# Tongue microbiome of smokeless tobacco users

**Esam Halboub** ( mhelboub@gmail.com )
Jazan University Faculty of Dentistry https://orcid.org/0000-0002-1894-470X
**Mohammed Alakhali**
Jazan University Faculty of Dentistry
**Abdulwahhab H. Al-Amir**
Jazan University Faculty of Dentistry
**Husham E. Homeida**
Jazan University Faculty of Dentistry
**Divyashri Baraniya**
Temple University Maurice H Kornberg School of Dentistry
**Tsute Chen**
The Forsyth Institute
**Nezar Noor Al-Hebshi**
Temple University Maurice H Kornberg School of Dentistry

**Research**

**Keywords:** Rothia mucilaginosa, High-Throughput Nucleotide Sequencing; Microbiota; Mouth neoplasms; Tobacco; Tongue

**Posted Date:** February 24th, 2020

**DOI:** https://doi.org/10.21203/rs.2.24331/v1



**Version of Record:** A version of this preprint was published at BMC Microbiology on July 8th, 2020. See the published version at https://doi.org/10.1186/s12866-020-01883-8.

# Abstract

Background: The possibility that smokeless tobacco may contribute to oral carcinogenesis by influencing the oral microbiome has not been explored. This cross sectional study sought to assess the effect of using hammah, a form of smokeless tobacco prevalent in Arabia, on the tongue microbiome. Tongue scarping samples were obtained from twenty-nine shammah users (SU; 27.34±6.9 years) and 23 shammah non-users (SNU; 27.7±7.19 years) and analyzed with 16S rRNA gene sequencing (V1-V3). Species-level taxonomy assignment of the high-quality, merged reads was obtained using a previously described BLASTn-based algorithm. Downstream analyses were performed with QIIME, LEfSe, and R.

Results: A total of 178 species, belonging to 62 genera and 8 phyla were identified. Genera Streptococcus , Leptotrichia , Actinomyces , Veillonella , Haemophilus , Prevotella and Neisseria accounted for more than 60% of the average microbiome. There were no differences between the two groups in species richness and alpha-diversity, but PCoA showed significant separation (P=0.015, ANOSIM). LEfSe analysis identified 22 species to be differentially abundant between the SU and SNU. However, only 7 species maintained a false discovery rate of $\leq 0.2$ and could cluster the two groups separately: Rothia mucilaginosa , Streptococcus sp. oral taxon 66, Actinomyces meyeri , Streptococcus vestibularis Streptococcus sanguinis and a potentially novel Veillonella species in association with SU, and Oribacterium asaccharolyticum with SNU.

Conclusion: Shammah use induces tongue microbiome changes that may be relevant to oral carcinogenesis, namely enrichment of species with high acetaldehyde production potential, which warrants further investigation.


# Background

Tobacco use is highly prevalent and remains a major global health threads worldwide, being responsible for killing 8 million people annually [1]. Based on how it is used, there are two major forms of tobacco: smoked, which is the most common form, and smokeless form which is used without burning [1]. Smokeless tobacco (ST) products are typically chewed, dipped, sucked or applied as a paste to the gingiva [2]. Both forms of tobacco are major risk factors of oral cancer, with pooled odds ratios of 3.6 and 7.9 for smoking and ST, respectively, according to one metaanalysis [3]. It is estimated that 90% of the global use burden of ST is in South Asia, where oral cancer ranks among the most common cancers (first or second in some countries like India) [4].

The carcinogenic effect of tobacco is ascribed to a wide range of carcinogens such as the tobacco-specific nitrosamine (e.g. N'-nitrosonornicotine), polycyclic aromatic hydrocarbons, metals and metalloids, and aldehydes, in addition to many co-carcinogens and toxicants [5]. Basically, these carcinogens undergo metabolic activation to intermediates that react with DNA to form what is called DNA adducts. The latter in turn, when cellular repair mechanisms fail, can result in permanent mutations in oncogenes and tumor suppressor genes, leading to development of oral cancer [5]. In addition to

genetic aberration, tobacco also contributes to oral carcinogenesis by inducing epigenetic alterations and immune dysfunctions [6].

