visit 1), eligible participants signed a consent form and filled out two structured questionnaires, the History of Tobacco Use Questionnaire and the Medical History Questionnaire. The questionnaires (questions and available responses) are presented in Appendix I. Collected data from the questionnaires reflected surrogate measures of psychosocial stress related to racial inequality, relationships, finances, education, and employment.

Responses were re-categorized into different categories to facilitate statistical analysis. We recategorized marital status questionnaire responses into four main categories: "Currently Married"; "Legally Divorced"; "Not Married" and "Separated". In the category "Currently Married", respondents who were "Married for the first time" as well as "Remarried" were grouped together. "Widowed" responses were grouped under "Not Married". The responses about living situations "With friends/other relatives"

and "With parents" were also grouped together. The "Other" category has been reclassified accordingly based on the responses provided. Data from the annual income survey was also re-categorized to create the category called "$50,000 or more" by combining "$50,000-69,999", "$70,000-99,999", and "$100,000 or more". Additionally, the Education categories of "High School Graduate/Equivalent" and "Some High School" were combined into "High School"; "Some College" and "Associate Degree" were combined into "Associate Degree"; "College Graduate/4-year Degree" and "Graduate-Professional Degree" were combined under "Bachelor's and Graduate Degree". A new category was created for the current employment category by grouping the responses for “Homemaker” and “Other” together as “Homemaker and other.” In the same way, data on employment during the past 3 years was recategorized by grouping responses of “Employed at least half of the time" and "Employed less than half of the time" as "Part-time", responses of "Employed virtually all of the time" as "Full-time" and "Not employed at all" as "Unemployed". In general, the re-categorization of data simplifies analysis and facilitates a more meaningful interpretation of results.

Oral cell collection

Oral cells were collected as previously described (Chapter3). In brief, participants were asked to brush their teeth one hour before their visit to Tobacco Research Center. They were also asked to refrain from smoking, drinking, and eating during that time. Participants were given two sterile cytobrushes and instructed to brush the insides of their left and right cheeks with a new brush each. Each brush was placed in a separate polypropylene tube containing 10 ml of Scope mouthwash and frozen until DNA extraction was completed.

DNA isolation from oral mucosa cells and microbiome analyses

DNA from oral swabs was previously described by the modified Qiagen DNA isolation protocol for DNA isolation from tissues and blood.[35,36] Briefly, the samples were thawed at room temperature and centrifuged at 1500 × g for 15 minutes. Afterward, the supernatant was discarded, the cell pellet was resuspended in 3 ml of Qiagen kit cell lysis solution, homogenized, and the DNA was isolated and purified according to the manufacturer's instructions. An Illumina MiSeq Sequencing platform was used to amplify and sequence the V3-V4 region of the 16S rRNA gene barcode primer pairs (515f-GTGCCAGCMGCCGGTAA and 806r-GGACTACHVGGGTWTCTAAT) for assessment of the oral microbiome.

Urinary total nicotine equivalents (TNE)

TNE is a urinary biomarker of daily nicotine intake. It is expressed as a molar sum of nicotine and its known metabolites, accounting for approximately 73-96% of the nicotine dose.[67,68] Urinary TNE was analyzed as previously described in Chapter 3.

Bioinformatics and statistical analyses

All quality control, trimming, and merging were completed in QIIME2 using the DADA2 plugin. The raw reads were trimmed to remove primers using cutadapt and filtered to remove low-quality reads (less than Q ¼ 20) using fastx_toolkit. Only high-quality reads are considered for downstream analysis. Bacterial taxonomy was assigned using a pre-built classifier on Greengenes 16SrRNA sequences.

The R statistical interface was used to perform all microbial community ecology analyses. The R phyloseq package[71] was used for the analysis of microbiome diversity characteristics: alpha diversity indices (Observed Richness and Shannon's indices), distance matrices (Bray-Curtis and UniFrac), ordination analyses [principal coordinates analysis (PCoA)]. Permutational multivariate analysis of variance (PERMANOVA) was calculated with the adonis and anosim functions within the R vegan package.[72] The Kruskal-Wallis and Wilcoxon rank-sum tests were used to determine the differences across groups. To asses the discriminant features Indicator species analyses within the labdsv package was preformed [119] and edgeR package.[120] Machine learning methods such as Random Forest (RF) classification[94] was also used to assess the classification accuracy of race based on taxonomic features. MicrobiomeSeq package[121] and Spearman correlation were used to assess the relationship between the most abundant taxa and smoking related variables across different groups, the threshold p-value was set to 0.05 adjusting for multiple comparisons using Benjamin and Hochberg. All plots were created using the R package ggplots.[74]

Results

This study included 146 participants out of 162 who had completed the study. There were 16 participants excluded because of missing data (relevant biomarker or questionnaire data) or low-quality sequence reads. The summary of participants' characteristics and biomarker levels by race is presented in Table 4-1. The age of all participants ranged from 37 to 52, and there was even distribution by biological sex (71 male and 75 female) and race (75 AA and 71 WH). In all study participants, the median self-reported cigarettes smoked per day (CPD) were 15 (12-20). AA participants smoked fewer CPD when compared to WH ($p < 0.001$). However, there was no difference in urinary biomarkers of tobacco smoke exposure, the median TNE were 54 (36.03-78.06) in AA and 54.36 (35.41-81.87) in WH (p=0.789). In terms of marital status and living situation, there was no difference between AA and WH (p=0.489 and p=0.263,

respectively). AA had lower educational levels, lower incomes, and higher unemployment when compared to WH ($p<0.001$ for all).

Oral microbiome profiles

After quality filtering, 11,405,326 16S rRNA sequence reads were retained, reflecting 2477 unique amplicon sequence variants (ASVs), each representing a unique taxon. Using further filters (e.g., dropping taxa present in fewer than 3 samples), 595 unique ASVs were included. As shown in **Figure 4-1**, the distribution of bacterial genera by race is generally similar across races. The dominant genera were *Actinomyces, Fusobacterium, Leptotrichia, Prevotella, Rothia, Fusobacterium, Streptococcus, Unclasified Gemellacea* and *Veillonella*,

Taxonomic composition and diversity

There was a significant difference in alpha-diversity metrics Observed Richness and Shannon indexes between AA and WH individuals who smoke (FDR-corrected pairwise Wilcoxon rank-sum test, p =0.0035 and p =0.0234, respectively); **(Fig. 4-2A)**. In addition, alpha-diversity metrics differed by living situation, individuals who lived alone had significantly lower alpha diversity (Observed Richness and Shannon indexes) when compared to those who lived with parents, friends, or other relatives (p=0.012 and p=0.0005 respectively); **(Fig. 4-3B)**. These metrices did not differ by biological sex, BMI, marital status, education, income nor by employment in this cohort of participants **(Fig. 4-2; Fig. 4-3)**.

To understand the variation of the microbiome between the groups with different sociodemographic characteristics Bray–Curtis dissimilarity and UniFrac distances were computed. There were significant differences in beta diversity metrices between AA and WH individuals who smoke for both indices Weighted Bray-Curtis (PERMANOVA: R^2=0.02, p=0.008; ANOSIM: R^2=0.02, p=0.008) and weighted UniFrac (PERMANOVA: R^2=0.03, p=0.008; ANOSIM: R^2=0.043, p=0.001, **[Fig. 4-4]**). Significant differences in Weighted Bray-Curtis by biological sex and BMI were also found (PERMANOVA: R^2=0.03, p=0.044, PERMANOVA: R^2=0.03, p=0.042, respectively). However, there were no statistically significant differences in beta diversity across other sociodemographic factors **(Fig. 4-5)**.

Association between race and microbial composition

To assess the relationship between race and microbial composition, Taxomomic Summary Score (TSS), was calculated as described in Chapter 3. The relation between the continuous variables (TSS, response variable) and race was modeled using a multivariable regression model. **Table 4-2** summarizes the results of the multivariable analysis. Regression model showed that there was a statistically significant

association between race and TSS (β = -27.78, p = 0.008), suggesting that White individuals who smoke are associated with lower levels of TSS (β = -0.24, p = 0.008) after accounting for the effects of smoking (TNE, NNAL and years of smoking).

Differential bacterial abundance

As AA individuals who smoke are disproportionately affected by smoking-related harm, the goal was to identify specific taxa that are significantly different from those of WH individuals who smoke. Indicator Species Analysis was performed and indicator values were calculated based on relative abundance data using the labdsv package (v2.0-1). Significant indicators are presented in **Supplemental Table 4-1**. To determine statistical significance, a p-value of 0.05 was used. Here, the p-value corresponds to confidence in determining an indicator value, not a difference in means. This analysis identified 73 significant taxa indicators at ASV level, of which only *Prevotella, Rothia*, and *Unclassified TM7.3* were associated with WH individuals who smoke. In order to determine if any differentially relative abundant taxa were identified as the result of autocorrelation with other taxa, a pair-wise comparison with edgeR, corrected for false discovery rate (FDR), and an acceptable threshold of p<0.001 was conducted. This analysis showed that several taxa belonging to the genus *Haemophilus, Fusobacterium, Prevotela, Streptococcus* and *Vellonela* to be significantly increased in AA when compared to WH individuals who smoke. **(Fig. 4-6)**. In addition, machine learning classification approaches like Random Forest (RF) classification models (using mean decrease accuracy index) were used to identify discriminatory features between AA and WH who smoke. For the RF model the dataset was split 70% for training and 30% retained for testing, the model was cross-validated 10-fold and repeated 5 times. Based on the taxonomic features of the oral microbiome of AA and WH individuals who smoke, the model performed poorly (with error rates of 22.31 and 17.30%, respectively, and an AUC of 0.576 **[Fig. 4-7]**).

Correlations between most abundant taxa at genus-level ASV and sociodemographic characteristics

Correlation between the top thirty most abundant taxa and relevant smoking biomarkers is shown in **Supplemental Figure 4-1.** It is interesting that *Atopobium, Lactobacilus, Megasphaera, Scardovia, Tanerella, Veillonellaceae* were positively related to smoked CPD in WH, whereas these taxa were negatively related to TNE. The most abundant taxa in AA smokers showed similar pattern of associations with smoking-related biomarkers. CPD and TNE were positively related to *Campylobacter, Fuscobacterium, Moryella, Tannerella, Treponema* and *Veillonella* while *Granulicatella, Lactobacilus, Rothia, Seleomonas and Streptococcus* were negatively related. TNE was positively correlated with *Actinomyces, Aggregatibacter,*

Atopobium, Campylobacter, Capnocytophaga, Corynebacterium, Fuscobacterium, Porphyromonas, Prevotella, Tannerella, Treponema in males, but not in females. There was a positive correlation between *Megasphera* and CPD and TNE, and a negative correlation between *Actinomyces, Aggregatibacter, Atopobium, Campylobacter* in separated individuals. In those who live alone, the majority of the taxa were negatively correlated with CPD, but not with TNE. The most abundant taxa were positively correlated with CPD in those with incomes of $20.000 to $50,000. The most abundant taxa showed a weak correlation (both positive and negative) between smoking-related biomarkers and employment status. After applying Benjamin and Hochberg's correction for multiple comparisons, it was found that none of the correlations had achieved statistical significance with a p-value below 0.05.

Discussion

In this study, the relationship between sociodemographic factors and the oral microbiome in individuals who smoke was explored. The study included 146 healthy individuals who smoked and showed that most of their oral microbiomes were comprised of eight genera (*Actinomyces, Fusobacterium, Leptotrichia, Prevotella, Rothia, Streptococcus, Unclassified Gemellaceae*, and *Veillonella*) belonging to three distinct phyla (*Actinobacteria, Fusobacteria, Bacteroidetes* **[Fig.4-1]**). In 99% of the samples, a total of 128 ASVs were identified, contributing to 46.5% of the global composition of the oral microbiome, indicating that the oral microbiome tend to share similar genera among individuals. Similarly, the Human Oral Microbiome Project found that *Streptococcus* dominates the oral microbiom, followed in abundance by *Haemophilus, Actinomyces, and Prevotella.*[94] A recent study that re-analyzed raw 16S rRNA amplicon sequencing data from 47 studies with 2206 saliva samples and found 68 core bacterial taxa that were consistently detected among different populations.[122]

In this study, AA individuals exhibited higher alpha diversity, as indicated by increased Observed Richness and Shannon Index values **(Fig.4-2).** Furthermore, there were significant differences in beta diversity metrics between AA and WH individuals who smoked, as measured by both the Weighted Bray-Curtis Index and the Weighted UniFrac Index **(Fig. 4-4)**. The Weighted UniFrac Index incorporates phylogenetic relationships among microbial taxa, considering their relative abundances. This metric operates on the premise that closely related taxa share ecological traits and respond similarly to environmental changes, leading to comparable community structures. The observed differences in Weighted UniFrac between AA and WH individuals who smoke suggest substantial dissimilarities in both

taxonomic composition and phylogenetic relatedness within their microbial communities, highlighting significant variations in the overall composition and structure of the microbiomes between AA and WH individuals who smoke.

These results are consisted with previous findings from the NIH Human Microbiome Project. They investigated whether differences in microbiome ecology are associated health-related outcomes and concluded that "ethnic/racial background proved to be one of the strongest associations" of microbiome composition.[94] Also, an investigation of racial differences in the oral microbiome in a subset of the Southern Community Cohort based in the Southeastern region of the US (N=1058 African Americans, N=558 European Americans), reported differences in overall microbial composition (African Americans were more likely to have higher alpha diversity than European Americans). In that study, African Americans were classified by both self-reported race and percentage of genetic African ancestry, and the microbial differences were consistent using both metrics.[39] Besides the difference in overall microbial composition, several common taxa were differentially abundant between AA and WH **(Fig. 4-6; Supplementary Tables 4-1 and 4-2)**. Overall, AA had a higher abundance of *Proteobacteria* and a lower abundance of *Actinobacteria* and *Bacteroidetes*. Indeed, the regression model showed that WH individuals who smoke had significantly lower TSS than AA adjusted for smoking. **(Table 4-2).** It is important to note that race is a social construct. Therefore, self-identified race in this study does not denote biological or genetic differences. Instead, it reflects the cumulative impact of intersecting social, environmental, and cultural factors. There is a need to further explore the factors contributing to the observed racial differences in the taxonomic composition of the oral microbiome among individuals who smoke. When using machine learning classification models such as the Random Forest (RF) approach to identify discriminatory features, the model's performance was poor with error rates of 22.31% and 17.30% for AA and WH individuals, respectively, and an AUC of 0.576 **(Fig. 4-7)**. These findings suggest that the oral microbiome's taxonomic features may not serve as effective discriminatory features for identifying differences between AA and WH. However, an earlier study used a machine-learning classifier (RF) and was able to identify an individual's race/ethnicity by analyzing their oral microbiome with higher specificity, 74%.[115] There could be several reasons for this discrepancy. That study had a different population group (non-Hispanic AA, non-Hispanic WH, Chinese, and Latinos) and sequencing technique (PCR with fluorescent-labeled broad-range bacterial primers A18-FAM and 317-HEX), which could have contributed to the differences in discriminatory features identified by the machine learning classifier. Also, the current study focused on identifying the differences between AA and WH individuals

who smoke, unlike the previous study that did not consider smoking status, which could have made it more difficult to detect discriminatory features since smoking has been shown to alter the oral microbiome composition significantly. Finally, there is the possibility that taxonomic characteristics of the oral microbiome are not reliable discriminators. Other functional features such as metabolic pathways may be needed to identify differences in the oral microbiome between AA and WH individuals who smoke.

Two types of differential abundance analyses consistently identified nine differentially abundant taxa between AA and WH individuals who smoke suggesting that these taxa may have biological importance (Supplementary Tables 4-1 and 4-2). These nine taxa were *Abiotrophia (ASV 121), Prevotella intermedia (ASV 50), Unclassified Leptotrichia (ASV 65), Unclassified Leptotrichia (ASV 72), Unclassified Porphyromonas (ASV 80), Unclassified Prevotella (ASV 57), Unclassified TM7.3 (ASV 67), Unclassified TM7.3 (ASV 217), and Unclassified Veillonella (ASV 37)*. Oral *Abiotrophia* species have been identified as causative agents of infections in distant body parts, such as the joints, the urinary tract, the cardiovascular and the central nervous system.[123-125] The gram-negative bacteria *Prevotella* has been associated with several systemic diseases, including cardiovascular disease (CVD) and cancer.[118,126] Oral *Prevotella* has been found to be more present in atherosclerotic plaques[127] suggesting a potential role in CVD development and progression. *Leptotrichia* species are anaerobic bacteria commonly found in the oral cavity and have been associated with the development of CVD.[128] Moreover, some studies have shown that *Leptotrichia* may play a role in the development of cancer, such as oral squamous cell carcinoma.[129] Another genus of gram-negative anaerobic bacteria commonly found in the mouth is *Porphyromonas*. *Porphyromonas gingivalis*, a species within the *Porphyromonas* genus, has been connected to high plasma levels of C-reactive protein (CRP), a biomarker of cardiovascular disease risk.[130] Furthermore, studies have shown that *P. gingivalis* in people who smoke increases inflammation and contributes to endothelial dysfunction, a major cause of atherosclerosis. This suggests yet another link between smoking, periodontal disease, and cardiovascular risk mediated by inflammation.[126] The Gram-negative bacterium *Veillonella* is generally considered commensal, but research suggests that it is associated with smoking-mediated responses to nitric oxide-associated metabolic pathways plausibly impacting cardiovascular risk.[131] It is interesting to note, however, that many of these nine taxa were positively correlated with smoked CPD in both AA and WH individuals who smoke (Supplemental Figure 4-1). Since these bacteria have shared associations with CVD and/or cardiovascular risk, they may be of relevance to future studies seeking to understand the relationship between smoking and elevated CVD risk through the oral microbiome.

