Revealing Oral Microbiota Composition and Functionality Associated with Heavy Cigarette Smoking.

Mohammad Tahseen AL Bataineh (malbataineh@sharjah.ac.ae)
University of Sharjah College of Medicine https://orcid.org/0000-0002-8855-5320
Nihar Ranjan Dash
Clinical Sciences Department, College of Medicine, University of Sharjah, United Arab Emirates
Mohammed Elkhazendar
Clinical Sciences Department, College of Medicine, University of Sharjah, United Arab Emirates
Dua'a Mohammad Hasan Alnusairat
Clinical Sciences Department, College of Medicine, University of Sharjah, United Arab Emirates
Islam Mohammad Ismail Darwish
Clinical Sciences Department, College of Medicine, University of Sharjah, United Arab Emirates
Mohamed Saleh Al-Hajjaj
Clinical Sciences Department, College of Medicine, University of Sharjah, United Arab Emirates. University Hospital Sharjah, Sharjah, United Arab Emirates
Qutayba Hamid
Sharjah Institute for Medical Research, University of Sharjah, United Arab Emirates. Meakins-Christie Laboratories, McGill University, Montreal, QC, Canada

Research

Keywords: Fagerström test, Nicotine dependence, Oral microbiota, Shotgun metagenomics

Posted Date: June 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-32914/v1

Version of Record: A version of this preprint was published on November 10th, 2020. See the published version at https://doi.org/10.1186/s12967-020-02579-3.

Abstract

Background: Heavy tobacco smoking, a hallmark feature of lung cancer is drastically predominant in Middle Eastern populations. The precise links between nicotine dependence and the functional contribution of the oral microbiota remain unknown in these populations.

Methods: We evaluated the functional capabilities of the oral microbiota with relation to cigarette smoking in 105 adults through shotgun metagenomics.

Results: The four major enterotypes initially described in westernized cohorts were retrieved in this population. Differential relative abundance testing unveiled relative abundance of Streptobacillus hongkongensis (Log2FoldChange 4.78, P. adjusted value < 0.00004), Fusobacterium massiliense (Log2FoldChange 4.63, P. adjusted value < 0.00000004), Prevotella bivia (Log2FoldChange 2.46, P. adjusted value < 0.00024) in high nicotine dependent compared to low nicotine dependent profiles based on Fagerström test for Nicotine Dependence. Functional profiling showed marked differences between smokers and non-smokers controls with an enrichment of Tricarballylate utilization (Log2FoldChange 2.52, P. adjusted value < 0.0013) and Lactate racemization (Log2FoldChange 1.003, P. adjusted value < 0.0001) among others in smokers vs . non-smokers group. According to nicotine dependence, we detected enrichment of Xanthosine utilization (Log2FoldChange 3.38, P. adjusted value < 0.00007), p-Aminobenzoyl-Glutamate utilization (Log2FoldChange 1.33, P. adjusted value < 0.00056), and Multidrug efflux pump in Campylobacter jejuni (Log2FoldChange 1.14, P. adjusted value < 0.00007) biosynthesis modules in the high nicotine dependent group.

Conclusions: These differences provide a critical insight on how variations in the oral microbiota may predispose to smoke cessation relapse, serious respiratory illnesses, and lung cancer in heavy cigarette smokers. The observed enrichment of Fusobacterium and Prevotella suggest an intriguing linkage to lung and gut cancers. This information may eventually lead to the development of screening biomarkers to predict early cancer development.

Introduction

The oral microbiota is the second most complex microbial ecosystem after the gut flora, consisting of a dynamic spectrum of microorganisms residing in the oral cavity and its interaction with host genetics, diet, immune system, and many other factors [1]. The bacterial microbiome are the predominant species consisting mainly of obligate aerobes such as *Neisseria* and *Rothia*, facultative aerobes such as *Streptococcus* and *Actinomyces* and obligate anaerobes including *Firmicutes*, *Bacteroidetes* and *Spirochaetes* among others [2]. The communities composition although similar amongst the buccal mucosa, gingiva, & hard palate; yet is different from the soft surfaces, saliva and gingival plaques [3]. In addition, saprophytic protozoa such as *Entamoeba gingivalis* and *Trichomonas tenax* and fungi such as *Candida albicans* and *Saccharomyces cerevisiae* are native residents of oral microbiota [1].

In spite of similarities in the core microbiota composition existing within oral cavities, the species vary depending on the host's diet & nutrition, genetic predisposition, hormonal factors, antibiotic exposure, alcohol consumption and repeated infections by pathogenic bacteria. This alteration, if pathogenic, is termed Dysbiosis [4, 5]. Dysbiosis can contribute to and cause several alterations to the host affecting both oral and systemic health through multiple pathophysiological processes. Dysbiosis of oral microbiota has been reported to be involved in the aetiology of oral diseases such as dental caries, gingivitis and periodontitis; and systemic diseases spanning from infections to cancers, such as respiratory tract infections, gastric ulcers, irritable bowel disease, rheumatoid arthritis, infective endocarditis, and cancers [1, 4, 6].

Tobacco smoking is a well-known preventable cause of death and affects nearly every organ system of the body [7]. The oral cavity is the first regions exposed to cigarette smoke and at a prime disadvantage for increased carcinogenesis, impaired mucosal immunity, and alteration of the oral microbiome [8-10]. Smoking increases colonization of the oral cavity by pathogenic bacteria and reduces colonization by commensal bacteria [11, 12]. Smoking enhances biofilm formation and results in greater epithelial adherence by certain pathogens, including *Steptococcus pneumonia*, *Staphlylococcus aureus*, *Sterptococcus mutans*; thereby, increasing susceptibility to respiratory infections and dental caries respectively in those smokers [8, 10, 12]. Furthermore, smoking contributes to the alteration in the oxygen tension of the oral and upper gastrointestinal microenvironment that encourages persistence of microaerophilic bacteria replacing the commensal beneficial species [12, 13]. Previous studies have shown an increased prevalence of the genera *Atopobium*, *Campylobacter*, and *Prevotella* among smokers and selective depletion of certain phyla including *Proteobacteria* [12, 14-16] .Thus, tobacco smoking creates a unique dysbiotic environment in the oral cavity influencing the microbiota composition that has ~~a~~ far reaching consequences in the local and systemic health of the host [8]. In this study, we intend to decipher our understanding of the oral microbiota's composition and its alteration due to tobacco smoking and smoking severity (nicotine dependence level). Further, we evaluated metabolic capabilities of the oral microbiota using shotgun metagenomic sequencing to determine microbial biodiversity and functional capabilities that associate with tobacco smoking in the oral cavity.