One possible, yet unexplored, mechanism by which tobacco may further contribute to the development of oral cancer is through disruption of the oral microbiome. This is important in view of the increasing evidence indicating that compositional and functional disturbances in the oral microbiome (dysbiosis) may play a role in oral cancer [7, 8]. In fact, few recent studies have shown that current smokers have an significantly altered oral microbiome compared to non- or former smokers [9, 10], Furthermore, in a Syrian Golden hamster cheek pouch carcinogenesis model, 4-week application of ST was shown to significantly disrupted the oral microbiota [11]. How ST products affect the human oral microbiome, and whether that may play a role in their carcinogenicity has not been studied.

Shammah, also known as Arabian snuff, is a form of ST that is used in Yemen and Saudi Arabia, and is strongly associated with oral potentially malignant lesions and oral cancer [12, 13]. On the grounds that the tongue is densely populated by a diverse microbial community [14], while it is also the most commonly affected site by oral cancer, including that associated with shammah use [15], the objective of this study was to explore the potential effect of chronic use of shammah, as an example of ST product, on the tongue microbiome in comparison to the tongue microbiome of shammah non-users.

# Results

### *Characterization of the study sample*

The characteristics of the study groups are presented in **Table 1**. Fifty-two males participated in this study: 29 SU and 23 SNU with a comparable mean age (27.34±6.9 and 27.7±7.19 years, respectively). Among the former, 17 reported using the black shammah type, while the rest reported using the white type. The mean duration of shammah use was 8.66±7.11 years, with a mean frequency of use of 11.86±4.43 times per a day. Seven subjects reported chewing qat (a plant with amphetamine-like effect that is habitually chewed) in addition to using shammah. There were 7 and 3 cigarette smokers among the SU and SNU groups, respectively. The DMFT scores did not differ between the two groups (5.41±4.38 and 5.78±4.11 for SU and SNU, respectively).

### *Sequencing and data preprocessing statistics*

A total of 3,088,631 raw reads were obtained (publically available from Sequence Read Archive; Project ID PRJNA605810) of which 90% could be successfully merged. Nearly 52% of the merged reads were filtered out during the highly stringent quality-filtration step, and additional 11% were identified as chimeric sequences and thus removed. About 80% of the high-quality, non-chimeric sequences were successfully classified to the species-level (mean of 14957± 5335 reads per sample). The detailed sequencing and data preprocessing statistics are provided in **Supplementary Dataset 1**

### *General microbiological profiles*

A total of 178 species, including 24 potentially novel species, belonging to 62 genera and 8 phyla were identified in the samples. The detection frequency and per-sample relative abundances of each taxon are presented in **Supplementary Datasets 2-4**. On average, 109 species (range 80-143) and 44 genera (range 32-55) were detected per subject. Fifty-three species and 26 genera were identified in more than 90% of the samples (i.e. can be defined as core taxa of the dorsum of the tongue). The average relative abundances of the phyla and top genera and species (those present at an average abundance of ≥ 2% in the control group) in each of the study groups are shown in **Figure 1**. Firmicutes, Actinobacteria, Proteobacteria, Fusobacteria, and Bacteroidetes were, in order, the most abundant phyla accounting for at least 97% of the reads in each sample. The top 13 genera accounted for more than 80% of the average microbiome, with *Streptococcus*, *Leptotrichia* and *Actinomyces* alone comprising ~ 40%. The top 15 species constituted ~ 55% of the reads on average, with *Neisseria flavescens/subflava*, *Haemophilus parainfluenzae*, *Rothia mucilaginosa*, *Veillonella parvula group*, *Streptococcus salivarius*, *Leptotrichia sp.* oral taxon 417, *Leptotrichia sp.* oral taxon 215 and *Actinomyces graevenitzii* making ~ 37%.