Similar to previously reported findings,[40, 96, 113] compositional differences in the oral microbiome by biological sex and BMI were found in this study. However, unlike other studies, differences in microbial profiles across other sociodemographic factors (marital status, education, income or employment) in this cohort of participants were not detected. There may be a number of reasons for this difference. First, differences in study populations between this study and other studies (which included a Canadian cohort [96], a cohort of residents of New Yorker city [40], and Danish cohort [132], and the related potential differences in environmental, dietary, and other life-style factors may have played a role. Different methodologies for microbiome analysis were also used across the studies. For instance, differences in sequencing platforms and bioinformatic pipelines may affect taxonomic classifications and abundance estimates of microbes. The heterogeneity and variations in the assessments of sociodemographic variables investigated in these studies and current study could also account for the differences in findings. In the Danish population-based study[132], for example, socioeconomic status was determined based on average annual income, and the percentage of unemployed inhabitants, while in the two other studies, sociodemographic and lifestyle variables were collected via questionnaires and interviews.[40,96]

It is interesting to note that among the participants in this study, those who live alone had significantly lower alpha diversity than those who live with parents, friends, or other relatives (Observed Richness and Shannon indexes). This suggests that living environments may greatly influence oral microbiome composition. Several factors may have contributed to these findings. Living alone could potentially affect dietary patterns and other lifestyle factors and environmental exposures and may also be a psychosocial stressor for some individuals. Such factors are known to affect the oral microbiome.[110-112,133] The limitations of this study are mainly related to the lack of data on certain potentially important confounders. For example, we did not have access to oral health and dental histories of the participants, their dietary patterns, and other factors. It is well established that AA tend to have a higher prevalence of periodontal disease compared to other racial groups.[39,134] A variety of complex factors may be contributing to this, including socioeconomic status and access to dental care. In order to place the findings of studies investigating racial differences in the oral microbiome into context, oral health and other related metrics must be carefully evaluated as potential contributors. In addition, because the study sampled only people living in the Twin Cities metro area, generalization of conclusions may be limited. Finally, the current study characterized the oral microbiome's taxonomic composition. While this approach provides valuable information, it does not provide insight into the functional capabilities of microorganisms. It is important

to understand the potential functional implications of observed taxonomic differences through functional metabolic profiling.

In summary, this study showed that sociodemographic factors may be important determinants of the oral microbiome of individuals who smoke. Oral microbiome and sociodemographic factors may play a role in mediating health inequities (particularly between AA and WH) in tobacco-associated diseases such as CVD and cancer. To better understand these associations, future research should examine multiple levels of exposure. Among these are macro-level social and health policies, psychosocial stressors, outdoor and built environment features, and social interactions. A better understanding of how social environments affect oral microbiomes can lead to the development of effective interventions to reduce the health inequities associated with tobacco use and improve public health.

TABLES AND FIGURES FOR CHAPTER 4

Table 4-1. Study participant characteristics and biomarkers – summary statistics

	Total (N=146)	Black or African American (N=75)	White (Caucasian) (N=71)	p value
Age (years)	45.5	45 (39-52)	45 (37-51)	0.458
Sex (female)	75 (51.4%)	36 (48.0%)	39 (54.9%)	0.402
Cigarettes per day (CPD)	15 (12-20)	15 (11-17)	18 (13-20)	**0.001**
TNE, pmol/mL urine	53.11 (35.92-81.10)	52.54 (36.03-78.06)	54.36 (35.41-81.87)	0.789
Alcohol use (drinks per day)	3 (2-3)	3 (2-3)	3 (2-3.5)	0.848
Marital Status				0.489
Currently Married	16 (11.0%)	6 (8.0%)	10 (14.1%)	
Legally Divorced	31 (21.2%)	14 (18.7%)	17 (23.9%)	
Not Married	88 (60.2%)	48 (64.0%)	40 (56.3%)	
Separated	11 (7.5%)	7 (9.3%)	4 (5.6%)	
Current Living Situation				0.263
Alone	59 (40.4%)	34 (45.3%)	25 (35.2%)	
Friends, other relatives, parents	37 (25.3%)	15 (20.0%)	22 (31.0%)	
Spouse, partner, kids	50 (34.2%)	26 (34.7%)	24 (33.8%)	
Education				**< 0.001**
High School Degree	67 (45.9%)	47 (62.7%)	20 (28.2%)	
Associate Degree	61 (41.8%)	26 (34.7%)	35 (49.3%)	
Bachelors or Graduate Degree	18 (12.3%)	2 (2.7%)	16 (22.5%)	
Annual Personal Income				**< 0.001**
Less than $10,000	75 (51.4%)	52 (69.3%)	23 (32.4%)	
$10,000 - $19,999	37 (25.3%)	19 (25.3%)	18 (25.4%)	
$20,000 - $29,999	10 (6.8%)	0 (0.0%)	10 (14.1%)	
$30,000 - $49,999	14 (9.6%)	3 (4.0%)	11 (15.5%)	
$50,000 and greater	10 (6.8%)	1 (1.3%)	9 (12.7%)	
Current Employment Status				**0.001**
Full Time	28 (19.2%)	9 (12.0%)	19 (26.8%)	
Part Time	23 (15.8%)	6 (8.0%)	17 (23.9%)	
Homemaker and other	20 (13.7%)	11 (14.7%)	9 (12.7%)	
Unemployed	75 (51.4%)	49 (65.3%)	26 (36.6%)	
Employment in past 3 years				**< 0.001**
Full Time	44 (30.1%)	13 (17.3%)	31 (43.7%)	
Part Time	54 (37.0%)	27 (36.0%)	27 (38.0%)	
Unemployed	48 (32.9%)	35 (46.7%)	13 (18.3%)	

*Summaries shown are median (1st quartile, 3rd quartile) or N (percent)

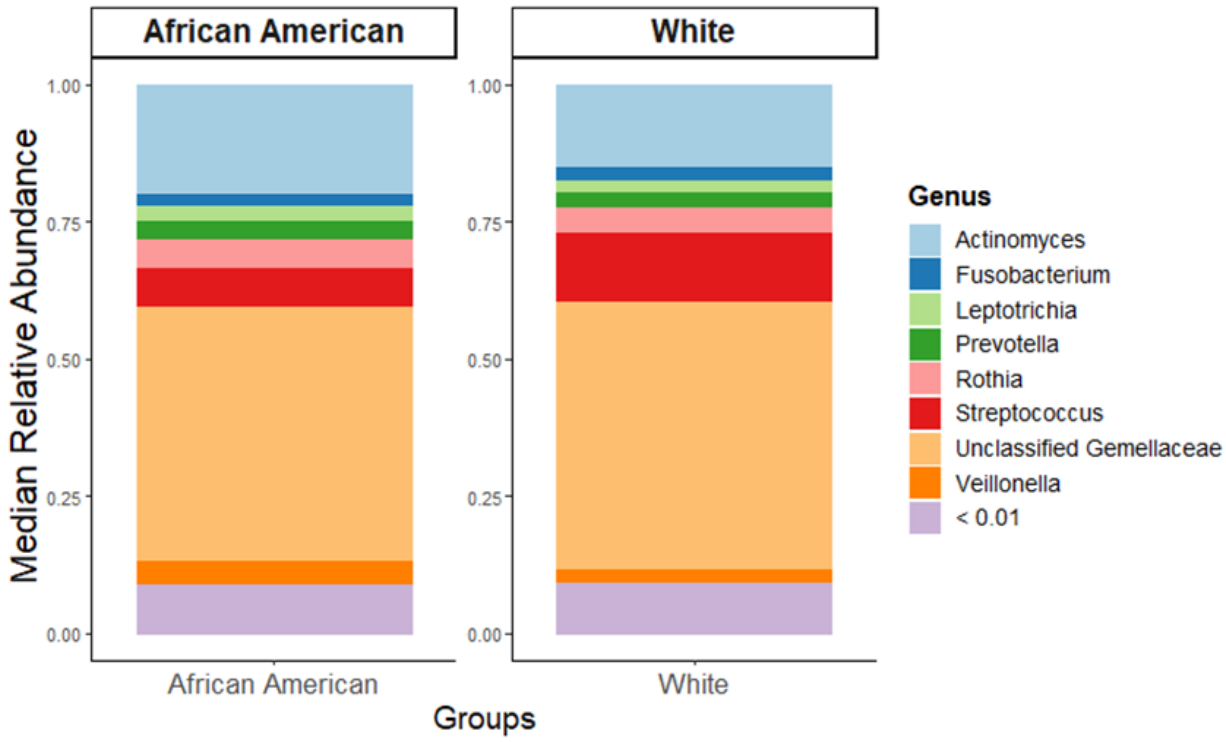


Figure 4-1. Median relative abundance of bacterial genera by race

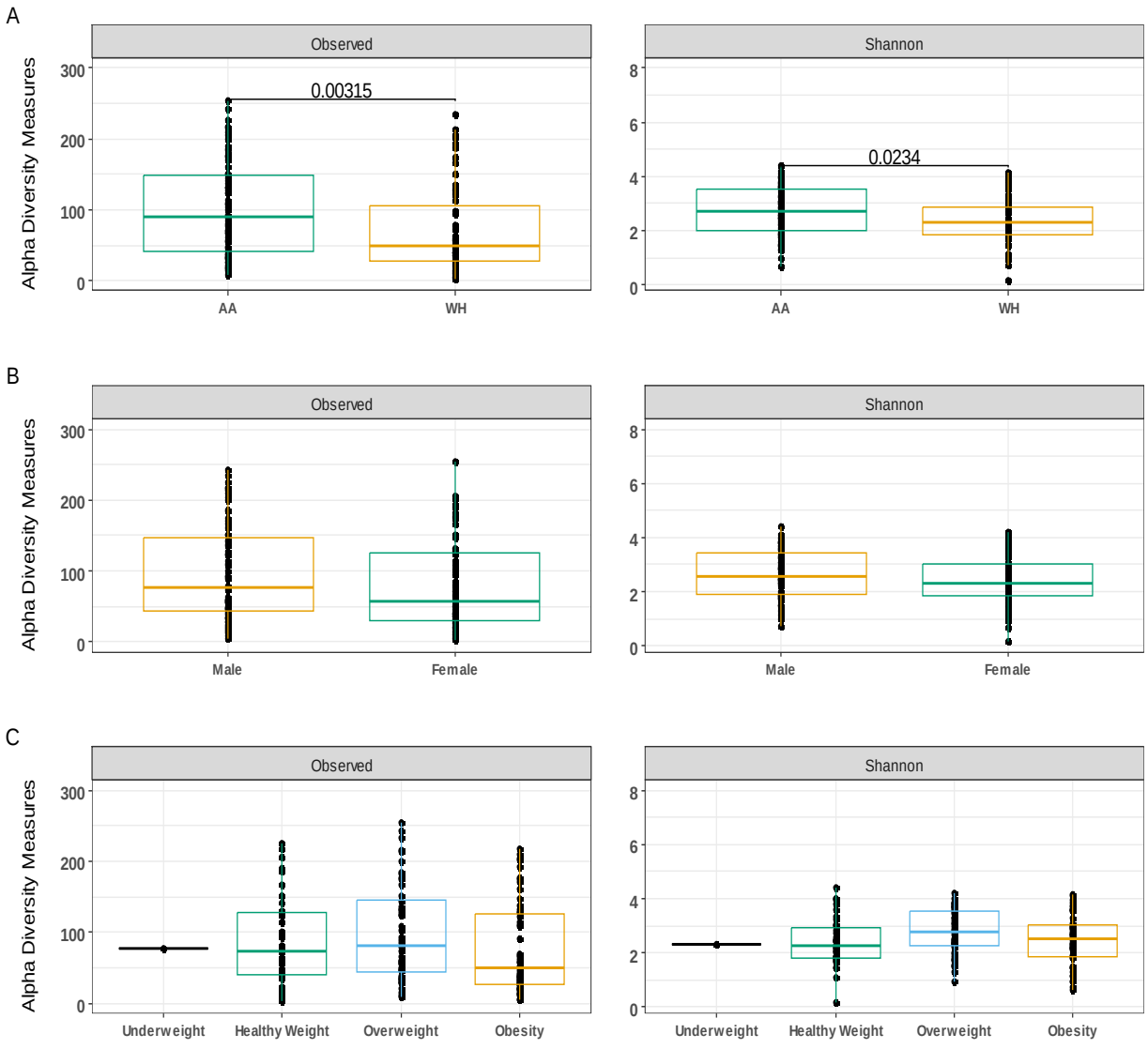
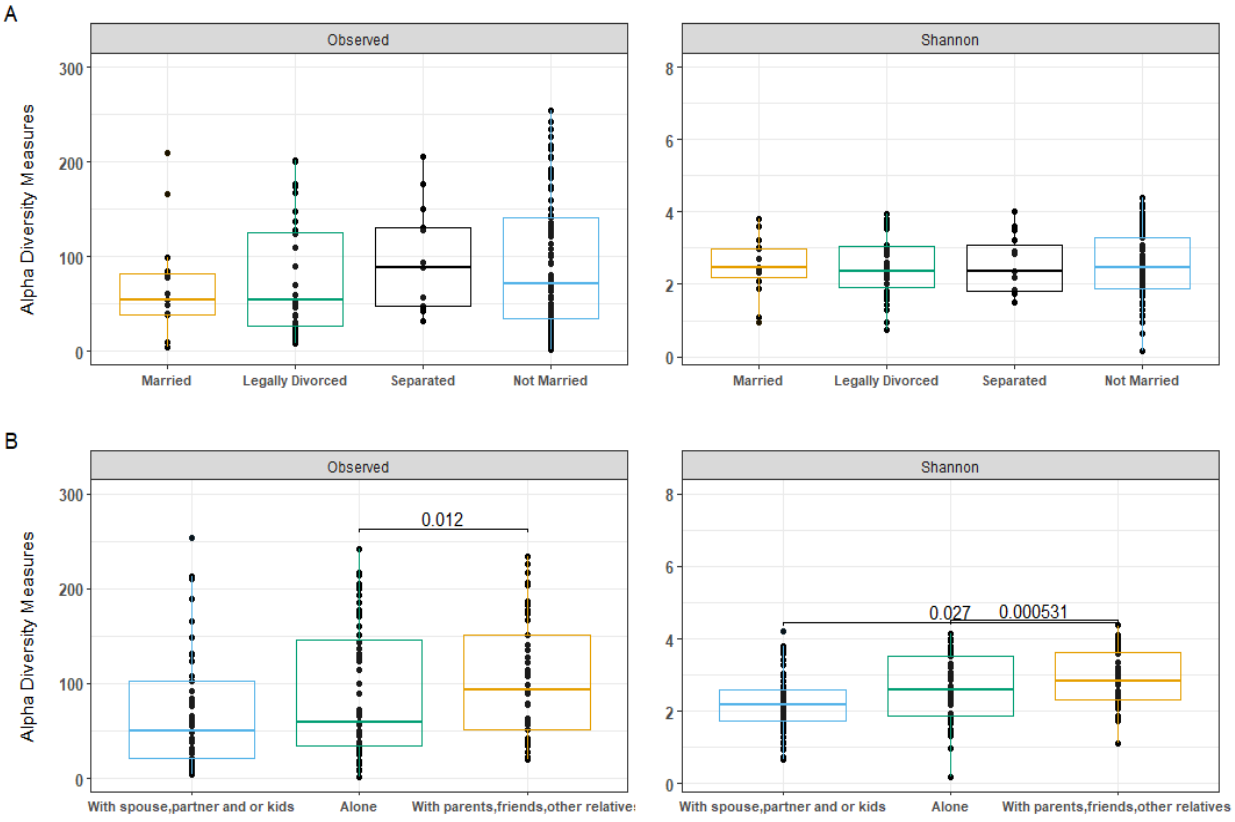
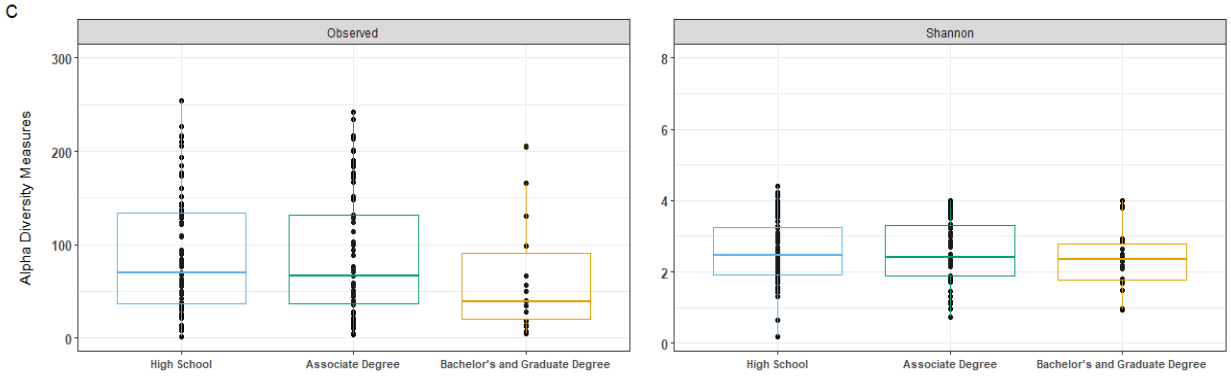