Materials And Methods

Study population

In this case-control study, we recruited participants over an eight-month period between June 2019 and February 2020 in the emirates of Dubai, Sharjah, and Ajman in the United Arab Emirates. Participants completed self-administered questionnaires that included comprehensive demographic, social, and medical history among another lifestyle information. Tobacco smokers were defined as those individuals that reported as exclusively cigarette smokers, on an average, for 11.8 years. Non-smoker controls were defined as individuals who did not report smoking cigarette or any other tobacco products and otherwise healthy. We excluded those who reported antibiotic or prescribed probiotic use in the past 3 months, and

those with preexisting respiratory illness such as asthma and chronic obstructive pulmonary disease in this study.

We have also assessed nicotine dependence by collecting participants' self-administered Fagerström Test for Nicotine Dependence (FTND) scale as previously described [17]. Briefly, yes or no items are scored from 0 to 1 and multiple-choice items are scored from 0 to 3. The items are summed to yield a total score of 0-10. Higher FTND scores indicate greater physical dependence on nicotine. For further validation, participants also completed the Short Nicotine Dependence Syndrome Scale [18, 19].

During data collection phase, we collected 539 buccal swabs from 428 non-smokers and 111 smokers, using Isohelix DNA/RNA Buccal Swabs (Isohelix Ltd. Harrietsham, United Kingdom) following the manufacturer's instruction (Isohelix Ltd.). Case-control matching of tobacco smokers and non-smokers group yielded 105 participants consisting of 50 non-smokers and 55 smokers. The swabs then collected in a sterile container, stored immediately into liquid nitrogen, and then transferred to a −80°C freezer until further analysis. Swabs from these 105 participants were further processed for analysis. All participants in the study read and signed an informed consent; and the Research Ethics Committee at University of Sharjah approved the study protocol.

DNA extraction and library preparation

DNA was extracted using the Qiagen MagAttract PowerSoil DNA KF kit (Formerly MO Bio PowerSoil DNA Kit) using a KingFisher robot. DNA quality was evaluated visually *via* gel electrophoresis and quantified using a Qubit 3.0 fluorometer (Thermo-Fischer, Waltham, MA, USA). Libraries were prepared with the Illumina Nextera library preparation kit using an in-house protocol (Illumina, San Diego, CA, USA).

Sequencing, data curation, and sequence processing

Paired-end sequencing (150 bp x 2) was done on a NextSeq 500 in medium-output mode. Next, shotgun metagenomic sequence reads were processed with the Sunbeam pipeline [20]. Initial quality evaluation was done using FastQC v0.11.5 (Bioinformatics Group at the Babraham Institute. Software available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Processing took part in four steps: adapter removal, read trimming, low-complexity reads removal, and host-sequence read removal. First, adapter removal was done using cutadapt v2.6 [21]. Next, trimming was done with Trimmomatic v0.36 [22] using custom parameters (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36). Then, low-complexity sequences were detected with Komplexity v0.3.6 [20]. Last, read decontamination and removal because they matched one of the pre-specified host/contaminant genomes or due to low complexity or quality. At the end of quality control, the median number of quality-filtered reads per samples was 5062550. The remaining reads were taxonomically classified using Kraken2 with the MiniKraken2_v1 database [23] and with the Genome Taxonomy Database (v. 89).

For functional profiling, high-quality (filtered) reads were aligned against the SEED database *via* translated homology search and annotated to Subsystems, or functional levels, 1-3 using Super-Focus [24].

Statistical analysis

We assessed the alpha diversity with Shannon and Chao1 indices after filtering out spurious OTUs, and then the significance of diversity changes was tested with Wilcoxon rank sum test (Mann Whitney test). Next, we evaluated beta diversity, underscoring differences across samples; a non-metric multidimensional scaling analysis was used to visualize microbiome similarities. Permutational analysis of variance (PERMANOVA) was used to test for the significance of overall microbiome differences. All analyses were conducted in the R environment.

Results

Bacterial summary taxonomic composition

We analyzed buccal swab samples from 105 participants for taxonomic composition, differential abundance and functional profiling of their oral microbiota. The subjects' characteristics in this study such as age, gender, body mass index (BMI), ethnicity, and medical history among others have been provided (Table 1).

Characteristics	Smokers	Non-smokers	p-value
	(n =55)	(n =50)	
Age, years	30.40 (9.508, 41)	30.30 (11.196, 39)	0.961 [a]
Mean (SD, range)			
Gender (M%, F %)	92.7%, 7.3%	90.0%, 10.0%	0.618 [b]
Ethnicity (%)			
MENA	78.20%	76.00%	0.798 [b]
Asians	20.00%	20.00%	
Africans	1.80%	4.00%	
BMI (Kg/m^2)	24.97 (4.65, 16.55)	24.92 (3.33, 14.45)	0.948 [a]
Mean (SD, range)			
Prescribed probiotics use (yes %)	0.00%	0.00%	-
Exercise (yes %)	61.10%	72.00%	0.24 [b]
Animal exposure (yes %)	14.50%	20.00%	0.459 [b]
Antibiotics use (past 3 months) (yes %)	0%	0.00%	-
Family History			
Cancer	12.70%	6.00%	0.241 [b]
HTN	41.80%	30.00%	0.208 [b]
Diabetes	50.90%	30.00%	0.03 [b]
Asthma	5.50%	2.00%	0.356 [b]
Household Smoker (yes %)	61.80%	30.00%	0.001 [b]
Family Smoker (yes %)	65.50%	36.00%	0.003 [b]
Smoking Duration	11.80 (8.065, 38)		
Mean (SD, range)			
FTND	4.82 (2.427, 9)		
Mean (SD, range)			
Low dependence	18.2%		
Low to moderate dependence	32.7%		