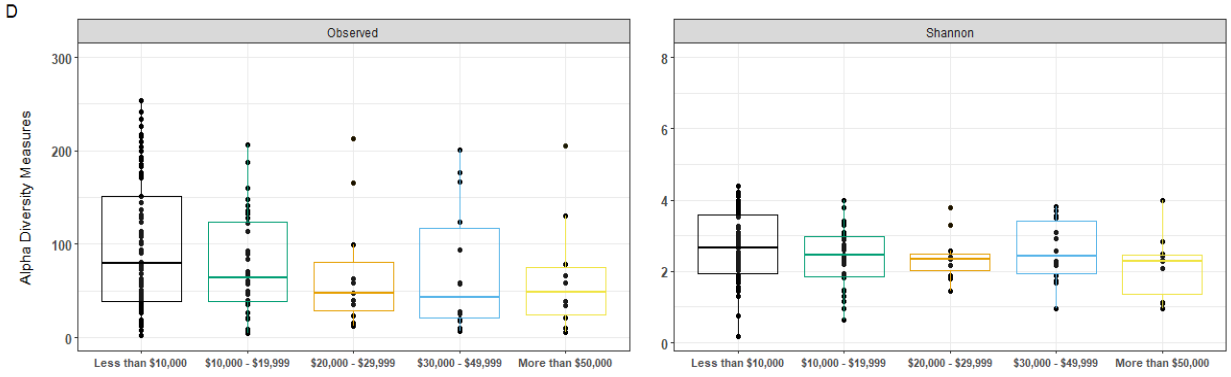
Figure 4-2. Alpha diversity measures (Observed Richness and Shannon Index); The boxplots show differences in each alpha diversity measure by: A) Race B) Biological sex C) Body Mass Index (BMI)


A
Observed
Shannon
Alpha Diversity Measures
300
200
100
0
8
6
4
2
0
Married
Legally Divorced
Separated
Not Married
B
Observed
Shannon
0.012
0.027
0.000531
With spouse,partner and or kids
Alone
With parents,friends,other relatives

C


Observed
Shannon
Alpha Diversity Measures
300
200
100
0
8
6
4
2
0
High School
Associate Degree
Bachelor's and Graduate Degree

D


Observed
Shannon
Alpha Diversity Measures
300
200
100
0
8
6
4
2
0
Less than $10,000
$10,000 - $19,999
$20,000 - $29,999
$30,000 - $49,999
More than $50,000

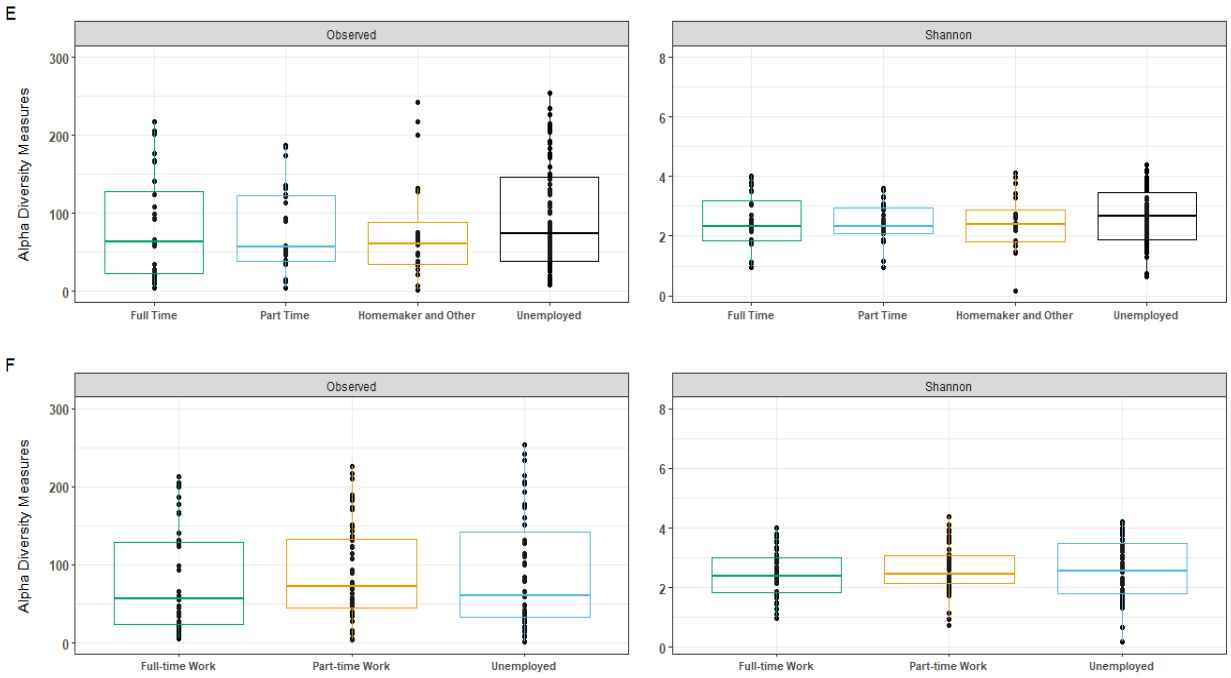


Figure 4-3. Alpha diversity measures (Observed Richness and Shannon Index); The boxplots show differences in each alpha diversity measure by sociodemographic factors in all study participants; A) Marital Status; B) Living Situation; C) Education; D) Annual Income; E) Current Empolyment; F) Employment in the past 3 years

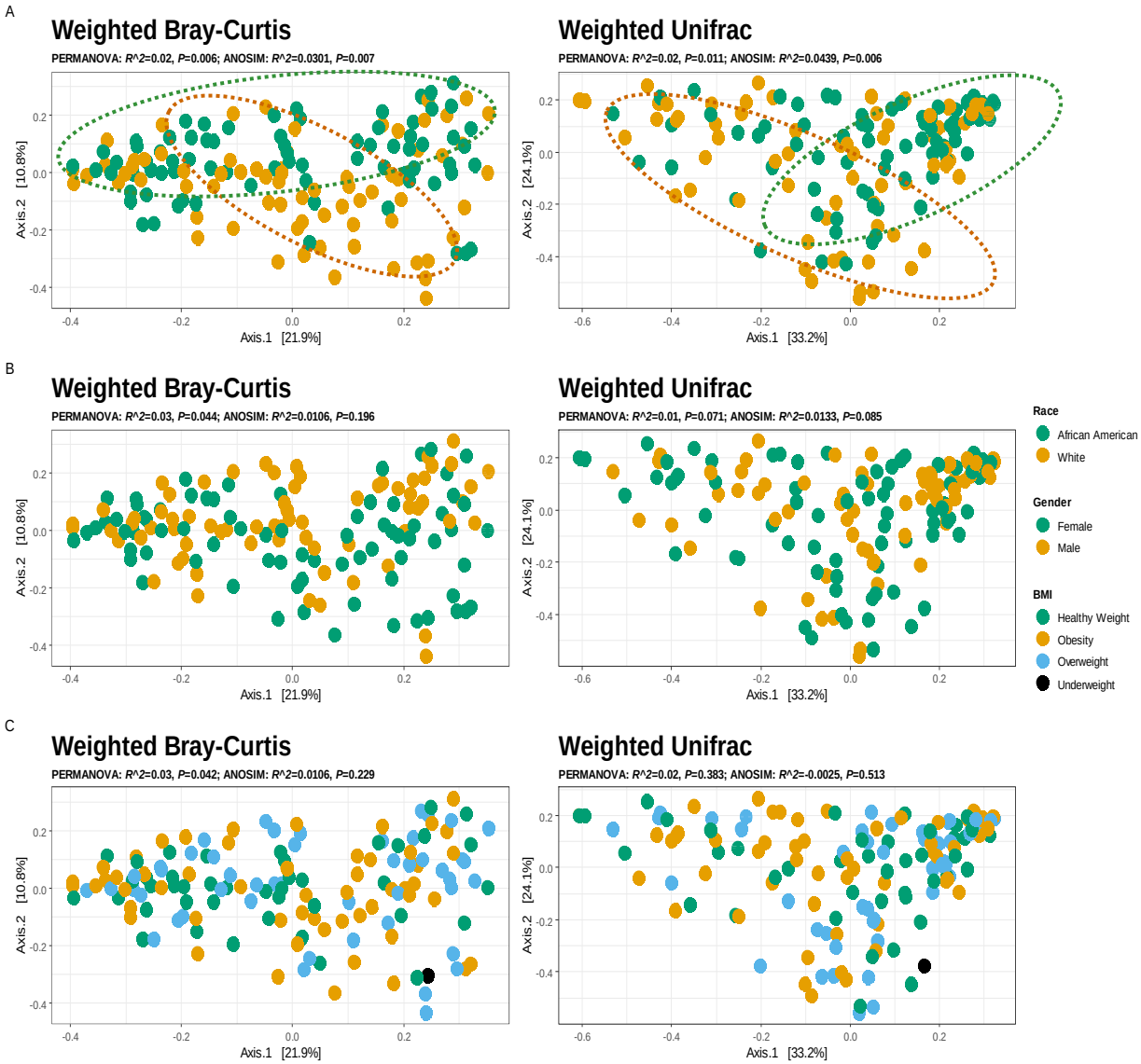
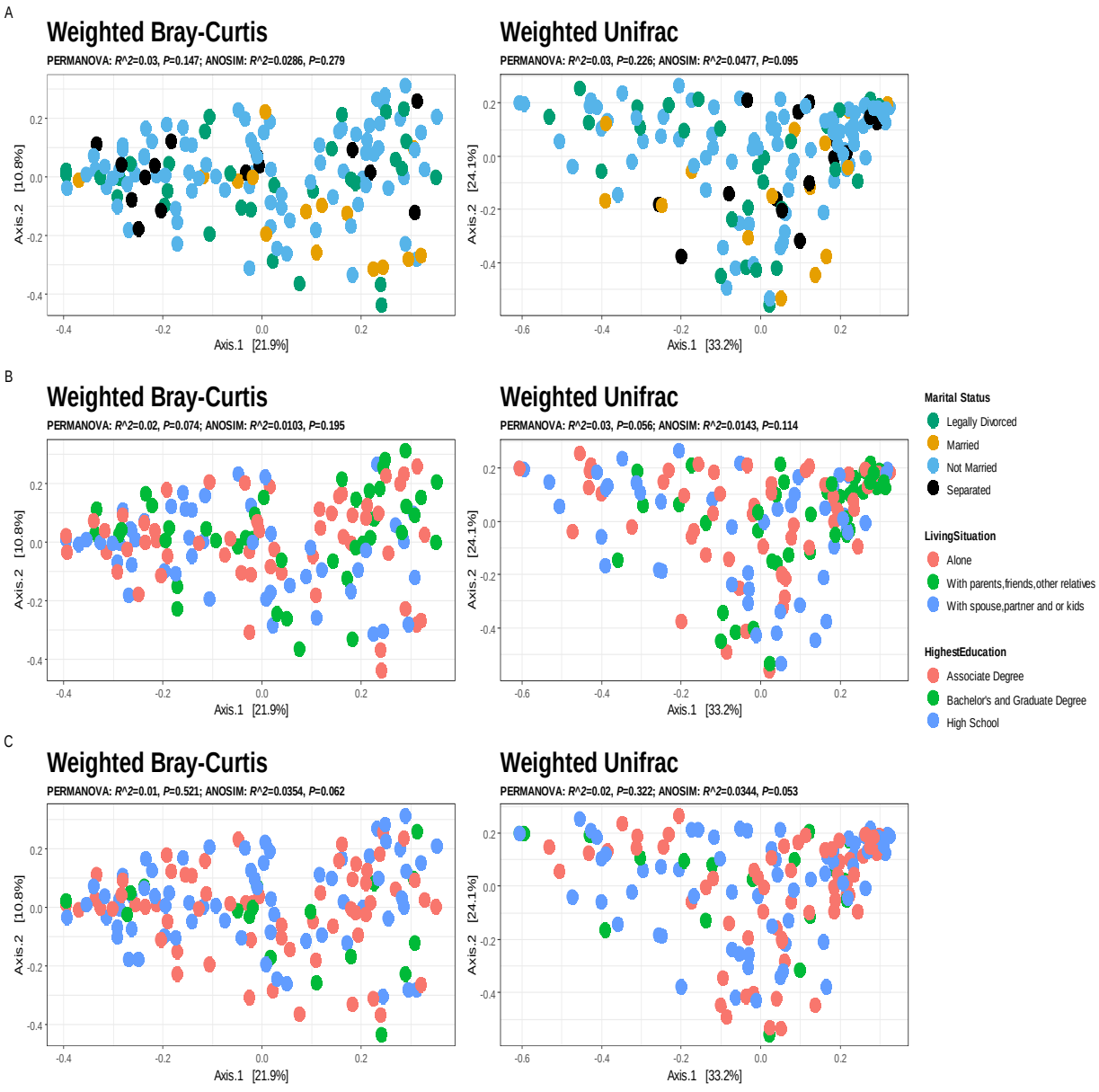


Figure 4-4. Beta diversity based on Weighted Bray-Curtis and Unifrac distances by A) Race; B) Biological sex; C) BMI; Scatter plots shows significant differences for both metrics (Weighted Bray-Curtis and Unifrac distances) by race according to PERMANOVA ($p<0.05$) and ANOSIM ($p<0.05$)


A
Weighted Bray-Curtis
PERMANOVA: R^2=0.03, P=0.147; ANOSIM: R^2=0.0286, P=0.279
Weighted Unifrac
PERMANOVA: R^2=0.03, P=0.226; ANOSIM: R^2=0.0477, P=0.095
B
Weighted Bray-Curtis
PERMANOVA: R^2=0.02, P=0.074; ANOSIM: R^2=0.0103, P=0.195
Weighted Unifrac
PERMANOVA: R^2=0.03, P=0.056; ANOSIM: R^2=0.0143, P=0.114
C
Weighted Bray-Curtis
PERMANOVA: R^2=0.01, P=0.521; ANOSIM: R^2=0.0354, P=0.062
Weighted Unifrac
PERMANOVA: R^2=0.02, P=0.322; ANOSIM: R^2=0.0344, P=0.053
Axis.1 [21.9%]
Axis.2 [10.8%]
Axis.1 [33.2%]
Axis.2 [24.1%]
Marital Status
Legally Divorced
Married
Not Married
Separated
LivingSituation
Alone
With parents,friends,other relatives
With spouse,partner and or kids
HighestEducation
Associate Degree
Bachelor's and Graduate Degree
High School

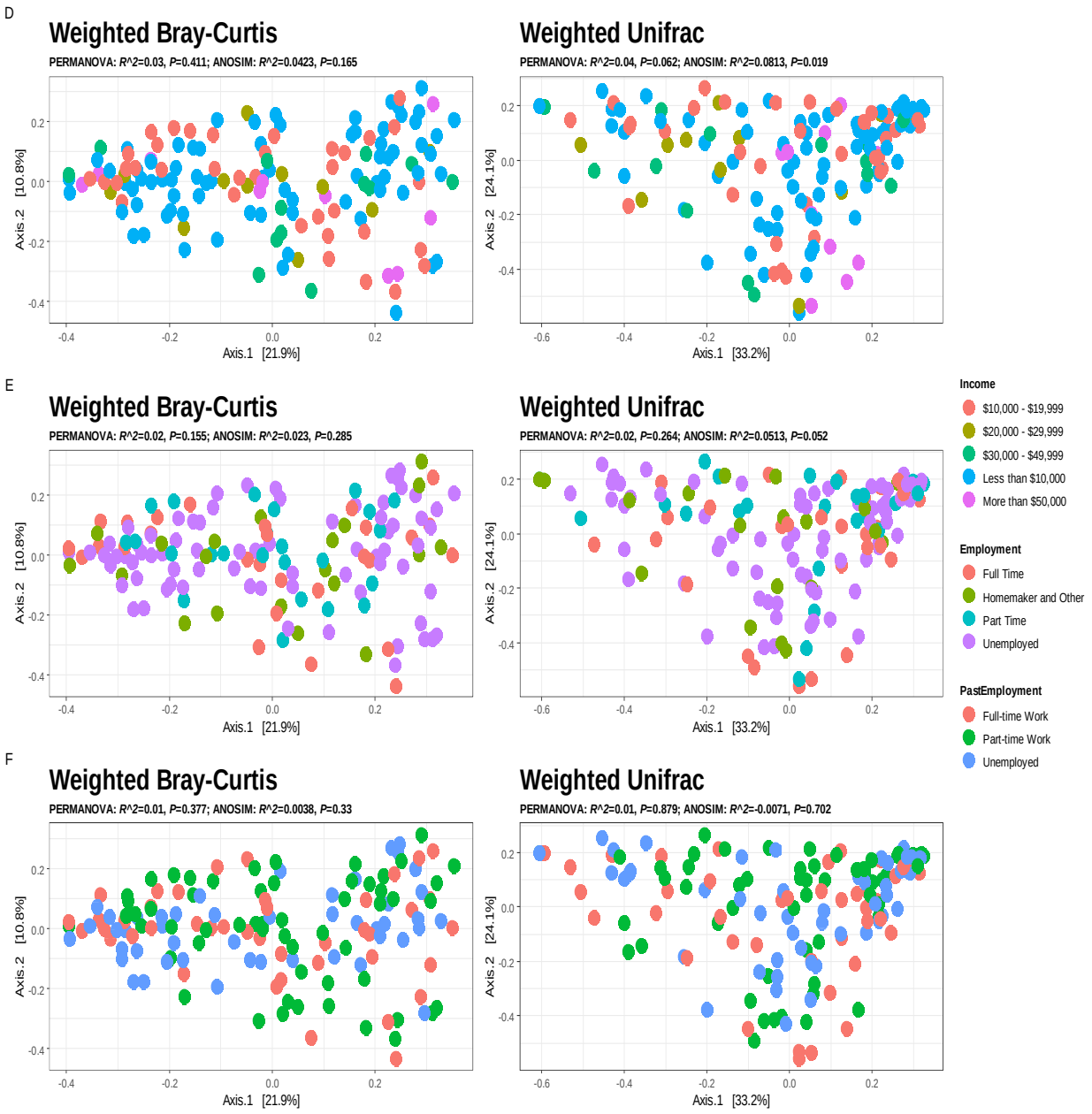


Figure 4-5. Beta diversity based on Weighted Bray-Curtis and Unifrac distances, revealing no clustering by sociodemographic factors in all study participants; A) Marital Status; B) Living Situation; C)Education; D) Annual Income; E) Current Empolyment; F) Employment in the past 3 years.