Moderate dependence	**32.7%**
High dependence	**16.4%**
[a] : Independent t-test, [b] : Chi-squared test	

Table 1
Demographics of Study Cohort

First, we evaluated the taxonomic composition generated from high-quality reads and classified them using the MiniKraken2_v1 database [23] as the reference database for bacteria. We aggregated taxa abundances into genera and plotted the relative abundances of the most abundant ones (Fig. S1). Furthermore, we plotted the relative abundances of the most abundant taxa within the smokers' group based on their FTND score (nicotine dependence); 1 − 2 (low dependence), 3 − 4 (low to moderate dependence), 5 – 7 (moderate dependence), and ≥ 8 (high dependence) (Fig. S2). Nicotine dependence was further evaluated using the Short Nicotine Dependence Syndrome Scale (NDSS-S) [18, 19]. Pearson correlation suggested a significant positive correlation between FTND and NDDS-S for smokers (r=0.646) (p-value<0.01) (Data not shown). Next, we estimated alpha diversity (richness and evenness) from taxonomic profiles using Shannon's diversity index and Chao1 richness estimator. No significant differences across different groups were found (Fig. S3). Last, to assess the overall microbial community compositional changes, PERMANOVA was used to model effects of smoking and smoking dependence on oral microbiota composition. We observed a significant taxonomy difference between smoker and non-smoker groups (p-value <0.04) and a non-significant difference based on FTND dependence among smoker group (p-value <0.09).

Bacterial differential abundance based on smoking and smoking dependence levels.

In order to further assess possible compositional differences in the bacterial community as suggested in figure S1, we conducted negative binomial models (DESEq2 R package) of the form *~group ~dependence* and ran for differential abundance testing of taxonomic and subsystem level 3 features. P values were calculated with Likelihood Ratio Tests. First, the comparison of average relative abundance between smokers and non-smokers groups revealed that profiles obtained from smokers have a statistically significant abundance of *Veillonella dispar* (Log2FoldChange 2.327, P. adjusted value < 0.0000003), *Leptotrichia* sp000469385 (Log2FoldChange 1.913, P. adjusted value < 0.0013), and *Prevotella pleuritidis* (Log2FoldChange 1.896, P. adjusted value < 0.00019) among others. Whereas, we noted a statistically significant under-representation of *Haemophilus_A* (Log2FoldChange -2.33, P. adjusted value < 0.00007), *Gemella cuniculi* (Log2FoldChange -1.976, P. adjusted value < 0.00019), *Neisseria subflava_B* (Log2FoldChange -1.87, P. adjusted value < 0.00006), *Gemella haemolysans_B* (Log2FoldChange -1.75, P. adjusted value < 0.00085), *Neisseria perflava* (Log2FoldChange -1.73, P. adjusted value < 0.0012), *Streptococcus oralis_BA* (Log2FoldChange -1.56, P. adjusted value < 0.0004),

and *Streptococcus mitis_AT* (Log2FoldChange -1.39, P. adjusted value < 0.0013) among others in smokers (Fig. 1). Next, we evaluated average relative abundance among smokers based on nicotine dependence (Fagerström score) and showed that profiles obtained from high nicotine dependence smokers have a statistically significant abundance of *Streptobacillus hongkongensis* (Log2FoldChange 4.78, P. adjusted value < 0.00004), *Fusobacterium massiliense* (Log2FoldChange 4.63, P. adjusted value < 0.00000004), *Prevotella sp000163055* (Log2FoldChange 4.42, P. adjusted value < 0.00008), and *Prevotella bivia* (Log2FoldChange 2.46, P. adjusted value < 0.00024) among others (Fig. 2).

Functional profiling of oral microbiota in smoker vs. non-smokers

We used shotgun metagenomic sequencing to determine functional contribution of the oral microbiota in smokers *vs*. non-smokers using the SEED hierarchical categorization. Functional profiling showed significant enrichment of Tricarballylate utilization (Log2FoldChange 2.52, P. adjusted value < 0.0013), Aminoglycoside adenylyltransferases (Log2FoldChange 2.39, P. adjusted value < 0.002), Bacteriocins in Lactobacilli (Log2FoldChange 2.29, P. adjusted value < 0.0012), Lactate racemization (Log2FoldChange 1.003, P. adjusted value < 0.0001), and Methionine salvage (Log2FoldChange 0.7, P. adjusted value < 0.0004) among others in smokers. Whereas, we noted a significant depletion of Two-component Response Regulator of Virulence ResDE (Log2FoldChange -1.28, P. adjusted value < 0.0009), Listeria Pathogenicity Island LIPI-1 extended (Log2FoldChange -0.888, P. adjusted value < 0.00006), and CarD (Log2FoldChange -0.139, P. adjusted value < 0.0007) among others in smokers (Fig. 3).

Functional profiling of oral microbiota based on nicotine dependence severity

Finally, we examined differentially abundant gene functions based on Fagerström score for nicotine dependence among smokers. Pairwise functional differences determined significant difference between low and high nicotine dependence groups (p-value < 0.02, p-value FDR<0.05). For example, we show enrichment of Xanthosine utilization (xap region) (Log2FoldChange 3.38, P. adjusted value < 0.00007), p-Aminobenzoyl-Glutamate utilization (Log2FoldChange 1.33, P. adjusted value < 0.00056), Multidrug efflux pump in Campylobacter jejuni (CmeABC operon) (Log2FoldChange 1.14, P. adjusted value < 0.00007), Glycine biosynthesis (Log2FoldChange 1.02, P. adjusted value < 0.00062), Isoleucine degradation (Log2FoldChange 0.989, P. adjusted value < 0.00021) among others. We also noted depletion of Type VI secretion systems (Log2FoldChange -1.99, P. adjusted value < 0.00027), Rrf2 family transcriptional regulators (Log2FoldChange -0.598, P. adjusted value < 0.00067), and ABC transporter oligopeptide (TC 3.A.1.5.1) (Log2FoldChange -0.351, P. adjusted value < 0.00001) among others in the high nicotine dependence group (Fig. 4).

Discussion

In this report, we attempted to explore oral microbial profiles and functions that influence host homeostasis in relation to heavy cigarette smoking. We explored the oral microbiota of tobacco smokers in Middle-Eastern population and described, for the first time, the functional contribution of the oral bacterial community based on nicotine dependence assessed by the Fagerström scale [17]. Consistent with previous several reports, we detected a significant taxonomic difference between smoker and non-smoker groups, but no significant differences in terms of microbial diversity and richness as shown in figure S3 [25-27]. Interestingly, a previous study conducted in the UAE determined only marginal significance of the overall oral microbial differences in smokers compared with non-smokers, underscoring the geographic and ethnic contribution [15]. However, our finding was not consistent with other groups reporting significant change in richness and diversity [28, 29]. The observed fluctuations in oral microbiota richness and diversity reporting by several groups are not unusual and further assert the high complexity and major effects of several factors such as diet, geography, ethnicity, and host factors. That said, the oral microbiota in our study exhibit comparable dominance of phyla *Firmicutes*, *Proteobacteria*, *Actinobacteria*, *Bacteroidetes* and genera *Prevotella*, and *Veillonella* to that of oral microbiota in other populations across the globe [16, 29, 30].

Differential abundance testing of bacterial communities based on nicotine dependence scores revealed relative abundance of *Streptobacillus hongkongensis* (Log2FoldChange 4.78, P. adjusted value < 0.00004) among high nicotine dependent smokers (high Fagerström score). Previous studies reported isolation of *S. hongkongensis* from patients with quinsy, pneumonia, and septic arthritis [31, 32], which was later reported as part of the human oropharynx natural reservoir [33]. Therefore, high nicotine dependent smokers are poised to an increased risk to develop serious respiratory illnesses. Furthermore, complications of streptobacilliary infections may include endocarditis, brain abscesses, amnionitis, as well as persistent severe arthritis [34].

Smoking tobacco is the single largest risk factor for lung cancers. Several studies established that *Fusobacterium nucleatum* play a major role in colorectal carcinogenesis *via* Fap2 mediated binding to tumor-overexpressed Gal-GalNAc-binding lectin [35-37]. Therefore, *F. nucleatum* deemed useful as a microbial biomarker for colorectal cancer detection [38]. Interestingly, we discovered the phylogenetically similar *Fusobacterium massiliense* that exhibited substantial sequence similarity with *F. nucleatum*, have a significant relative abundance among high nicotine dependent smokers (Log2FoldChange 4.63, P. adjusted value < 0.00000004). Further, protein-protein BLAST analysis of the Fap2 surface protein of *F. nucleatum* ATCC 23726 produced a significant sequences alignment with pyridoxal phosphate-dependent aminotransferase of *F. massiliense* [36, 39], the active form of vitamin B6. A previous study examined over 44,000 individuals and evaluated their smoking history and B6 vitamin supplement use over 10 years; found that high dosages of vitamin B6 supplements were associated with 3-4 folds increase in lung cancer risk among smokers [40]. Altogether, perhaps enrichment of *F. massiliense* among high nicotine dependent smokers suggest a possible linkage to lung cancer in a pyridoxal phosphate-dependent manner. In addition, tobacco smoking associates with colorectal cancers and with an intriguing correlation with increased abundance of the gut *Prevotella* [41]. Here, we also noted an increase relative abundance of *Prevotella sp000163055* (Log2FoldChange 4.42, P. adjusted value < 0.00008) and

Prevotella bivia (Log2FoldChange 2.46, P. adjusted value < 0.00024) in oral microbiota of heavy smokers, thereby suggesting a possible downstream effect on the development of colorectal cancers.

Next, we evaluated metabolic capabilities of oral microbiota using shotgun metagenomic sequencing approach to determine microbial biodiversity and functional capabilities associated with tobacco smoking in the oral cavity. Functional profiling showed significant enrichment of Tricarballylate utilization (Log2FoldChange 2.52, P. adjusted value < 0.0013) among smokers *vs*. non-smokers group, a good chelator of magnesium leading to magnesium deficiency [42]. Magnesium plays an important role in tobacco addiction by inhibiting several essential steps of nicotine addiction such as dopamine secretion, NMDA receptor stimulation by glutamate, and the synthesis of substance P and nitric oxide [43, 44]. In fact, a previous study showed a significant decrease in the number of cigarettes smoked and Fagerström score after 28 days of magnesium therapy [45]. It is fascinating to observe a significant enrichment of bacterial genes involved in Tricarballylate utilization among smokers, suggesting an intriguing role for the dysbiotic oral microbiota in maintaining nicotine addiction and perhaps influence smoke cessation relapse. Moreover, a significant increase in the nickel-dependent lactate racemase enzymes was observed in smokers (Log2FoldChange 1.003, P. adjusted value < 0.0001), consistent with the toxic nickel exposure from tobacco smoking [46, 47].