Table 4-2. Assosiation between TSS and Race ; summary of multivariable analyses

Model	Race (African American)			
	Coefficient (β)	Std.Error	95% CI	p- value
Model 1	-26.835	9.012	(-44.40, -9.26)	**0.003**
Model 2	-25.404	9.012	(-43.90, -6.87)	**0.005**
Model 3	-24.787	9.245	(-42.89,-6.67)	**0.008**

Race (reference group: Black or African American) as a predictor for taxonomic summary score (TSS)

Model 1: crude; Model 2: adjusted for age, sex; Model 3: Model 2 + smoking (TNE, NNAL, years of smoking)

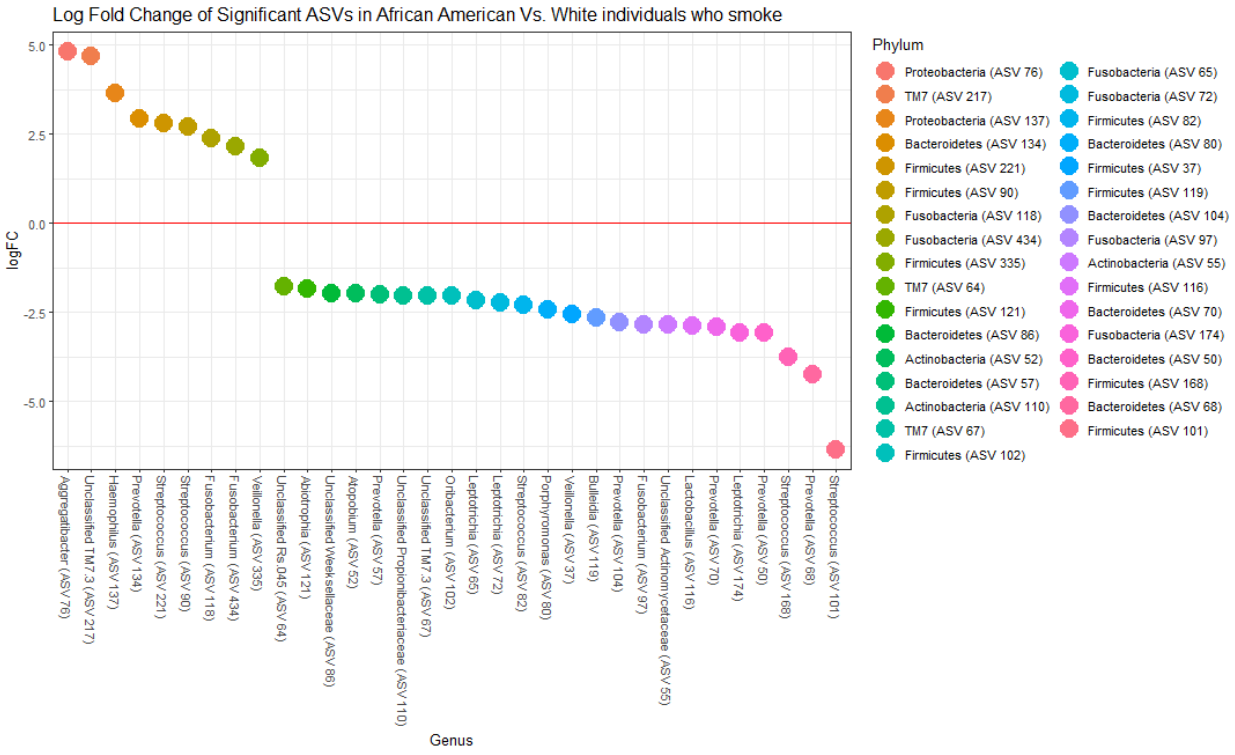


Figure 4-6. Differential taxa abundances; log2-fold changes for genus-level bacterial ASVs that were statistically significant at p≤0.001 (FDR-corrected) according to the pairwise test as implemented in edgeR. Each data point represents a genus-level ASVs (x-axis) identified as significantly different along with the

log2 fold change (y-axis). Positive log2-fold changes indicate increased abundance in AA when compared to WH individuals who smoke

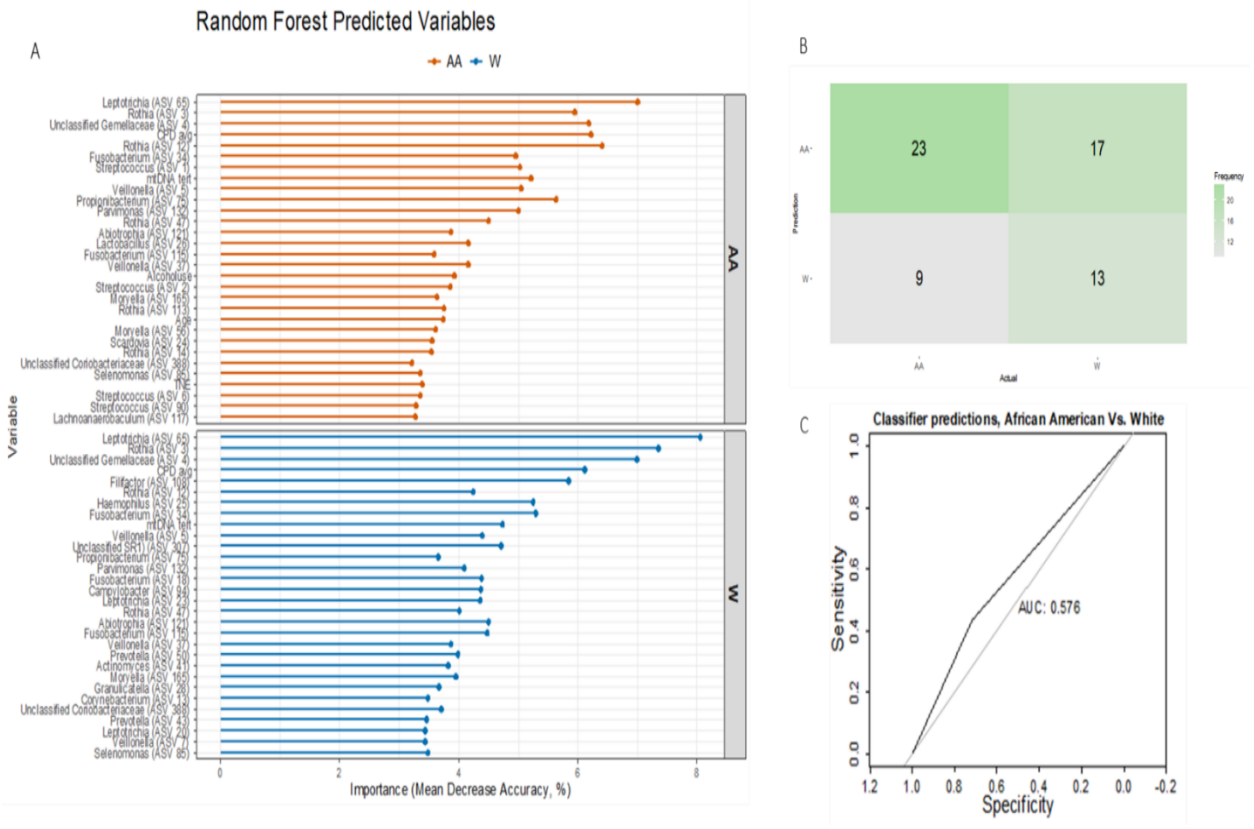


Figure 4-7. Random forest classification: A) Important features for each group based on the mean decrease accuracy are shown. B) Confusion matrix of model predictions C) AUROC results of the final model

SUPPLEMENTAL DATA FOR CHAPTER 4

Supplemental Table 4-1. Differently abundant taxa between AA and WH by Indicator Species Analysis

Taxa at genus-level ASV	Group	Iv*	p-value**
Unclassified Gemellaceae (ASV 4)	AA	0.601	0.001
Rothia (ASV 12)	AA	0.449	0.035
Haemophilus (ASV 25)	AA	0.38	0.002
Abiotrophia (ASV 121)	AA	0.373	0.003
Campylobacter (ASV 94)	AA	0.365	0.011
Granulicatella (ASV 28)	AA	0.36	0.007
Prevotella (ASV 57)	AA	0.36	0.008
Prevotella (ASV 50)	AA	0.353	0.001
Unclassified TM7.3 (ASV 67)	AA	0.316	0.014
Veillonella (ASV 37)	AA	0.315	0.015
Veillonella (ASV 32)	AA	0.314	0.046
Fusobacterium (ASV 115)	AA	0.305	0.003
Selenomonas (ASV 85)	AA	0.305	0.048
Leptotrichia (ASV 88)	AA	0.304	0.002
Campylobacter (ASV 53)	AA	0.303	0.032
Veillonella (ASV 39)	AA	0.296	0.022
Moryella (ASV 165)	AA	0.287	0.003
Filifactor (ASV 108)	AA	0.285	0.006
Tannerella (ASV 92)	AA	0.27	0.037
Atopobium (ASV 136)	AA	0.268	0.039
Leptotrichia (ASV 65)	AA	0.266	0.012
Peptostreptococcus (ASV 240)	AA	0.26	0.006
Unclassified Actinomycetaceae (ASV 130)	AA	0.25	0.006
TG5 (ASV 159)	AA	0.248	0.007
TG5 (ASV 172)	AA	0.244	0.005
Porphyromonas (ASV 80)	AA	0.241	0.003
Leptotrichia (ASV 72)	AA	0.228	0.026
Unclassified SR1 (ASV 307)	AA	0.224	0.001
Prevotella (ASV 128)	AA	0.217	0.006
Lautropia (ASV 160)	AA	0.216	0.039
Treponema (ASV 173)	AA	0.212	0.017
Actinomyces (ASV 207)	AA	0.207	0.011
Megasphaera (ASV 145)	AA	0.198	0.004
Oribacterium (ASV 351)	AA	0.197	0.012
Selenomonas (ASV 158)	AA	0.196	0.005
Parvimonas (ASV 273)	AA	0.184	0.041

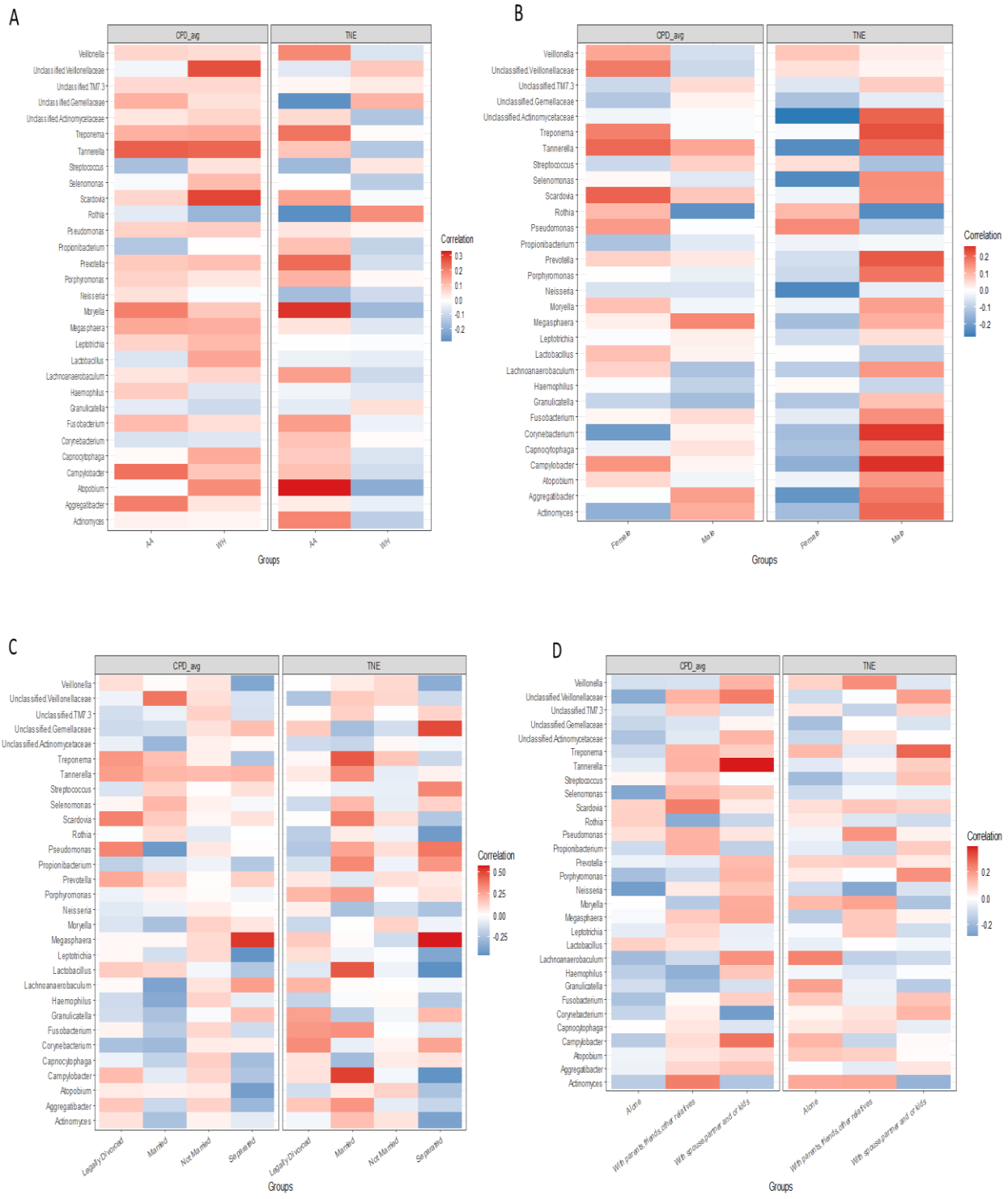
Actinomyces (ASV 319)	AA	0.183	0.031
Eikenella (ASV 271)	AA	0.177	0.026
Unclassified Acidaminobacteraceae (ASV 321)	AA	0.173	0.036
Aggregatibacter (ASV 291)	AA	0.163	0.003
Leptotrichia (ASV 230)	AA	0.145	0.037
Actinomyces (ASV 343)	AA	0.136	0.007
Unclassified Bacteroidales (ASV 494)	AA	0.133	0.023
Actinomyces (ASV 417)	AA	0.13	0.007
Actinomyces (ASV 402)	AA	0.129	0.013
Unclassified Lachnospiraceae.1 (ASV 324)	AA	0.125	0.019
Selenomonas (ASV 315)	AA	0.125	0.041
Unclassified Propionibacteriaceae (ASV 386)	AA	0.122	0.036
Prevotella (ASV 490)	AA	0.121	0.023
Tannerella (ASV 236)	AA	0.12	0.005
Pseudoramibacter_Eubacterium (ASV 481)	AA	0.119	0.043
Lachnoanaerobaculum (ASV 226)	AA	0.117	0.018
Prevotella (ASV 290)	AA	0.116	0.02
Leptotrichia (ASV 242)	AA	0.113	0.022
Selenomonas (ASV 393)	AA	0.112	0.043
Selenomonas (ASV 365)	AA	0.11	0.036
Unclassified TM7.3 (ASV 347)	AA	0.107	0.049
Leptotrichia (ASV 306)	AA	0.104	0.019
Fusobacterium (ASV 279)	AA	0.103	0.016
Atopobium (ASV 313)	AA	0.101	0.016
Treponema (ASV 535)	AA	0.099	0.012
Treponema (ASV 401)	AA	0.097	0.026
Unclassified Weeksellaceae (ASV 538)	AA	0.093	0.019
Prevotella (ASV 592)	AA	0.093	0.019
Unclassified Leptotrichiaceae (ASV 224)	AA	0.08	0.026
Unclassified Unassigned.1 (ASV 594)	AA	0.08	0.026
Campylobacter (ASV 575)	AA	0.08	0.027
Leptotrichia (ASV 423)	AA	0.08	0.029
Unclassified CW040 (ASV 512)	AA	0.08	0.032
Haemophilus (ASV 255)	AA	0.067	0.047
Rothia (ASV 3)	WH	0.652	0.003
Unclassified TM7.3 (ASV 217)	WH	0.096	0.04
Prevotella (ASV 525)	WH	0.07	0.02