Finally, we examined differentially abundant gene functions based on Fagerström score for nicotine dependence among smokers. Heavy smokers consume more coffee than others to obtain the same satisfactory effect of caffeine, as reported in a study of two European cohorts, which determined a positive association between smoking and coffee consumption [48]. Remarkably, we observed a significant enrichment of xanthosine utilization (Log2FoldChange 3.38, P. adjusted value < 0.00007) among high nicotine dependent smokers, a catabolite of purine nucleotides that leads to caffeine synthesis [49]. Perhaps xanthosine utilization may further contribute to caffeine toxicity among heavy smokers, which subsequently could make smoking cessation even more difficult [50]. Furthermore, smokers have lower levels of folic acid and this is reflective in our study by folate catabolism *via* upregulation of p-Aminobenzoyl-Glutamate utilization (Log2FoldChange 1.33, P. adjusted value < 0.00056), especially among heavy smokers [51, 52]. Lastly, we noted an enrichment of Multidrug efflux pump in *Campylobacter jejuni* (CmeABC operon) (Log2FoldChange 1.14, P. adjusted value < 0.00007) biosynthesis module in the heavy smokers group, an important component of bacterial virulence that predispose heavy smokers to additional risk of tobacco-related morbidity and mortality [53]. It is important to mention that our findings need further validation on a larger cohort. The data obtained from self-administered questionnaires was subject to self-reporting bias; however, a study staff was available during the questionnaire to answer any questions.

Conclusions

We used shotgun metagenomic approach to shed a new light into the complex functional profiles of the oral microbiota in tobacco smokers from the Middle East. To the best of our knowledge, this is the first report on oral microbiota role in heavy smoking among Middle Eastern populations based on nicotine

dependence assessed by the Fagerström test. Despite the need of further investigations, our data identified important compositional and functional variations in microbial communities in correlation with various clinicopathological predispositions in those with high nicotine dependence (heavy smokers); including more serious respiratory illnesses, lung cancer, smoke cessation relapse, and caffeine intoxication leading to increased morbidity and mortality. This information may eventually lead to screening biomarkers to predict early cancer development and improve tobacco control strategies.

Abbreviations

FTND: Fagerström Test for Nicotine Dependence

NDDS-S: Short Nicotine Dependence Syndrome Scale

HTN: Hypertension

MENA: Middle East and North Africa

SEED: Categorization system that organizes gene functional categories into a hierarchy

SUPER-FOCUS: A Tool for Agile Functional Analysis of Shotgun Metagenomic Data

BMI: Body mass index

Declaration

Ethics approval and consent to participate:

All participants in the study read and signed an informed consent; and the Research Ethics Committee at University of Sharjah approved the study protocol, reference # REC-18-02-18-01.

Consent for publication

There are no individual person identifiers in this manuscript. Consent for publication was not sought.

Competing interests

All authors declare they have no competing interests.

Funding

The authors were supported by UAE funding agencies; MTA funded by University of Sharjah (Grant code: 1901090253). QH funded by University of Sharjah (Grant code: 150317).

Authors' contributions

Conception and design: MTA, NRD, MSA, and QH. Acquisition of data: ME, DMHA, IMID. Processing of specimens and generation of data: MTA, ME, DMHA, IMID Analysis and interpretation of data: MTA, NRD, ME, DMHA. Drafting or revising of manuscript: MTA, NRD, MSA, QH, ME, DMHA, IMID. Final approval of manuscript: MTA, NRD, MSA, QH, ME, DMHA, IMID. MTA has access to all study data and takes responsibility for the data integrity and accuracy.

Acknowledgements

We thank Nouran Nada, M.Besher Adi, Aya Fathy, Rushud Abdulsalam, and Rafla Ali for helping in data and sample collection phases of the study. Sequencing data provided by microbiome insights, Canada.

Availability of data and material

The data are all published. The data that support the findings of this study are available on request from the corresponding author.

References

1. Wade WG: **The oral microbiome in health and disease.** *Pharmacol Res* 2013, **69:**137-143.
2. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner ACR, Yu W-H, Lakshmanan A, Wade WG: **The Human Oral Microbiome.** *J Bacteriol* 2010, **192:**5002-5017.
3. Segata N, Haake SK, Mannon P, Lemon KP, Waldron L, Gevers D, Huttenhower C, Izard J: **Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples.** *Genome Biol* 2012, **13**.
4. Appanna V: **Dysbiosis, Probiotics, and Prebiotics: In Diseases and Health.** In; 2018: 81-122
5. Zaura E, Nicu EA, Krom BP, Keijser BJF: **Acquiring and maintaining a normal oral microbiome: current perspective.** *Front Cell Infect Microbiol* 2014, **4:**85.
6. Kilian M, Chapple ILC, Hannig M, Marsh PD, Meuric V, Pedersen AML, Tonetti MS, Wade WG, Zaura E: **The oral microbiome - an update for oral healthcare professionals.** *British Dental Journal* 2016, **221:**657-666.
7. World Health O: **WHO report on the global tobacco epidemic 2019: Offer help to quit tobacco use.** 2019.
8. Huang C, Shi G: **Smoking and microbiome in oral, airway, gut and some systemic diseases.** *Journal of Translational Medicine* 2019, **17**.
9. West R: **Tobacco smoking: Health impact, prevalence, correlates and interventions.** *Psychology & Health* 2017, **32:**1018-1036.