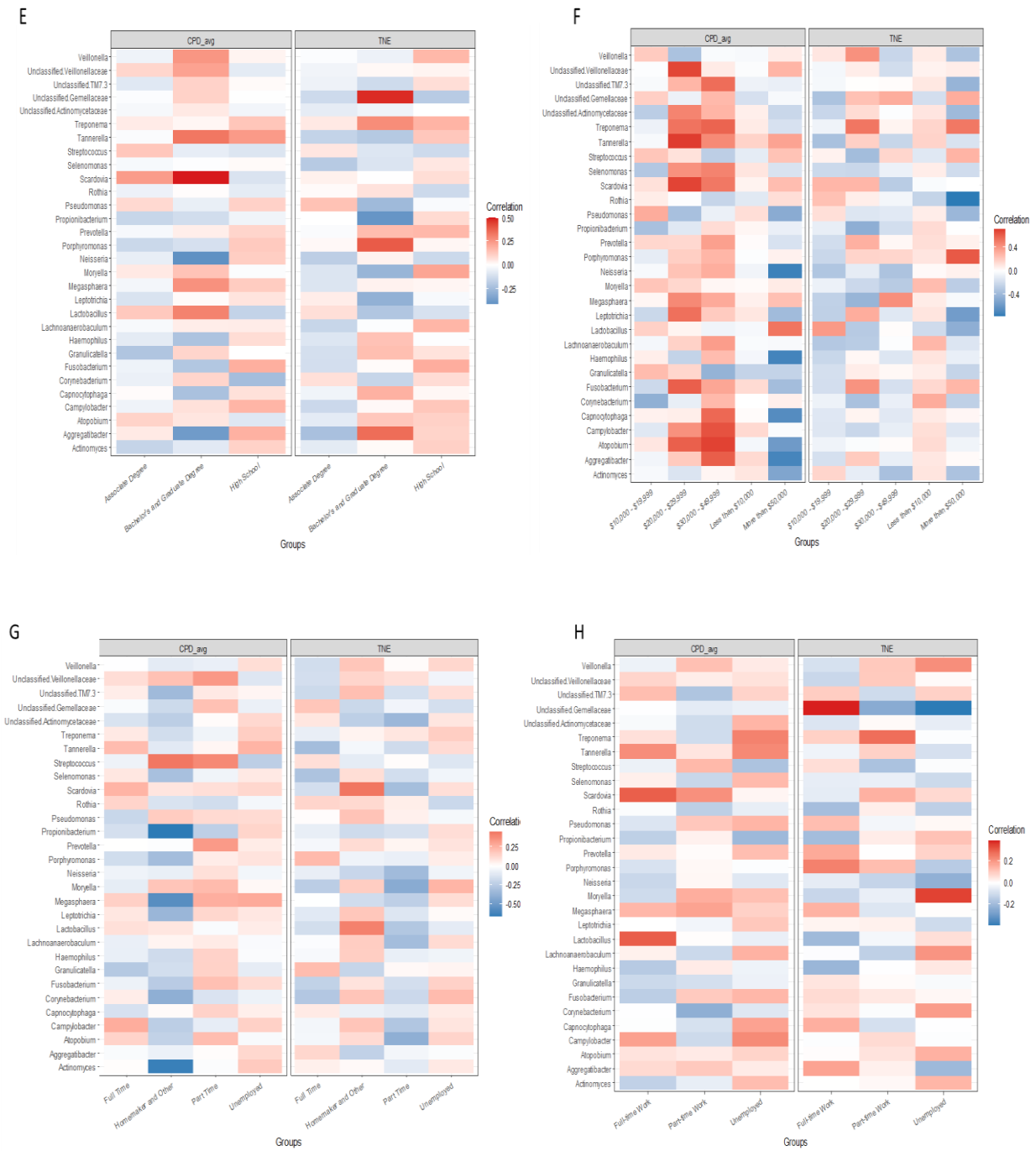
*Indicator value or IV value is a quantitative measure used to assess the degree to which a particular microbial taxon is associated with a specific group. The IV value combines both the abundance and specificity of a taxon within a given group. The IV value can range from 0 to 1, where higher values indicate a stronger association of the taxon with the specific group; **p-value of <0.05 was used to determine statistical significance. Here, the p-value corresponds to confidence in determining an indicator value, not a difference in means

Supplemental Table 4-2. Differently abundant taxa between AA and WH as assessed by differentialy abbundance taxa analysis

ASV level agglomeration	logFC	logCPM	PValue	FDR
Unclassified Streptococcus (ASV 101)	-6.3484	10.7969	8.48E-24	1.05E-21
Prevotella intermedia (ASV 68)	-4.2349	10.9933	5.40E-13	1.68E-11
Unclassified Streptococcus (ASV 168)	-3.7699	10.276	1.51E-10	3.12E-09
Prevotella intermedia (ASV 50)	-3.0823	11.1162	2.28E-09	3.14E-08
Unclassified Leptotrichia (ASV 174)	-3.0675	9.7081	5.65E-09	7.00E-08
Prevotella melaninogenica (ASV 70)	-2.9131	9.4874	1.03E-09	1.82E-08
Lactobacillus helveticus (ASV 116)	-2.8992	10.7187	4.48E-07	3.09E-06
Unclassified Actinomycetaceae (ASV 55)	-2.8477	11.1468	1.12E-08	1.27E-07
Unclassified Fusobacterium (ASV 97)	-2.8368	10.0552	1.93E-07	1.53E-06
Unclassified Prevotella (ASV 104)	-2.8002	9.8236	1.70E-09	2.64E-08
Bulleidia moorei (ASV 119)	-2.664	10.5624	6.98E-08	6.66E-07
Unclassified Veillonella (ASV 37)	-2.5614	11.9162	1.64E-06	1.07E-05
Unclassified Porphyromonas (ASV 80)	-2.423	10.4128	2.92E-06	1.81E-05
Streptococcus anginosus (ASV 82)	-2.2966	11.505	0.0001	0.0004
Unclassified Leptotrichia (ASV 72)	-2.2333	10.5765	9.39E-06	5.06E-05
Unclassified Leptotrichia (ASV 65)	-2.167	11.2567	6.46E-05	0.00028
Unclassified Oribacterium (ASV 102)	-2.054	10.1129	3.48E-06	2.05E-05
Unclassified TM7.3 (ASV 67)	-2.0339	11.3322	5.47E-05	0.00024
Unclassified Propionibacteriaceae (ASV 110)	-2.0297	10.0891	4.55E-05	0.00023
Unclassified Prevotella (ASV 57)	-1.9954	10.9629	1.63E-05	8.40E-05
Unclassified Atopobium (ASV 52)	-1.9858	11.5152	9.57E-05	0.0004
Unclassified Weeksellaceae (ASV 86)	-1.9681	11.0016	5.37E-05	0.00024
Unclassified Abiotrophia (ASV 121)	-1.8545	9.9917	5.12E-05	0.00024
Unclassified Rs.045 (ASV 64)	-1.7754	10.6892	0.00024	0.00091
Unclassified Veillonella (ASV 335)	1.835	8.8859	0.00017	0.00065
Unclassified Fusobacterium (ASV 434)	2.1646	7.6733	3.78E-07	2.76E-06
Unclassified Fusobacterium (ASV 118)	2.3937	10.1599	6.40E-06	3.61E-05
Streptococcus anginosus (ASV 90)	2.7111	11.4147	1.28E-07	1.13E-06
Unclassified Streptococcus (ASV 221)	2.8134	9.8167	1.98E-07	1.53E-06
Unclassified Prevotella (ASV 134)	2.9221	9.8565	4.57E-08	4.72E-07
Haemophilus parainfluenzae (ASV 137)	3.6636	10.0077	6.58E-11	1.63E-09
Unclassified TM7.3 (ASV 217)	4.7003	9.2349	1.86E-21	1.16E-19
Unclassified Aggregatibacter (ASV 76)	4.8263	11.9687	1.54E-13	6.39E-12

*log2-fold changes for genus-level bacterial ASVs that were statistically significant at p≤0.001 (FDR-corrected); Positive log2-fold changes indicate increased abundance in AA when compared to WH individuals who smoke


A
CPD_avg
TNE
Veillonella
Unclassified.Veillonellaceae
Unclassified.TM7.3
Unclassified.Gemellaceae
Unclassified.Actinomycetaceae
Treponema
Tannerella
Streptococcus
Selenomonas
Scardovia
Rothia
Pseudomonas
Propionibacterium
Prevotella
Porphyromonas
Neisseria
Moryella
Megasphaera
Leptotrichia
Lactobacillus
Lachnoanaerobaculum
Haemophilus
Granulicatella
Fusobacterium
Corynebacterium
Capnocytophaga
Campylobacter
Atopobium
Aggregatibacter
Actinomyces
AA
WH
Correlation
0.3
0.2
0.1
0.0
-0.1
-0.2
Groups
B
CPD_avg
TNE
Female
Male
Correlation
0.2
0.1
0.0
-0.1
-0.2
Groups
C
CPD_avg
TNE
Legally Divorced
Married
Not Married
Separated
Correlation
0.50
0.25
0.00
-0.25
Groups
D
CPD_avg
TNE
Alone
With parents,friends,other relatives
With spouse,partner and or kids
Correlation
0.2
0.0
-0.2
Groups



Supplemental Figure 4-1. Relationship between most abundant taxa and smoking dose (assessed as CPD or TNE) in participant subgroups by A) Race; B) Sex; C) Marital Status; D) Living Situation; E) Education; F) Income; G) Current Employment; H) Past Employment; none of the correlations had achieved statistical significance at threshold p-value for the Spearman correlation test set to 0.05 adjusting for multiple comparisons using Benjamin and Hochberg.

CHAPTER 5

Summary of findings, implications, and future directions

This dissertation aimed to explore the potential utility of the oral microbiome as a comprehensive and robust biological matrix for studies of tobacco product use and disease susceptibility by examining the responsiveness of the oral microbiome to changes in tobacco-related exposures; its association with the host biological effects in tobacco users; and the potential effect of sociodemographic stressors linked to increased susceptibility to tobacco-associated diseases on the oral microbiome of persons who smoke. Overall, these results showed that the oral microbiome could detect changes in tobacco exposure over time and capture biological effects indicative of increased risk for tobacco-related diseases.

Specific Aim 1 demonstrated that transitioning to exclusive use of e-cigarettes, a switch that was biochemically confirmed, over an 8-week period led to notable changes in the oral microbial composition and provided valuable insights into the potential usefulness of the microbiome as a receptor matrix for testing tobacco-associated exposures.

Results from Specific Aim 2 showed that specific taxa of the oral microbiome were associated with high levels of HPB-releasing DNA adducts in oral cells of persons who smoke. Many of such taxa contribute to nitrate metabolism and subsequently to endogenous nitrosamine formation, which is the potential mechanism by which the oral microbiome is contributing to tobacco-related cancer risks.

Lastly, Specific Aim 3 showed that sociodemographic factors may contribute to inter-individual differences in microbial composition and function among persons who smoke. The most significant findings were for self-identified race, potentially reflecting a cumulative effect of social and environmental factors on the oral microbiome.

Public Health Implications

The findings of this dissertation are multifaceted and could have significant public health implications. The oral microbiome could effectively detect changes in tobacco exposure over time, therefore, it could have role in monitoring and evaluating tobacco product use, especially in a marketplace

of diverse products, and serve as an asset in tobacco regulatory science. Additionally, by analyzing the oral microbiome individuals' tobacco use status could be assessed and the impact of interventions, such as transitioning to e-cigarettes, on reducing tobacco exposure could be evaluated . This information can inform tobacco cessation programs and policy-making efforts, providing further evidence for the inclusion of e-cigarettes as a tool for harm reduction in tobacco control strategies. The association between the oral microbiome and biological effects such as tobacco-induced DNA damage provides insights into the potential use of the microbiome as a marker for assessing disease susceptibility. Identifying these microbial markers can help in early detection, risk stratification, and personalized prevention strategies for individuals at higher risk for tobacco-related diseases. Lastly, understanding the impact of sociodemographic stressors on the oral microbiome in facilitating health inequities related to tobacco-associated diseases, particularly between African American and White individuals could be helpful in developing and implementing specific interventions aimed at promoting equitable tobacco-related health outcomes.

Future Directions

There are several avenues for further research that can be explored based on the research findings of this dissertation. Future studies should focus on mechanistic investigations to unravel the precise mechanisms by which tobacco exposure influences the oral microbiome and its subsequent impact on health outcomes. Integrating multi-omics approaches, such as metagenomics, metabolomics, and transcriptomics, will provide a comprehensive understanding of the biological processes involved in the tobacco-oral microbiome interaction. Also, it is crucial to conduct long-term longitudinal studies to provide insight into how the oral microbiome changes following tobacco exposure, as well as the stability and persistence of these changes. These studies will also help establish temporal associations between alterations in the oral microbiome and biological effects linked to tobacco-related diseases. Longitudinal studies should include diverse populations, with adequate representation of different racial and ethnic groups, to capture variations in microbial dynamics and health outcomes. Such research will provide evidence-based guidance for individuals considering e-cigarettes as a harm reduction strategy and inform regulatory policies regarding these products. Research in the future should focus on developing targeted intervention strategies using the oral microbiome as a modifiable factor. Utilizing probiotics, prebiotics, or other microbiome-based interventions may reduce the risk of tobacco-associated diseases. The

implementation of these strategies in diverse populations, taking sociodemographic factors into account, can contribute to reducing health inequities associated with tobacco use.

Bibliography:

1. State of Tobacco Control 2023, 2023

2.Islami, F., Goding Sauer, A., Miller, K. D., Siegel, R. L., Fedewa, S. A., Jacobs, E. J., McCullough, M. L., Patel, A. V., Ma, J., Soerjomataram, I., Flanders, W. D., Brawley, O. W., Gapstur, S. M., & Jemal, A. (2018). Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States: Potentially Preventable Cancers in US. CA: A Cancer Journal for Clinicians, 68(1), 31–54. https://doi.org/10.3322/caac.2144

3.Cornelius, M. E. (2022). Tobacco Product Use Among Adults—United States, 2020. MMWR. Morbidit and Mortality Weekly Report, 71. https://doi.org/10.15585/mmwr.mm7111a1

4.Colditz, G. A. (2015). The SAGE Encyclopedia of Cancer and Society. SAGE Publications, Inc. https://doi.org/10.4135/9781483345758

5.Perfetti, T.A., & Rodgman, A. (2013). The Chemical Components of Tobacco and Tobacco Smoke (2nd ed.). CRC Press. https://doi-org.ezp3.lib.umn.edu/10.1201/b13973

6.IARC (2004). Tobacco Smoke and Involuntary Smoking. https://publications.iarc.fr/Book-And-Report-Series/Iarc-Monographs-On-The-Identification-Of-Carcinogenic-Hazards-To-Humans/Tobacco-Smoke-And-Involuntary-Smoking-2004

7.Benowitz, N. L. (2010). Nicotine Addiction. The New England Journal of Medicine, 362(24), 2295–2303. https://doi.org/10.1056/NEJMra0809890

8.Govind, A. P., Vezina, P., & Green, W. N. (2009). Nicotine-induced Upregulation of Nicotinic Receptors: Underlying Mechanisms and Relevance to Nicotine Addiction. Biochemical Pharmacology, 78(7), 756–765. https://doi.org/10.1016/j.bcp.2009.06.011

9.Lee, W. H., Chen, H. M., Yang, S. F., Liang, C., Peng, C. Y., Lin, F. M., Tsai, L. L., Wu, B. C., Hsin, C. H., Chuang, C. Y., Yang, T., Yang, T. L., Ho, S. Y., Chen, W. L., Ueng, K. C., Huang, H. D., Huang, C. N., & Jong, Y. J. (2017). Bacterial alterations in salivary microbiota and their association in oral cancer. Scientific reports, 7(1), 16540. https://doi.org/10.1038/s41598-017-16418-x

10.Schmidt, B. L., Kuczynski, J., Bhattacharya, A., Huey, B., Corby, P. M., Queiroz, E. L., Nightingale, K., Kerr, A. R., DeLacure, M. D., Veeramachaneni, R., Olshen, A. B., & Albertson, D. G. (2014). Changes in abundance of oral microbiota associated with oral cancer. PloS one, 9(6), e98741. https://doi.org/10.1371/journal.pone.0098741

11.Peters, B. A., Wu, J., Pei, Z., Yang, L., Purdue, M. P., Freedman, N. D., Jacobs, E. J., Gapstur, S. M., Hayes, R. B., & Ahn, J. (2017). Oral Microbiome Composition Reflects Prospective Risk for Esophageal Cancers. Cancer research, 77(23), 6777–6787. https://doi.org/10.1158/0008-5472.CAN-17-1296

12.Yan, X., Yang, M., Liu, J., Gao, R., Hu, J., Li, J., Zhang, L., Shi, Y., Guo, H., Cheng, J., Razi, M., Pang, S., Yu, X., & Hu, S. (2015). Discovery and validation of potential bacterial biomarkers for lung cancer. American journal of cancer research, 5(10), 3111–3122.

13.Atarashi, K., Suda, W., Luo, C., Kawaguchi, T., Motoo, I., Narushima, S., Kiguchi, Y., Yasuma, K., Watanabe, E., Tanoue, T., Thaiss, C. A., Sato, M., Toyooka, K., Said, H. S., Yamagami, H., Rice, S. A., Gevers,

D., Johnson, R. C., Segre, J. A., Chen, K., … Honda, K. (2017). Ectopic colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation. Science (New York, N.Y.), 358(6361), 359–365. https://doi.org/10.1126/science.aan4526

14.Thomas, A. M., Manghi, P., Asnicar, F., Pasolli, E., Armanini, F., Zolfo, M., Beghini, F., Manara, S., Karcher, N., Pozzi, C., Gandini, S., Serrano, D., Tarallo, S., Francavilla, A., Gallo, G., Trompetto, M., Ferrero, G., Mizutani, S., Shiroma, H., Shiba, S., … Segata, N. (2019). Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nature medicine, 25(4), 667–678. https://doi.org/10.1038/s41591-019-0405-7

15.Wu, J., Peters, B. A., Dominianni, C., Zhang, Y., Pei, Z., Yang, L., Ma, Y., Purdue, M. P., Jacobs, E. J., Gapstur, S. M., Li, H., Alekseyenko, A. V., Hayes, R. B., & Ahn, J. (2016). Cigarette smoking and the oral microbiome in a large study of American adults. The ISME journal, 10(10), 2435–2446. https://doi.org/10.1038/ismej.2016.37

16.Camelo-Castillo, A. J., Mira, A., Pico, A., Nibali, L., Henderson, B., Donos, N., & Tomás, I. (2015). Subgingival microbiota in health compared to periodontitis and the influence of smoking. Frontiers in Microbiology, 6. https://doi.org/10.3389/fmicb.2015.00119

17.Olapade, O. A., & Ronk, A. J. (2015). Isolation, Characterization and Community Diversity of Indigenous Putative Toluene-Degrading Bacterial Populations with Catechol-2,3-Dioxygenase Genes in Contaminated Soils. Microbial Ecology, 69(1), 59–65. https://doi.org/10.1007/s00248-014-0466-6

18.Gao, S., Seo, J. S., Wang, J., Keum, Y. S., Li, J., & Li, Q. X. (2013). Multiple degradation pathways of phenanthrene by Stenotrophomonas maltophilia C6. International biodeterioration & biodegradation, 79, 98–104. https://doi.org/10.1016/j.ibiod.2013.01.012

19.Sato, N., Kakuta, M., Uchino, E., Hasegawa, T., Kojima, R., Kobayashi, W., Sawada, K., Tamura, Y., Tokuda, I., Imoto, S., Nakaji, S., Murashita, K., Yanagita, M., & Okuno, Y. (2020). The relationship between cigarette smoking and the tongue microbiome in an East Asian population. Journal of oral microbiology, 12(1), 1742527. https://doi.org/10.1080/20002297.2020.1742527

20.Goniewicz, M. L., Knysak, J., Gawron, M., Kosmider, L., Sobczak, A., Kurek, J., Prokopowicz, A., Jablonska-Czapla, M., Rosik-Dulewska, C., Havel, C., Jacob, P., 3rd, & Benowitz, N. (2014). Levels of selected carcinogens and toxicants in vapour from electronic cigarettes. *Tobacco control*, *23*(2), 133–139. https://doi.org/10.1136/tobaccocontrol-2012-050859

21.Goniewicz, M. L., Smith, D. M., Edwards, K. C., Blount, B. C., Caldwell, K. L., Feng, J., Wang, L., Christensen, C., Ambrose, B., Borek, N., van Bemmel, D., Konkel, K., Erives, G., Stanton, C. A., Lambert, E., Kimmel, H. L., Hatsukami, D., Hecht, S. S., Niaura, R. S., Travers, M., … Hyland, A. J. (2018). Comparison of Nicotine and Toxicant Exposure in Users of Electronic Cigarettes and Combustible Cigarettes. JAMA network open, 1(8), e185937. https://doi.org/10.1001/jamanetworkopen.2018.5937

22.Chopyk, J., Bojanowski, C. M., Shin, J., Moshensky, A., Fuentes, A. L., Bonde, S. S., Chuki, D., Pride, D. T., & Alexander, L. E. C. (2021). Compositional Differences in the Oral Microbiome of E-cigarette Users. Frontiers in Microbiology, 12. https://doi.org/10.3389/fmicb.2021.599664

23.Pushalkar, S., Paul, B., Li, Q., Yang, J., Vasconcelos, R., Makwana, S., González, J. M., Shah, S., Xie, C.,

Janal, M. N., Queiroz, E., Bederoff, M., Leinwand, J., Solarewicz, J., Xu, F., Aboseria, E., Guo, Y., Aguallo,

D., Gomez, C., … Saxena, D. (2020). Electronic Cigarette Aerosol Modulates the Oral Microbiome and

Increases Risk of Infection. *IScience*, *23*(3). https://doi.org/10.1016/j.isci.2020.10088

24.Centers for Disease Control and Prevention (US), National Center for Chronic Disease Prevention and Health Promotion (US), & Office on Smoking and Health (US). (2010). How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Centers for Disease Control and Prevention (US). http://www.ncbi.nlm.nih.gov/books/NBK53017/

25.Hecht S. S. (1998). Biochemistry, biology, and carcinogenicity of tobacco-specific N-nitrosamines. Chemical research in toxicology, 11(6), 559–603. https://doi.org/10.1021/tx980005

26.Coussens, L. M., & Werb, Z. (2002). Inflammation and cancer. Nature, 420(6917), 860–867. https://doi.org/10.1038/nature01322

27.Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Molecular Biology of the Cell. https://www.ncbi.nlm.nih.gov/books/NBK21054/

28.Masayesva, B. G., Mambo, E., Taylor, R. J., Goloubeva, O. G., Zhou, S., Cohen, Y., Minhas, K., Koch, W., Sciubba, J., Alberg, A. J., Sidransky, D., & Califano, J. (2006). Mitochondrial DNA content increase in response to cigarette smoking. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 15(1), 19–24. https://doi.org/10.1158/1055-9965.EPI-05-0210

29.Tan, D., Goerlitz, D. S., Dumitrescu, R. G., Han, D., Seillier-Moiseiwitsch, F., Spernak, S. M., Orden, R. A., Chen, J., Goldman, R., & Shields, P. G. (2008). Associations between cigarette smoking and mitochondrial DNA abnormalities in buccal cells. Carcinogenesis, 29(6), 1170–1177. https://doi.org/10.1093/carcin/bgn034

30.Hosgood, H. D., 3rd, Liu, C. S., Rothman, N., Weinstein, S. J., Bonner, M. R., Shen, M., Lim, U., Virtamo, J., Cheng, W. L., Albanes, D., & Lan, Q. (2010). Mitochondrial DNA copy number and lung cancer risk in a prospective cohort study. Carcinogenesis, 31(5), 847–849. https://doi.org/10.1093/carcin/bgq045

31.Bonner, M. R., Shen, M., Liu, C. S., Divita, M., He, X., & Lan, Q. (2009). Mitochondrial DNA content and lung cancer risk in Xuan Wei, China. Lung cancer (Amsterdam, Netherlands), 63(3), 331–334. https://doi.org/10.1016/j.lungcan.2008.06.012

32.Reznik, E., Miller, M. L., Şenbabaoğlu, Y., Riaz, N., Sarungbam, J., Tickoo, S. K., Al-Ahmadie, H. A., Lee, W., Seshan, V. E., Hakimi, A. A., & Sander, C. (2016). Mitochondrial DNA copy number variation across human cancers. eLife, 5, e10769. https://doi.org/10.7554/eLife.10769

33.Pfeifer GP, Denissenko MF, Olivier M, et al. (2002) Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. Oncogene. 21:74357451.

34.Stepanov, I., Upadhyaya, P., Carmella, S. G., Feuer, R., Jensen, J., Hatsukami, D. K., & Hecht, S. S. (2008). Extensive metabolic activation of the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in smokers. Cancer epidemiology, biomarkers & prevention : a publication of the

American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 17(7), 1764–1773. https://doi.org/10.1158/1055-9965.EPI-07-2844

35.Stepanov, I., Muzic, J., Le, C. T., Sebero, E., Villalta, P., Ma, B., Jensen, J., Hatsukami, D., & Hecht, S. S. (2013). Analysis of 4-hydroxy-1-(3-pyridyl)-1-butanone (HPB)-releasing DNA adducts in human exfoliated oral mucosa cells by liquid chromatography-electrospray ionization-tandem mass spectrometry. Chemical research in toxicology, 26(1), 37–45. https://doi.org/10.1021/tx300282k

36.Ma, B., Ruszczak, C., Jain, V., Khariwala, S. S., Lindgren, B., Hatsukami, D. K., & Stepanov, I. (2016). Optimized Liquid Chromatography Nanoelectrospray-High-Resolution Tandem Mass Spectrometry Method for the Analysis of 4-Hydroxy-1-(3-pyridyl)-1-butanone-Releasing DNA Adducts in Human Oral Cells. Chemical research in toxicology, 29(11), 1849–1856. https://doi.org/10.1021/acs.chemrestox.6b00254

37.Khariwala, S. S., Ma, B., Ruszczak, C., Carmella, S. G., Lindgren, B., Hatsukami, D. K., Hecht, S. S., & Stepanov, I. (2017). High Level of Tobacco Carcinogen-Derived DNA Damage in Oral Cells Is an Independent Predictor of Oral/Head and Neck Cancer Risk in Smokers. Cancer prevention research (Philadelphia, Pa.), 10(9), 507–513. https://doi.org/10.1158/1940-6207.CAPR-17-0140

38.Sharma, A. K., DeBusk, W. T., Stepanov, I., Gomez, A., & Khariwala, S. S. (2020). Oral Microbiome Profiling in Smokers with and without Head and Neck Cancer Reveals Variations Between Health and Disease. *Cancer Prevention Research*, *13*(5), 463–474. https://doi.org/10.1158/1940-6207.CAPR-19-0459

39.Yang, Y., Zheng, W., Cai, Q., Shrubsole, M. J., Pei, Z., Brucker, R., Steinwandel, M., Bordenstein, S. R., Li, Z., Blot, W. J., Shu, X.-O., & Long, J. (2019). Racial Differences in the Oral Microbiome: Data from Low-Income Populations of African Ancestry and European Ancestry. *MSystems*, *4*(6), e00639-19. https://doi.org/10.1128/mSystems.00639-19

40. Renson, A., Jones, H. E., Beghini, F., Segata, N., Zolnik, C. P., Usyk, M., Moody, T. U., Thorpe, L., Burk, R., Waldron, L., & Dowd, J. B. (2017). Sociodemographic patterning in the oral microbiome of a diverse sample of New Yorkers [Preprint]. Epidemiology. https://doi.org/10.1101/189225

41. Caliri, A. W., Tommasi, S., & Besaratinia, A. (2021). Relationships among smoking, oxidative stress, inflammation, macromolecular damage, and cancer. *Mutation Research*, *787*, 108365. https://doi.org/10.1016/j.mrrev.2021.108365

42. Hecht, S. S., & Hoffmann, D. (1988). Tobacco-specific nitrosamines, an important group of carcinogens in tobacco and tobacco smoke. Carcinogenesis, 9(6), 875–884. doi: 10.1093/carcin/9.6.875

43. Ashley, D. L., Oconnor, R. J., Bernert, J. T., Watson, C. H., Polzin, G. M., Jain, R. B., … Mccraw, J. M. (2010). Effect of Differing Levels of Tobacco-Specific Nitrosamines in Cigarette Smoke on the Levels of Biomarkers in Smokers. Cancer Epidemiology Biomarkers & Prevention, 19(6), 1389–1398. doi: 10.1158/1055-9965.epi-10-0084

44.Stepanov, I., Knezevich, A., Zhang, L., Watson, C. H., Hatsukami, D. K., & Hecht, S. S. (2011). Carcinogenic tobacco-specific N-nitrosamines in US cigarettes: three decades of remarkable neglect by the tobacco industry. Tobacco Control, 21(1), 44–48. doi: 10.1136/tc.2010.042192

45.Edwards, S. H., Rossiter, L. M., Taylor, K. M., Holman, M. R., Zhang, L., Ding, Y. S., & Watson, C. H. (2016). Tobacco-Specific Nitrosamines in the Tobacco and Mainstream Smoke of U.S. Commercial Cigarettes. Chemical Research in Toxicology, 30(2), 540–551. doi: 10.1021/acs.chemrestox.6b00268

46.IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, & International Agency for Research on Cancer (Eds.). (2010). Some non-heterocyclic polycyclic aromatic hydrocarbons and some related occupational exposures. IARC Press ; Distributed by World Health Organization.

47.Feng Z, Hu W, Hu Y, Tang MS. (2006) Acrolein is a major cigarette-related lung cancer agent. Preferential binding at p53 mutational hotspots and inhibition of DNA repair. Proc Natl Acad Sci USA. doi: 103:15404-15409

48.Lee, H.-W., Wang, H.-T., Weng, M.-W., Hu, Y., Chen, W.-S., Chou, D., … Tang, M.-S. (2014). Acrolein- and 4-Aminobiphenyl-DNA adducts in human bladder mucosa and tumor tissue and their mutagenicity in human urothelial cells. Oncotarget, 5(11). doi: 10.18632/oncotarget.1954

49.Zhang, S., Villalta, P. W., Wang, M., & Hecht, S. S. (2006). Analysis of Crotonaldehyde- and Acetaldehyde-Derived 1,N2-Propanodeoxyguanosine Adducts in DNA from Human Tissues Using Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry. Chemical Research in Toxicology, 19(10), 1386–1392. doi: 10.1021/tx060154d

50.List of Classifications, Agents classified by the IARC Monographs, Volumes 1–124". *IARC Monographs on the Evaluation of Risk to Humans. IARC. July 7, 2019*.

51.Goniewicz, M. L., Smith, D. M., Edwards, K. C., Blount, B. C., Caldwell, K. L., Feng, J., Wang, L., Christensen, C., Ambrose, B., Borek, N., van Bemmel, D., Konkel, K., Erives, G., Stanton, C. A., Lambert, E., Kimmel, H. L., Hatsukami, D., Hecht, S. S., Niaura, R. S., Travers, M., … Hyland, A. J. (2018). Comparison of Nicotine and Toxicant Exposure in Users of Electronic Cigarettes and Combustible Cigarettes. JAMA network open, 1(8), e185937. https://doi.org/10.1001/jamanetworkopen.2018.5937

52.National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Committee on the Review of the Health Effects of Electronic Nicotine Delivery Systems, Eaton, D. L., Kwan, L. Y., & Stratton, K. (Eds.). (2018). Public Health Consequences of E-Cigarettes. National Academies Press (US).

53.Centers for Disease Control and Prevention (US), National Center for Chronic Disease Prevention and Health Promotion (US), & Office on Smoking and Health (US). (2010). How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Centers for Disease Control and Prevention (US). http://www.ncbi.nlm.nih.gov/books/NBK53017/

54.Murphy, S. E., Park, S. L., Balbo, S., Haiman, C. A., Hatsukami, D. K., Patel, Y., Peterson, L. A., Stepanov, I., Stram, D. O., Tretyakova, N., Hecht, S. S., & Le Marchand, L. (2018). Tobacco biomarkers and genetic/epigenetic analysis to investigate ethnic/racial differences in lung cancer risk among smokers. NPJ precision oncology, 2, 17. https://doi.org/10.1038/s41698-018-0057-y

55.Carmella, S. G., Chen, M., Han, S., Briggs, A., Jensen, J., Hatsukami, D. K., & Hecht, S. S. (2009). Effects of Smoking Cessation on Eight Urinary Tobacco Carcinogen and Toxicant Biomarkers. *Chemical Research in Toxicology*, *22*(4), 734. https://doi.org/10.1021/tx800479s

56.Yuan, J. M., Koh, W. P., Murphy, S. E., Fan, Y., Wang, R., Carmella, S. G., Han, S., Wickham, K., Gao, Y. T., Yu, M. C., & Hecht, S. S. (2009). Urinary levels of tobacco-specific nitrosamine metabolites in relation

to lung cancer development in two prospective cohorts of cigarette smokers. Cancer research, 69(7), 2990–2995. https://doi.org/10.1158/0008-5472.CAN-08-4330

57.Stepanov, I., Sebero, E., Wang, R., Gao, Y. T., Hecht, S. S., & Yuan, J. M. (2014). Tobacco-specific N-nitrosamine exposures and cancer risk in the Shanghai Cohort Study: remarkable coherence with rat tumor sites. International journal of cancer, 134(10), 2278–2283. https://doi.org/10.1002/ijc.28575

58.Park, S. L., Tiirikainen, M. I., Patel, Y. M., Wilkens, L. R., Stram, D. O., Le Marchand, L., & Murphy, S. E. (2016). Genetic determinants of CYP2A6 activity across racial/ethnic groups with different risks of lung cancer and effect on their smoking intensity. Carcinogenesis, 37(3), 269–279. https://doi.org/10.1093/carcin/bgw012

59.Park, S. L., Carmella, S. G., Chen, M., Patel, Y., Stram, D. O., Haiman, C. A., Marchand, L. L., & Hecht, S. S. (2015). Mercapturic Acids Derived from the Toxicants Acrolein and Crotonaldehyde in the Urine of Cigarette Smokers from Five Ethnic Groups with Differing Risks for Lung Cancer. PLOS ONE, 10(6), e0124841. https://doi.org/10.1371/journal.pone.0124841

60.Conklin, D. J., Ogunwale, M. A., Chen, Y., Theis, W. S., Nantz, M. H., Fu, X.-A., Chen, L.-C., Riggs, D. W., Lorkiewicz, P., Bhatnagar, A., & Srivastava, S. (2018). Electronic cigarette-generated aldehydes: The contribution of e-liquid components to their formation and the use of urinary aldehyde metabolites as biomarkers of exposure. Aerosol Science and Technology, 52(11), 1219–1232. https://doi.org/10.1080/02786826.2018.1500013

61.Li D, Sundar IK, McIntosh S, Ossip DJ, Goniewicz ML, O'Connor RJ, et al. Association of smoking and electronic cigarette use with wheezing and related respiratory symptoms in adults: cross-sectional results from the Population Assessment of Tobacco and Health (PATH) study, wave 2. Tob Control. 2019 Feb 13;tobaccocontrol-2018-054694.