10. Tomar SL, Hecht SS, Jaspers I, Gregory RL, Stepanov I: **Oral Health Effects of Combusted and Smokeless Tobacco Products.** *Adv Dent Res*, **30:**4-10.
11. Mason MR, Preshaw PM, Nagaraja HN, Dabdoub SM, Rahman A, Kumar PS: **The subgingival microbiome of clinically healthy current and never smokers.** *ISME J* 2015, **9:**268-272.
12. Brook I: **The impact of smoking on oral and nasopharyngeal bacterial flora.** *J Dent Res* 2011, **90:**704-710.
13. Jaspers I: **Cigarette smoke effects on innate immune mechanisms in the nasal mucosa. Potential effects on the microbiome.** *Ann Am Thorac Soc* 2014, **11 Suppl 1:**38-42.
14. Yang Y, Zheng W, Cai Q-Y, Shrubsole MJ, Pei Z, Brucker R, Steinwandel MD, Bordenstein SR, Li Z, Blot WJ, et al: **Cigarette smoking and oral microbiota in low-income and African-American populations.** *J Epidemiol Community Health* 2019, **73:**1108-1115.
15. Vallès Y, Inman CK, Peters BA, Ali R, Wareth LA, Abdulle A, Alsafar H, Anouti FA, Dhaheri AA, Galani D, et al: **Types of tobacco consumption and the oral microbiome in the United Arab Emirates Healthy Future (UAEHFS) Pilot Study.** *Sci Rep* 2018, **8:**1-11.
16. Wu J, Peters BA, Dominianni C, Zhang Y, Pei Z, Yang L, Ma Y, Purdue MP, Jacobs EJ, Gapstur SM, et al: **Cigarette smoking and the oral microbiome in a large study of American adults.** *ISME J* 2016, **10:**2435-2446.
17. Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom KO: **The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire.** *Br J Addict* 1991, **86:**1119-1127.
18. Shiffman S, Sayette MA: **Validation of the nicotine dependence syndrome scale (NDSS): a criterion-group design contrasting chippers and regular smokers.** *Drug Alcohol Depend* 2005, **79:**45-52.
19. Becona E, Fernandez del Rio E, Lopez A, Miguez Mdel C, Castro J, Nogueiras L, Florez G, Alvarez S, Vazquez D: **[The Short Nicotine Dependence Syndrome Scale (NDSS-S) in Spanish smokers].** *Psicothema* 2011, **23:**126-132.
20. Clarke EL, Taylor LJ, Zhao C, Connell A, Lee JJ, Fett B, Bushman FD, Bittinger K: **Sunbeam: an extensible pipeline for analyzing metagenomic sequencing experiments.** *Microbiome* 2019, **7:**46.
21. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** *2011* 2011, **17:**3.
22. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data.** *Bioinformatics* 2014, **30:**2114-2120.
23. Wood DE, Lu J, Langmead B: **Improved metagenomic analysis with Kraken 2.** *Genome Biol* 2019, **20:**257.
24. Silva GG, Green KT, Dutilh BE, Edwards RA: **SUPER-FOCUS: a tool for agile functional analysis of shotgun metagenomic data.** *Bioinformatics* 2016, **32:**354-361.
25. Jiang Y, Zhou X, Cheng L, Li M: **The Impact of Smoking on Subgingival Microflora: From Periodontal Health to Disease.** *Front Microbiol* 2020, **11:**66.

26. Bostrom L, Bergstrom J, Dahlen G, Linder LE: **Smoking and subgingival microflora in periodontal disease.** *J Clin Periodontol* 2001, **28:**212-219.
27. Gomes SC, Nonnenmacher C, Susin C, Oppermann RV, Mutters R, Marcantonio RA: **The effect of a supragingival plaque-control regimen on the subgingival microbiota in smokers and never-smokers: evaluation by real-time polymerase chain reaction.** *J Periodontol* 2008, **79:**2297-2304.
28. Joshi V, Matthews C, Aspiras M, de Jager M, Ward M, Kumar P: **Smoking decreases structural and functional resilience in the subgingival ecosystem.** *J Clin Periodontol* 2014, **41:**1037-1047.
29. Takeshita T, Kageyama S, Furuta M, Tsuboi H, Takeuchi K, Shibata Y, Shimazaki Y, Akifusa S, Ninomiya T, Kiyohara Y, Yamashita Y: **Bacterial diversity in saliva and oral health-related conditions: the Hisayama Study.** *Sci Rep* 2016, **6:**22164.
30. Contreras M, Costello EK, Hidalgo G, Magris M, Knight R, Dominguez-Bello MG: **The bacterial microbiota in the oral mucosa of rural Amerindians.** *Microbiology* 2010, **156:**3282-3287.
31. Woo PC, Wu AK, Tsang CC, Leung KW, Ngan AH, Curreem SO, Lam KW, Chen JH, Chan JF, Lau SK: **Streptobacillus hongkongensis sp. nov., isolated from patients with quinsy and septic arthritis, and emended descriptions of the genus Streptobacillus and Streptobacillus moniliformis.** *Int J Syst Evol Microbiol* 2014, **64:**3034-3039.
32. Eisenberg T, Ewers C, Rau J, Akimkin V, Nicklas W: **Approved and novel strategies in diagnostics of rat bite fever and other Streptobacillus infections in humans and animals.** *Virulence* 2016, **7:**630-648.
33. Lau SK, Chan JF, Tsang CC, Chan SM, Ho ML, Que TL, Lau YL, Woo PC: **Human oropharynx as natural reservoir of Streptobacillus hongkongensis.** *Sci Rep* 2016, **6:**24419.
34. Kerr J: **Manual of clinical microbiology.** BMJ Publishing Group; 2004.
35. Mima K, Cao Y, Chan AT, Qian ZR, Nowak JA, Masugi Y, Shi Y, Song M, da Silva A, Gu M, et al: **Fusobacterium nucleatum in Colorectal Carcinoma Tissue According to Tumor Location.** *Clin Transl Gastroenterol* 2016, **7:**e200.
36. Abed J, Emgard JE, Zamir G, Faroja M, Almogy G, Grenov A, Sol A, Naor R, Pikarsky E, Atlan KA, et al: **Fap2 Mediates Fusobacterium nucleatum Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc.** *Cell Host Microbe* 2016, **20:**215-225.
37. Brennan CA, Garrett WS: **Fusobacterium nucleatum — symbiont, opportunist and oncobacterium.** *Nature Reviews Microbiology* 2019, **17:**156-166.
38. Peng BJ, Cao CY, Li W, Zhou YJ, Zhang Y, Nie YQ, Cao YW, Li YY: **Diagnostic Performance of Intestinal Fusobacterium nucleatum in Colorectal Cancer: A Meta-Analysis.** *Chin Med J (Engl)* 2018, **131:**1349-1356.
39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215:**403-410.
40. Brasky TM, White E, Chen C-L: **Long-Term, Supplemental, One-Carbon Metabolism–Related Vitamin B Use in Relation to Lung Cancer Risk in the Vitamins and Lifestyle (VITAL) Cohort.** *Journal of Clinical Oncology* 2017, **35:**3440-3448.