62. Xie Z, Ossip DJ, Rahman I, Li D. Use of Electronic Cigarettes and Self-Reported Chronic Obstructive Pulmonary Disease Diagnosis in Adults. Nicotine & Tobacco Research. 2020 Jun 12;22(7):1155–61

63.Ganesan, S. M., Dabdoub, S. M., Nagaraja, H. N., Scott, M. L., Pamulapati, S., Berman, M. L., Shields, P. G., Wewers, M. E., & Kumar, P. S. (2020). Adverse effects of electronic cigarettes on the disease-naive oral microbiome. Science advances, 6(22), eaaz0108. https://doi.org/10.1126/sciadv.aaz0108\

64.Pushalkar, S., Paul, B., Li, Q., Yang, J., Vasconcelos, R., Makwana, S., González, J. M., Shah, S., Xie, C., Janal, M. N., Queiroz, E., Bederoff, M., Leinwand, J., Solarewicz, J., Xu, F., Aboseria, E., Guo, Y., Aguallo, D., Gomez, C., Kamer, A., … Saxena, D. (2020). Electronic Cigarette Aerosol Modulates the Oral Microbiome and Increases Risk of Infection. iScience, 23(3), 100884. https://doi.org/10.1016/j.isci.2020.10088442.42.

65.Chen, M., Carmella, S. G., Sipe, C., Jensen, J., Luo, X., Le, C. T., Murphy, S. E., Benowitz, N. L., McClernon, F. J., Vandrey, R., Allen, S. S., Denlinger-Apte, R., Cinciripini, P. M., Strasser, A. A., al'Absi, M., Robinson, J. D., Donny, E. C., Hatsukami, D., & Hecht, S. S. (2019). Longitudinal stability in cigarette smokers of urinary biomarkers of exposure to the toxicants acrylonitrile and acrolein. *PLoS ONE*, *14*(1), e0210104. https://doi.org/10.1371/journal.pone.0210104

66.Luo, X., Carmella, S. G., Chen, M., Jensen, J. A., Wilkens, L. R., Le Marchand, L., Hatsukami, D. K., Murphy, S. E., & Hecht, S. S. (2020). Urinary Cyanoethyl Mercapturic Acid, a Biomarker of the Smoke Toxicant Acrylonitrile, Clearly Distinguishes Smokers From Nonsmokers. Nicotine & tobacco research : official journal

of the Society for Research on Nicotine and Tobacco, 22(10), 1744–1747. https://doi.org/10.1093/ntr/ntaa080

67.Murphy SE, Park SS, Thompson EF, et al. Nicotine N-glucuronidation relative to N-oxidation and C-oxidation and UGT2B10 genotype in five ethnic/racial groups. *Carcinogenesis.* 2014;35(11):2526-2533.

68.von Weymarn, L. B., Lu, X., Thomson, N. M., LeMarchand, L., Park, S. L., & Murphy, S. E. (2023). Quantitation of Ten Urinary Nicotine Metabolites, Including 4-Hydroxy-4-(3-pyridyl) Butanoic Acid, a Product of Nicotine 2'-Oxidation, and CYP2A6 Activity in Japanese Americans, Native Hawaiians, and Whites. Chemical research in toxicology, 36(2), 313–321.

69.Chang, C. M., Edwards, S. H., Arab, A., Del Valle-Pinero, A. Y., Yang, L., & Hatsukami, D. K. (2017). Biomarkers of Tobacco Exposure: Summary of an FDA-Sponsored Public Workshop. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 26(3), 291–302. https://doi.org/10.1158/1055-9965.EPI-16-0675

70.R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. Available: https://www.R-project.org/.

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

72.Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB et al. vegan: community ecology package. R package version 2011;117–118.

73.Love MI, Huber W, Anders S (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." *Genome Biology*, **15**, 550. doi:10.1186/s13059-014-0550-8.

74.Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis* (2nd ed.) [PDF]. Springer International Publishing.

75.Guo, J., Ikuemonisan, J., Hatsukami, D. K., & Hecht, S. S. (2021). Liquid chromatography-nanoelectrospray ionization-high resolution tandem mass spectrometry analysis of apurinic/apyrimidinic sites in oral cell DNA of cigarette smokers, e-cigarette users and nonsmokers. *Chemical Research in Toxicology*, *34*(12), 2540. https://doi.org/10.1021/acs.chemrestox.1c00308

76.Nath, S., Pulikkotil, S. J., Weyrich, L., Zilm, P., Kapellas, K., & Jamieson, L. (2022). Effect of Periodontal Interventions on Characteristics of the Periodontal Microbial Profile: A Systematic Review and Meta-Analysis. *Microorganisms*, *10*(8). https://doi.org/10.3390/microorganisms10081582

77.Yang, I., Rodriguez, J., Wright, C. Y., & Hu, Y.-J. (2023). Oral microbiome of electronic cigarette users: A cross-sectional exploration. *Oral Diseases*, *29*(4), 1875–1884. https://doi.org/10.1111/odi.14186

78.Thomas, S. C., Xu, F., Pushalkar, S., Lin, Z., Thakor, N., Vardhan, M., Flaminio, Z., Khodadadi-Jamayran, A., Vasconcelos, R., Akapo, A., Queiroz, E., Bederoff, M., Janal, M. N., Guo, Y., Aguallo, D., Gordon, T., Corby, P. M., Kamer, A. R., Li, X., & Saxena, D. (2022). Electronic Cigarette Use Promotes a Unique Periodontal Microbiome. *MBio*. https://doi.org/10.1128/mbio.00075-22

79.Fangxi Xu, Smruti Pushalkar, Zi-Yan Lin, Scott Thomas, Julia Kishanie Persaud, Maria A. Sierra, Mridula Vardhan, Rebeca Vasconcelos, Adenike Akapo, Yuqi Guo, Terry Gordon, Patricia Corby, Angela R. Kamer, Xin Li, & Deepak Saxena. (2022). Electronic Cigarette Use Enriches Periodontal Pathogens. *Molecular Oral Microbiology*. https://doi.org/10.1111/omi.12361

80.Bagale Kamal, Paudel Santosh, Cagle Hayden, Sigel Erin, & Kulkarni Ritwij. (2020). Electronic Cigarette (E-Cigarette) Vapor Exposure Alters the Streptococcus pneumoniae Transcriptome in a Nicotine-Dependent Manner without Affecting Pneumococcal Virulence. *Applied and Environmental Microbiology*, *86*(3), e02125-19. https://doi.org/10.1128/AEM.02125-19

81.Brennan, P., Hainaut, P., & Boffetta, P. (2011). Genetics of lung-cancer susceptibility. *The Lancet Oncology*, *12*(4), 399–408. https://doi.org/10.1016/S1470-2045(10)70126-1

82.Harman, S. M., Liang, L., Tsitouras, P. D., Gucciardo, F., Heward, C. B., Reaven, P. D., Ping, W., Ahmed, A., & Cutler, R. G. (2003). Urinary excretion of three nucleic acid oxidation adducts and isoprostane F2α measured by liquid chromatography-mass spectrometry in smokers, ex-smokers, and nonsmokers. *Free Radical Biology and Medicine*, *35*(10), 1301–1309. https://doi.org/10.1016/j.freeradbiomed.2003.07.003

83.Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., & Darnell, J. (2000). Molecular Cell Biology (4th ed.). W. H. Freeman.

84.Ma, B., Ruszczak, C., Jain, V., Khariwala, S. S., Lindgren, B., Hatsukami, D. K., & Stepanov, I. (2016). Optimized Liquid Chromatography Nanoelectrospray-High-Resolution Tandem Mass Spectrometry Method for the Analysis of 4-Hydroxy-1-(3-pyridyl)-1-butanone-Releasing DNA Adducts in Human Oral Cells. Chemical research in toxicology, 29(11), 1849–1856. https://doi.org/10.1021/acs.chemrestox.6b00254

85.Balbo, S., & Brooks, P. J. (2015). Implications of Acetaldehyde-Derived DNA Adducts for Understanding Alcohol-Related Carcinogenesis. In *Biological Basis of Alcohol-Induced Cancer* (pp. 71–88). Springer, Cham. https://doi.org/10.1007/978-3-319-09614-8_5

86.Freire, M., Nelson, K. E., & Edlund, A. (2021). The Oral Host–Microbial Interactome: An Ecological Chronometer of Health? *Trends in Microbiology*, *29*(6), 551–561. https://doi.org/10.1016/j.tim.2020.11.004

87.Doel, J. J., Benjamin, N., Hector, M. P., Rogers, M., & Allaker, R. P. (2005). Evaluation of bacterial nitrate reduction in the human oral cavity. European journal of oral sciences, 113(1), 14–19. https://doi.org/10.1111/j.1600-0722.2004.00184.x

88.Stepanov, I., Carmella, S. G., Han, S., Pinto, A., Strasser, A. A., Lerman, C., & Hecht, S. S. (2009). Evidence for endogenous formation of N'-nitrosonornicotine in some long-term nicotine patch users. Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco, 11(1), 99–105. https://doi.org/10.1093/ntr/ntn004

89.Stepanov, I., Carmella, S. G., Briggs, A., Hertsgaard, L., Lindgren, B., Hatsukami, D., & Hecht, S. S. (2009). Presence of the carcinogen N'-nitrosonornicotine in the urine of some users of oral nicotine replacement therapy products. Cancer research, 69(21), 8236–8240. https://doi.org/10.1158/0008-5472.CAN-09-1084

90.Bustamante, G., Ma, B., Yakovlev, G., Yershova, K., Le, C., Jensen, J., Hatsukami, D. K., & Stepanov, I. (2018b). Presence of the Carcinogen N′-Nitrosonornicotine in Saliva of E-cigarette Users. Chemical Research in Toxicology, 31(8), 731–738. https://doi.org/10.1021/acs.chemrestox.8b000898

91.Karunadharma, P. P., Nordgaard, C. L., Olsen, T. W., & Ferrington, D. A. (2010). Mitochondrial DNA damage as a potential mechanism for age-related macular degeneration. Investigative ophthalmology & visual science, 51(11), 5470–5479. https://doi.org/10.1167/iovs.10-5429

92.Murphy SE, Park SS, Thompson EF, et al. Nicotine N-glucuronidation relative to N-oxidation and C-oxidation and UGT2B10 genotype in five ethnic/racial groups. *Carcinogenesis.* 2014;35(11):2526-2533

93.Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R. and Beiko, R.G. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature Biotechnology, 31(9), pp.814-821. doi:10.1038/nbt.2676

94.Consortium, T. H. M. P. (2012). Structure, Function and Diversity of the Healthy Human Microbiome. *Nature*, *486*(7402), 207. https://doi.org/10.1038/nature11234

95.Borbolis, F., Mytilinaiou, E., & Palikaras, K. (2023). The Crosstalk between Microbiome and Mitochondrial Homeostasis in Neurodegeneration. *Cells*, *12*(3), Article 3. https://doi.org/10.3390/cells12030429

96.Nearing, J. T., DeClercq, V., Limbergen, J. V., & Langille, M. G. I. (2020). Assessing the Variation within the Oral Microbiome of Healthy Adults. *MSphere*, *5*(5). https://doi.org/10.1128/mSphere.00451-20

97.Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer statistics, 2022. *CA: A Cancer Journal for Clinicians*, *72*(1), 7–33. https://doi.org/10.3322/caac.21708

98.*Monograph 22. A Socioecological Approach to Addressing Tobacco-Related Health Disparities*. (n.d.).

99.Zhang, Y.-B., Chen, C., Pan, X.-F., Guo, J., Li, Y., Franco, O. H., Liu, G., & Pan, A. (2021). Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: Two prospective cohort studies. *BMJ*, *373*. https://doi.org/10.1136/bmj.n604

100.Harvanek, Z. M., Fogelman, N., Xu, K., & Sinha, R. (2021). Psychological and biological resilience modulates the effects of stress on epigenetic aging. *Translational Psychiatry*, *11*(1), Article 1. https://doi.org/10.1038/s41398-021-01735-7

101.National Academies of Sciences, E., Division, H. and M., Practice, B. on P. H. and P. H., States, C. on C.-B. S. to P. H. E. in the U., Baciu, A., Negussie, Y., Geller, A., & Weinstein, J. N. (2017). The State of Health Disparities in the United States. In *Communities in Action: Pathways to Health Equity*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK425844/

102.Montez, J. K., & Berkman, L. F. (2014). Trends in the Educational Gradient of Mortality Among US Adults Aged 45 to 84 Years: Bringing Regional Context Into the Explanation. *American Journal of Public Health*, *104*(1), e82. https://doi.org/10.2105/AJPH.2013.301526

103.Hayward, M.D. and Farina, M.P. (2023), Dynamic Changes in the Association Between Education and Health in the United States. Milbank Quarterly., 101: 396-418. https://doi.org/10.1111/1468-0009.12611

104.Bor, J., Cohen, G.H., & Galea, S. (2017). Population health in an era of rising income inequality: USA, 1980–2015. *The Lancet, 389*, 1475-1490.

105.Stram, D. O., Park, S. L., Haiman, C. A., Murphy, S. E., Patel, Y., Hecht, S. S., & Le Marchand, L. (2019). Racial/Ethnic Differences in Lung Cancer Incidence in the Multiethnic Cohort Study: An Update. *JNCI: Journal of the National Cancer Institute*, *111*(8), 811–819. https://doi.org/10.1093/jnci/djy206

106.Haiman, C. A., Stram, D. O., Wilkens, L. R., Pike, M. C., Kolonel, L. N., Henderson, B. E., & Le Marchand, L. (2006). Ethnic and racial differences in the smoking-related risk of lung cancer. *The New England Journal of Medicine*, *354*(4), 333–342. https://doi.org/10.1056/NEJMoa033250

107.Siegel, R., Ward, E., Brawley, O., & Jemal, A. (2011). Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA: a cancer journal for clinicians*, *61*(4), 212–236. https://doi.org/10.3322/caac.2012

108.Conway, D. I., Brenner, D. R., McMahon, A. D., Macpherson, L. M. D., Agudo, A., Ahrens, W., Bosetti, C., Brenner, H., Castellsague, X., Chen, C., Curado, M. P., Curioni, O. A., Maso, L. D., Daudt, A. W., Filho, J. F. de G., D'Souza, G., Edefonti, V., Fabianova, E., Fernandez, L., … Brennan, P. (2015). Estimating and explaining the effect of education and income on head and neck cancer risk: INHANCE consortium pooled analysis of 31 case-control studies from 27 countries. *International Journal of Cancer. Journal International Du Cancer*, *136*(5), 1125. https://doi.org/10.1002/ijc.29063

109.Slopen, N., Kontos, E. Z., Ryff, C. D., Ayanian, J. Z., Albert, M. A., & Williams, D. R. (2013). Psychosocial stress and cigarette smoking persistence, cessation, and relapse over 9–10 years: A prospective study of middle-aged adults in the United States. *Cancer Causes & Control : CCC*, *24*(10), 1849. https://doi.org/10.1007/s10552-013-0262-5

110.Blekhman, R., Goodrich, J. K., Huang, K., Sun, Q., Bukowski, R., Bell, J. T., Spector, T. D., Keinan, A., Ley, R. E., Gevers, D., & Clark, A. G. (2015). Host genetic variation impacts microbiome composition across human body sites. *Genome Biology*, *16*(1), 191. https://doi.org/10.1186/s13059-015-0759-1

111.Gomez, A., Espinoza, J. L., Harkins, D. M., Leong, P., Saffery, R., Bockmann, M., Torralba, M., Kuelbs, C., Kodukula, R., Inman, J., Hughes, T., Craig, J. M., Highlander, S. K., Jones, M. B., Dupont, C. L., & Nelson, K. E. (2017). Host Genetic Control of the Oral Microbiome in Health and Disease. *Cell Host & Microbe*, *22*(3), 269-278.e3. https://doi.org/10.1016/j.chom.2017.08.013

112.Stahringer, S. S., Clemente, J. C., Corley, R. P., Hewitt, J., Knights, D., Walters, W. A., Knight, R., & Krauter, K. S. (2012). Nurture trumps nature in a longitudinal survey of salivary bacterial communities in twins from early adolescence to early adulthood. *Genome Research*, *22*(11), 2146–2152. https://doi.org/10.1101/gr.140608.112

113.Burcham, Z. M., Garneau, N. L., Comstock, S. S., Tucker, R. M., Knight, R., Metcalf, J. L., Genetics of Taste Lab Citizen Scientists, Miranda, A., Reinhart, B., Meyers, D., Woltkamp, D., Boxer, E., Hutchens, J., Kim, K., Archer, M., McAteer, M., Huss, P., Defonseka, R., Stahle, S., … Reusser, W. (2020). Patterns of Oral Microbiota Diversity in Adults and Children: A Crowdsourced Population Study. *Scientific Reports*, *10*(1), 2133. https://doi.org/10.1038/s41598-020-59016-0

114.Mukherjee, C., Moyer, C. O., Steinkamp, H. M., Hashmi, S. B., Beall, C. J., Guo, X., Ni, A., Leys, E. J., & Griffen, A. L. (2021). Acquisition of oral microbiota is driven by environment, not host genetics. *Microbiome*, *9*. https://doi.org/10.1186/s40168-020-00986-8

115.Mason, M. R., Nagaraja, H. N., Camerlengo, T., Joshi, V., & Kumar, P. S. (2013). Deep Sequencing Identifies Ethnicity-Specific Bacterial Signatures in the Oral Microbiome. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0077287