41. Stewart CJ, Auchtung TA, Ajami NJ, Velasquez K, Smith DP, De La Garza R, 2nd, Salas R, Petrosino JF: **Effects of tobacco smoke and electronic cigarette vapor exposure on the oral and gut microbiota in humans: a pilot study.** *PeerJ* 2018, **6:**e4693.
42. Lewis JA, Escalante-Semerena JC: **The FAD-dependent tricarballylate dehydrogenase (TcuA) enzyme of Salmonella enterica converts tricarballylate into cis-aconitate.** *J Bacteriol* 2006, **188:**5479-5486.
43. Hiranita T, Soto PL, Kohut SJ, Kopajtic T, Cao J, Newman AH, Tanda G, Katz JL: **Decreases in cocaine self-administration with dual inhibition of the dopamine transporter and sigma receptors.** *J Pharmacol Exp Ther* 2011, **339:**662-677.
44. Veeneman MM, Broekhoven MH, Damsteegt R, Vanderschuren LJ: **Distinct contributions of dopamine in the dorsolateral striatum and nucleus accumbens shell to the reinforcing properties of cocaine.** *Neuropsychopharmacology* 2012, **37:**487-498.
45. Nechifor M, Chelarescu D, Mandreci I, Cartas N: **Magnesium influence on nicotine pharmacodependence and smoking.** *Magnes Res* 2004, **17:**176-181.
46. Desguin B, Goffin P, Viaene E, Kleerebezem M, Martin-Diaconescu V, Maroney MJ, Declercq J-P, Soumillion P, Hols P: **Lactate racemase is a nickel-dependent enzyme activated by a widespread maturation system.** *Nat Commun* 2014, **5:**3615.
47. Torjussen W, Zachariasen H, Andersen I: **Cigarette smoking and nickel exposure.** *J Environ Monit* 2003, **5:**198-201.
48. Treur JL, Taylor AE, Ware JJ, McMahon G, Hottenga JJ, Baselmans BM, Willemsen G, Boomsma DI, Munafo MR, Vink JM: **Associations between smoking and caffeine consumption in two European cohorts.** *Addiction* 2016, **111:**1059-1068.
49. Ashihara H, Kato M, Crozier A: **Distribution, Biosynthesis and Catabolism of Methylxanthines in Plants.** In *Methylxanthines.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2011: 11-31
50. Ashihara H, Kato M, Crozier A: **Distribution, biosynthesis and catabolism of methylxanthines in plants.** *Handb Exp Pharmacol* 2011**:**11-31.
51. Vardavas CI, Linardakis MK, Hatzis CM, Malliaraki N, Saris WH, Kafatos AG: **Smoking status in relation to serum folate and dietary vitamin intake.** *Tob Induc Dis* 2008, **4:**8.
52. Green JM, Hollandsworth R, Pitstick L, Carter EL: **Purification and characterization of the folate catabolic enzyme p-aminobenzoyl-glutamate hydrolase from Escherichia coli.** *J Bacteriol* 2010, **192:**2407-2413.
53. Warren GW, Alberg AJ, Kraft AS, Cummings KM: **The 2014 Surgeon General's report:"The Health Consequences of Smoking–50 Years of Progress": a paradigm shift in cancer care.** *Cancer* 2014, **120:**1914-1916.

Figures

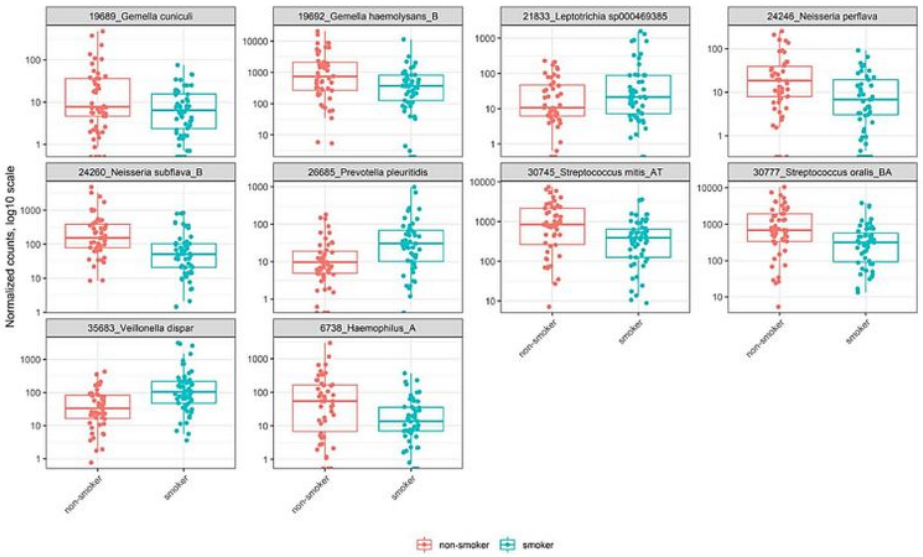

Figure 1

Differentially abundant taxa between smokers and non-smokers group. Panel shows relative abundance of normalized counts for the top 10 taxa. Results were calculated by negative binomial models (DESEq2 R package) of the form ~group for differential abundance testing of taxonomic and subsystem level 3 features. P values were calculated with Likelihood Ratio Tests method. Smoker and non-smoker corresponding abundance are colored in blue and red, respectively.

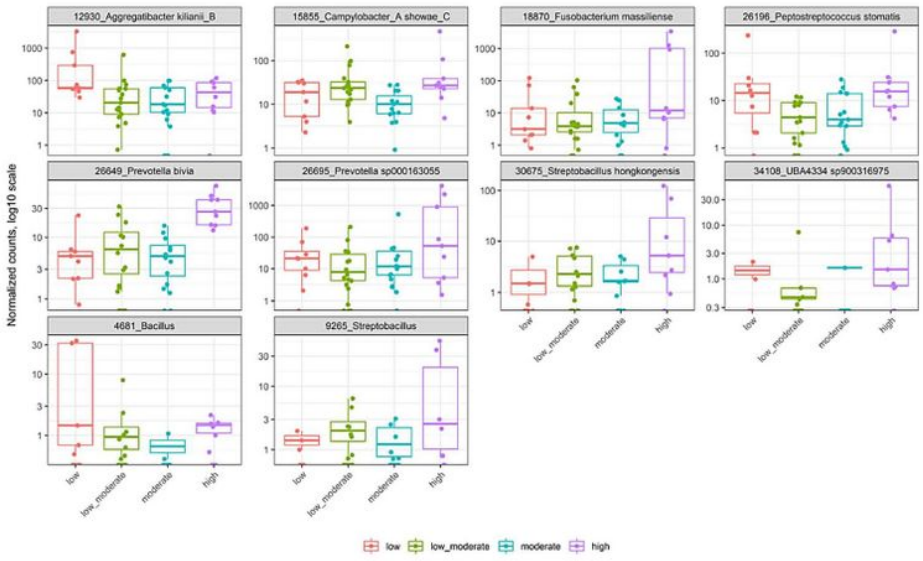

Figure 2

Differentially abundant taxa based on FNTD nicotine dependence score. Panel shows relative abundance of normalized counts for the top 10 taxa. Results were calculated by negative binomial models (DESEq2 R package) of the form ~group for differential abundance testing of taxonomic and subsystem level 3 features. P values were calculated with Likelihood Ratio Tests method. Nicotine dependence FTND scores; low, low to moderate, moderate, and high are colored in red, green, blue and pink, respectively.

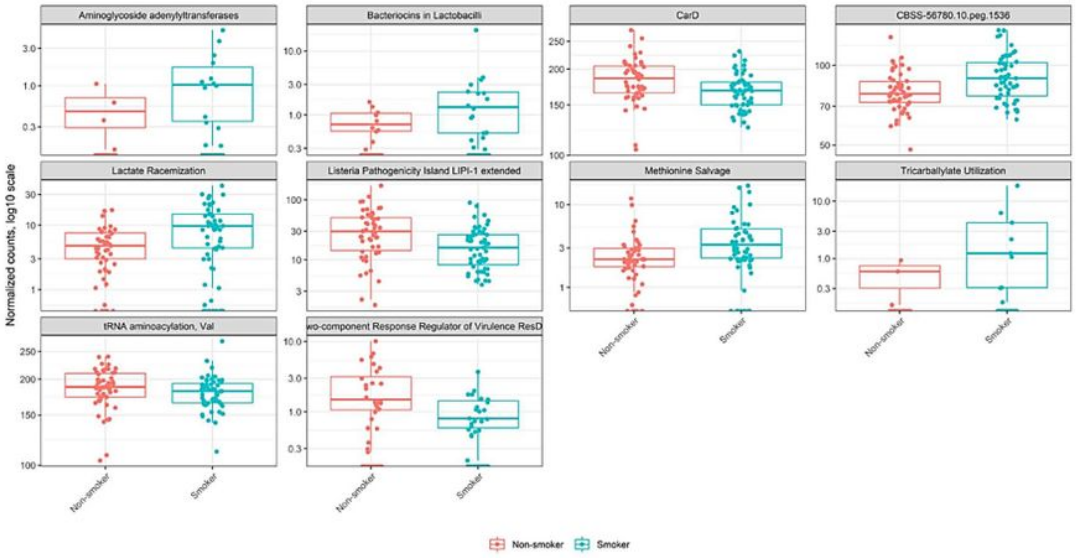

Figure 3

Differentially abundant gene functions of smokers vs. non-smokers group. Panel shows relative abundance of normalized counts for functional genes using SEED hierarchical categorization. Smoker and non-smoker corresponding abundance are colored in blue and red, respectively.

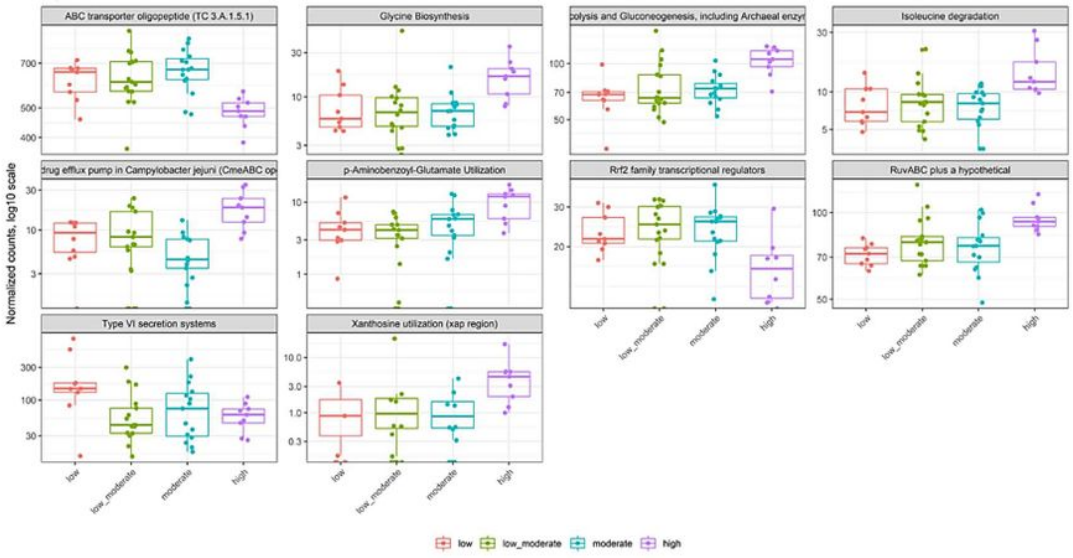

Figure 4

Differentially abundant gene functions based on FNTD nicotine dependence score. Panel shows relative abundance of normalized counts for functional genes using SEED hierarchical categorization. Smoking dependence, low, low to moderate, moderate, and high are colored in red, green, blue and pink, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- suppl..pdf