116.Marotz, C., Molinsky, R., Martino, C., Bohn, B., Roy, S., Rosenbaum, M., Desvarieux, M., Yuzefpolskaya, M., Paster, B. J., Jacobs, D. R., Colombo, P. C., Papapanou, P. N., Knight, R., & Demmer, R. T. (2022). Early microbial markers of periodontal and cardiometabolic diseases in ORIGINS. *Npj Biofilms and Microbiomes*, *8*(1), Article 1. https://doi.org/10.1038/s41522-022-00289-w

117.Sedghi, L. M., Bacino, M., & Kapila, Y. L. (2021). Periodontal Disease: The Good, The Bad, and The Unknown. *Frontiers in Cellular and Infection Microbiology*, *11*. https://doi.org/10.3389/fcimb.2021.766944

118.Teles, F. R. F., Alawi, F., Castilho, R. M., & Wang, Y. (2020). Association or Causation? Exploring the Oral Microbiome and Cancer Links. *Journal of Dental Research*, *99*(13), 1411–1424. https://doi.org/10.1177/0022034520945242

119.Roberts, D.W. (2016). labdsv: Ordination and Multivariate Analysis for Ecology. R package version 1.8-0. Available at: https://CRAN.R-project.org/package=labdsv

120. McDermaid et al. (2018): Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1), 139-140

121. Lahti, L., Shetty, S., and Blake, T. (2017). MicrobiomeSeq: An R package for microbial community analysis. MicrobiomeSeq. Available at: https://github.com/umerijaz/microbiomeSeq

122.Ruan, X., Luo, J., Zhang, P., & Howell, K. (2022). The salivary microbiome shows a high prevalence of core bacterial members yet variability across human populations. *Npj Biofilms and Microbiomes*, *8*(1), Article 1. https://doi.org/10.1038/s41522-022-00343-7

123.Alberti, M. O., Hindler, J. A., & Humphries, R. M. (2015). Antimicrobial Susceptibilities of Abiotrophia defectiva, Granulicatella adiacens, and Granulicatella elegans. *Antimicrobial Agents and Chemotherapy*. https://doi.org/10.1128/AAC.02645-15

124.Danilkowicz, R. M., Lachiewicz, A. M., Lorenzana, D. J., Barton, K. D., & Lachiewicz, P. F. (2021). Prosthetic Joint Infection After Dental Work: Is the Correct Prophylaxis Being Prescribed? A Systematic Review. *Arthroplasty Today*, *7*, 69–75. https://doi.org/10.1016/j.artd.2020.11.007

125.Elashery, A. R., Stratidis, J., & Patel, A. D. (2020). Double-Valve Heart Disease and Glomerulonephritis Consequent to Abiotrophia defectiva Endocarditis. *Texas Heart Institute Journal*, *47*(1), 35–37. https://doi.org/10.14503/THIJ-17-6575

126.Maki, K. A., Ganesan, S. M., Meeks, B., Farmer, N., Kazmi, N., Barb, J. J., Joseph, P. V., & Wallen, G. R. (2022). The role of the oral microbiome in smoking-related cardiovascular risk: A review of the literature exploring mechanisms and pathways. *Journal of Translational Medicine*, *20*(1), Article 1. https://doi.org/10.1186/s12967-022-03785-x

127.Filho, W. S. e S., Casarin, R. C. V., Junior, E. L. N., Passos, H. M., Sallum, A. W., & Gonçalves, R. B. (2014). Microbial Diversity Similarities in Periodontal Pockets and Atheromatous Plaques of Cardiovascular Disease Patients. *PLOS ONE*, *9*(10), e109761. https://doi.org/10.1371/journal.pone.0109761

128.Eribe, E. R. K., & Olsen, I. (2017). Leptotrichia species in human infections II. *Journal of Oral Microbiology*, *9*(1), 1368848. https://doi.org/10.1080/20002297.2017.1368848

129.Mougeot, J.-L. C., Beckman, M. F., Langdon, H. C., Lalla, R. V., Brennan, M. T., & Bahrani Mougeot, F. K. (2022). Haemophilus pittmaniae and Leptotrichia spp. Constitute a Multi-Marker Signature in a Cohort of Human Papillomavirus-Positive Head and Neck Cancer Patients. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.794546

130.Pietiäinen, M., Liljestrand, J. M., Kopra, E., & Pussinen, P. J. (2018). Mediators between oral dysbiosis and cardiovascular diseases. *European Journal of Oral Sciences*, *126*, 26–36. https://doi.org/10.1111/eos.12423

131.Goh, C. E., Bohn, B., Marotz, C., Molinsky, R., Roy, S., Paster, B. J., Chen, C.-Y., Rosenbaum, M., Yuzefpolskaya, M., Colombo, P. C., Desvarieux, M., Papapanou, P. N., David R. Jacobs, J., Knight, R., & Demmer, R. T. (2022). Nitrite Generating and Depleting Capacity of the Oral Microbiome and Cardiometabolic Risk: Results from ORIGINS. *Journal of the American Heart Association: Cardiovascular and Cerebrovascular Disease*, *11*(10). https://doi.org/10.1161/JAHA.121.023038

132.Belstrøm, D., Holmstrup, P., Nielsen, C. H., Kirkby, N., Twetman, S., Heitmann, B. L., Klepac-Ceraj, V., Paster, B. J., & Fiehn, N.-E. (2014). Bacterial profiles of saliva in relation to diet, lifestyle factors, and socioeconomic status. *Journal of Oral Microbiology*, *6*. https://doi.org/10.3402/jom.v6.23609

133.Santonocito, S., Giudice, A., Polizzi, A., Troiano, G., Merlo, E. M., Sclafani, R., Grosso, G., & Isola, G. (2022). A Cross-Talk between Diet and the Oral Microbiome: Balance of Nutrition on Inflammation and Immune System's Response during Periodontitis. *Nutrients*, *14*(12). https://doi.org/10.3390/nu14122426

134.Naorungroj, S., Slade, G. D., Divaris, K., Heiss, G., Offenbacher, S., & Beck, J. D. (2017). Racial Differences in Periodontal Disease and 10-Year Self-Reported Tooth Loss among Late Middle-Aged and Older Adults: The Dental ARIC Study. *Journal of Public Health Dentistry*, *77*(4), 372. https://doi.org/10.1111/jphd.12226

Appendix I History of Tobacco Use Questionnaire

Subject #☐☐☐☐	Subject Initials ☐☐☐	Date ☐☐/☐☐/☐☐

A. Demographics

A1. Date of Birth ☐☐/☐☐/☐☐☐☐

month day year

A2. Age ☐☐ years

A3. Gender ◯ 1. Male ◯ 2. Female

A4. a. Race ☐

1. Black or African American
2. American Indian, Alaskan Native
3. Asian/Pacific Islander
4. More than one race_____(Specify)
5. White (Caucasian)
6. Other _____(Specify)

a1. Are all 4 of your grandparents of the same race as identified for yourself above?

◯ 1. Yes ◯ 2. No

b. Ethnicity ☐

1. Hispanic
2. Non-Hispanic

A5. Marital Status ☐

1. Never Married

A8. Education ☐

Indicate highest level completed.

1. 8th grade or less
2. Some High School
3. High School Graduate/Equivalent
4. Some College/2-year Degree
5. College Graduate/4-year Degree
6. Graduate - Professional Degree

A9. Annual Personal Income ☐

(from all sources)

1. Less than $10,000
2. $10,000 - $19,999
3. $20,000 - $29,999
4. $30,000 - $49,999
5. $50,000 - $69,999
6. $70,000 - $99,999
7. $100,000 or more

A10. Occupation : Job description

_____

_____

2. Married for the first time
3. Remarried
4. Separated
5. Divorced
6. Widowed

A6. Current Living Situation ☐

1. Alone
2. With spouse/partner and/or kids
3. With parents
4. With friends/other relatives
5. Other: _____ (Specify)

A7. Children Currently Living With You

☐☐ Indicate in the boxes the total number of children currently living with you. Include step-children and adopted children.

_____

_____

☐ for office use

A11. Current Employment Status ☐

1. Regular full-time work
2. Part-time work
3. Homemaker
4. Unemployed
5. Other:_____

A12. Employment: Last 3 Years ☐

1. Employed virtually all of the time
2. Employed at least half of the time
3. Employed less than half of the time
4. Not employed at all

Subject # ☐☐☐☐ **Subject Initials** ☐☐☐

B. Current Use of Cigarettes and Other Tobacco Products

B1. How many cigarettes per day do you smoke? ☐☐☐ per day

B2. How long have you smoked at this rate? ☐☐ years

or check here ☐ if less than 1 year

B3. Have you ever smoked more than your current rate of smoking?

☐ 1. Yes

☐ 2. No *(go to Question B4)*

a. When you were smoking the heaviest, how many cigarettes did you smoke per day?

☐☐☐ cigarettes per day

b. How long did you smoke at that rate?

☐☐ years or ☐☐ months (enter 00 if less than 1 month)

c. When was the last time you smoked that much on a regular basis?

☐☐/☐☐

month year

B4. Which brand of cigarettes do you smoke? (If you usually buy generic cigarettes, which brand do you buy if generic are unavailable?)

_____

a. Are your cigarettes: ☐ 1. Non filtered ☐ 2. Filtered

b. Do you smoke menthols? ☐ 1. Yes ☐ 2. No

c. Do you buy a: ☐ 1. Hard pack ☐ 2. Soft pack

d. What size are your cigarettes? ☐ 1. Regular ☐ 2. King size (80-85mm)

☐ 3. 100 mm ☐ 4. 120 mm

e. Are your cigarettes: ☐1. Regular ☐2. Lights

☐3. Medium ☐4. Ultralights

Subject # ☐☐☐☐ Subject Initials ☐☐☐

Other Tobacco Use

B5. Do you currently smoke a pipe regularly? ☐ 1. Yes ☐ 2. No *(if no go to Question B6)*

a. How often do you smoke a pipe? ☐ ☐ times per

(check one) ☐ 1. week or ☐ 2. month or ☐ 3. year

(Not eligible if 2 or more times per week)

b. When did you last smoke a pipe? ☐☐/☐☐/☐☐

month day year

B6. Do you currently smoke cigars regularly? ☐ 1. Yes ☐ 2. No *(if no go to Question B7)*

a. How many cigars do you smoke? ☐ ☐ times per

(check one) ☐ 1. week or ☐ 2. month or ☐ 3. year

(Not eligible if 2 or more times per week)

b. When did you last smoke a cigar? ☐☐/☐☐/☐☐

month day year

B7. Do you currently use chewing tobacco? ☐ 1. Yes ☐ 2. No *(if no go to Question B8)*

a. How often do you use chewing tobacco or snuff? ☐☐times per

(check one) ☐1. week or ☐ 2. month or ☐ 3. year

(Not eligible if 2 or more times per week)

b. When did you last chew tobacco? ☐☐/☐☐/☐☐

month day year

Smoking History

B8. At what age did you smoke your first cigarette? ☐☐ years old

B9. At what age did you become a regular smoker, that is smoked daily?

☐☐ years old

B10. How many times have you made a serious attempt to quit smoking

(you quit for 24 hours or longer)? ☐☐quit attempts

(If none, enter 00. Go to Question B14)

B11. When was the last time you tried to quit? ☐☐/☐☐/☐☐

Subject # ☐☐☐☐	Subject Initials ☐☐☐

B12. What was the longest time you ever quit smoking? (Fill in one period)

☐☐☐ days OR ☐☐☐ months OR☐☐☐ years

B13. What is the **MAIN** reason you went back to smoking? (**CHECK ONLY ONE**)

☐1. Weight gain ☐4. Withdrawal symptoms ☐7. Thought I'd have just one

☐2. Stress ☐5. Liked smoking ☐8. Spouse/close friends smoke

☐3. As a reward ☐6. Was using alcohol ☐9. Other_____

B14. Do you wake up at night to smoke? (i.e. after you go to bed)

◯ 1 Yes ◯ 2 No

B15. Have you ever been without cigarettes for 24 hours or more?

◯ 1 Yes ◯ 2 No *(if no go to Question B16)*

Did you experience any of the following symptoms as a result of being without cigarettes?

	Yes 1	No 2
a. Craving for cigarettes	◯	◯
b. Depressed or sad mood	◯	◯
c. Insomnia	◯	◯
d. Irritability, frustration, or anger	◯	◯
e. Anxiety	◯	◯

f. Difficulty concentrating ○ ○

g. Restlessness ○ ○

h .Increased appetite or weight gain ○ ○

NWSC TOTAL ☐☐ *(Office use only)*

Subject # ☐☐☐☐ Subject Initials ☐☐☐

B16. FTND

a. How soon after you wake up do you smoke your first cigarette?

1☐ 0 - 5 minutes

2☐ 6 - 30 minutes

3☐ 31 - 60 minutes

4☐ More than 60 minutes

b. Do you find it difficult to refrain from smoking in places where it is forbidden (such as in church, at the library, or theater)?

1☐ Yes

2☐ No

c. Which cigarette would you hate most to give up?

1☐ The first one in the morning

2☐ Any other

d. How many cigarettes a day do you smoke?

1☐ 10 or less

2☐ 11-20 cigarettes a day

3☐ 21-30 cigarettes a day

4☐ 30 or more

e. Do you smoke more frequently during the first hours after waking than during the rest of the day?

1☐ Yes

2☐ No

f. Do you smoke if you are so ill that you are in bed most of the day?

1☐ Yes

2☐ No

g. How often do you inhale the smoke from your cigarette?

1☐ Never

2☐ Sometimes

3☐Always

Office Use Only Nicotine Content ☐

FTND Total ☐ ☐

Subject # ☐☐☐☐ Subject Initials ☐☐☐

B17. Have you ever used any of the methods listed below to stop smoking?

Method	# of times tried this method	Ever quit smoking for 24 hours or more with method?	What is the longest period you have used this method?	What is the longest period you stayed smoke-free?

a. Cold turkey		☐1 yes ☐2 no		
b. Nicotine patch		☐1 yes ☐2 no		
c. Nicotine gum		☐1 yes ☐2 no		
d. Nicotine lozenge		☐1 yes ☐2 no		
e. Nicotine nasal spray		☐1 yes ☐2 no		
f. Nicotine inhaler		☐1 yes ☐2 no		
g. Zyban (wellbutrin)		☐1 yes ☐2 no		
h. Chantix (varenicline)		☐1 yes ☐2 no		
g. Clonidine patches		☐1 yes ☐2 no		
h. Hypnosis		☐1 yes ☐2 no		
i. Acupuncture		☐1 yes ☐2 no		
j. Combination of methods (specify:_____)		☐1 yes ☐2 no		
k. Stop smoking clinic		☐1 yes ☐2 no		
i. Other:_____		☐1 yes ☐2 no		
l. Other:_____		☐1 yes ☐2 no		

B18. Have you ever tried to cut down before? ☐ 1. Yes ☐ 2. No *(If no, Go to C1)*

a. Were you successful?

☐ 1. Yes, for how long: ☐☐☐ days OR☐☐☐ months OR☐☐☐ years

☐ 2. No

b. How many cigarettes did you cut out per day? ☐☐

Subject # ☐☐☐☐ Subject Initials ☐☐☐

C. Caffeine and alcohol use

C1. Do you currently drink caffeinated beverages? ☐ 1. Yes

☐ 2. No *(If No go to Question C2)*

a. On average, how many cups of caffeinated beverages do you drink per day?

☐☐ cups of coffee/tea per day

☐☐ cans of soda per day

C2. How often do you drink alcohol?

☐ 1. Never *(If No go to Question D1)*

☐ 2. Monthly or less ☐ 3. 2-4 times per month

☐ 4. 2-3 times per week ☐ 5. 4 or more times per week

C3. On a day that you do drink, how many drinks do you have on average?

☐☐ cans of beers, glasses of wine, and shots of liquor combined

C4. How often do you have 6 or more drinks on one occasion?

☐ 1. never ☐ 2. less than monthly

☐ 3. weekly ☐ 4. daily or almost daily

C5. If you do not currently drink alcohol, how long have you been abstinent?

(If you drink go Question C6) ☐ 1. lifelong ☐ 3. 2-5 years

☐ 2. more than 5 years ☐ 4. 1-2 years

☐ 5. less than 1 year

Why are you abstinent from alcohol?

☐ 1. never drank ☐ 4. quit due to health problems

☐ 2. just not interested ☐ 5. family history of problems with alcohol

☐ 3. alcohol became a problem ☐ 6. other_____

C6. Have you ever been in treatment for problems with alcohol or drugs?

☐ 1. Yes: last time: ☐☐/☐☐/☐☐

☐ 2. No

Subject # ☐☐☐☐	Subject Initials ☐☐☐

D. General Health and Social Environment

D1. Now using a 0-10 scale, with 0 = the worst your health has ever been and 10 =the best your health has ever been, how would you rate your overall health? ☐☐

D2. Has your doctor ever told you to quit smoking? ☐ 1. Yes ☐ 2. No

D3. How many people live with you ? ☐☐ people

D4. How many of the people living with you smoke cigarettes? ☐☐ people

D5. If you have a spouse or significant other, do they smoke cigarettes?

☐ 1. Yes ☐ 2. No ☐ 3. Not applicable

D6. How many of your friends smoke cigarettes?

☐ 1. Almost all ☐ 2. More than half ☐ 3. About half ☐ 4. Less than half ☐ 5. None

D7. Do you smoke (check all that apply)

☐ 1. anywhere at home ☐ 2. only certain rooms at home ☐ 3. smoke outside only

D8. Do you smoke in the car? ☐ 1. Yes ☐ 2. No

D9. Does your employer have any smoking restrictions? ☐ 1. Yes ☐ 2. No