***Diversity and differentially abundance taxa***

There were no statistically significant differences between the two study groups in species richness or alpha diversity indices as illustrated in **Figure 2A**. However, analysis of beta diversity by PCoA (based on Bray-Curtis distance matrix) showed significant (P=0.015, Analysis of Similarities), but not complete, separation between the two groups **(Figure 2B).**

LEfSe analysis identified 7 genera and 22 species to be differentially abundant between the two groups (Figure 3). Most of them maintained the associations when the qat chewers were excluded (**Supplementary Figure 1**). However, after adjustment for multiple comparisons using the Benjamini-Hochberg method, only seven species had a false discovery rate (FDR) of less than 0.2, namely *R. mucilaginosa*, *Streptococcus sp.* oral taxon 66, *Actinomyces meyeri*, *Streptococcus vestibularis* *Streptococcus sanguinis*, a potentially novel *Veillonella* species and *Oribacterium asaccharolyticum*. Centroid-based hierarchal clustering of the samples by the relative abundance of these species resulted in separation between the shammah users and non-users (**Figure 4**). The relative abundances of the 7 species in individual samples are shown in **Figure 5**.

# Discussion

The current study characterized the tongue microbiome associated with use of shammah, as a highly carcinogenic ST product, with the aim of identifying shifts that may be relevant to development of oral cancer. In other words, the study is based on the premise that, in addition to inducing genetic and epigenetic aberrations, tobacco can contribute to oral carcinogenesis through disrupting tongue microbiome. Indeed, the study found, after adjustment for multiple comparisons, 6 species to be enriched in the tongue microbiome of SU, namely R. mucilaginosa, Streptococcus sp. oral taxon 66, A. meyeri, S. vestibularis, S. sanguinis and a potentially novel Veillonella species.

Enrichment of R. mucilaginosa is particularly relevant. This species has been found to be significantly more abundant in tongue leukoplakia lesions compared to contralateral side, and to tongue swab samples collected from healthy controls [16]. More importantly, most strains of R. mucilaginosa has been recently found to produce high levels of acetaldehyde from ethanol comparable to that of Candida and Neisseria spp., and to lack genes encoding acetaldehyde dehydrogenases- a group of enzymes that detoxify acetyl aldehyde [17]. Acetaldehyde is well known carcinogenic compound, and its production has been proposed as a mechanism by which bacteria can contribute to oral and gastrointestinal carcinogenesis [18, 19]. Interestingly, Rothia mucilaginosa have been reported to be reduced in established oral cancer lesions [20, 21], suggestive of a possible role only in early stages.

Streptococci are also known to produce acetaldehyde, although there are variations among the different species, with Streptococcus salivarius, Streptococcus intermedius and Streptococcus mitis having the highest acetaldehyde-producing potential [22]. Indeed S. mitis, was among the species identified by LEfSe analysis in this study as overabundant among the SU, but it did not stand adjustment for multiple comparisons. No information is available in the literature about acetaldehyde-producing potential of the other three Streptococcus species found here to be enriched in SU, except S. Sanguinis that has been shown to produce relatively smaller amounts of acetaldehyde but to encode non-functional acetaldehyde dehydrogenase genes [23]. The acetaldehyde-producing abilities of S. vestibularis and Streptococcus sp. oral taxon 66 needs to be experimentally assessed.

Based on the above, it may be hypothesized that use of smokeless tobacco (probably tobacco in general) indirectly contributes to initiation of oral cancer by enrichment of acetaldehyde-producing bacterial species. This may be particularly relevant in people who consume alcohol in addition to using tobacco. In fact, this could be one mechanism underlying the known interaction between alcohol and tobacco in head and neck cancers [24]. The hypothesis is supported by the fact that healthy individuals without clinically detectable lesions were recruited to this study, and thus the changes observed represent early events prior to development of malignancy. This model, however, needs to be validated in future studies.

The tongue microbiome has not been extensively explored, with the few available studies focusing on how it relates to halitosis [14, 25, 26]. This is, thus, probably the first study to assess the effect of tobacco in general, and ST in particular, on the tongue microbiome within the context of oral carcinogenesis. The study, however, has some limitations to note. First, the variation in effect of shammah use by gender could not be assessed due to exclusion of females. This, however, was because it was difficult to recruit females who are willing to report using shammah due to the social stigma associated with using it. Secondly, and despite all efforts made, it was extremely difficult to exclude qat chewing due to the strong association between the two habits, so some of the shammah users recruited were also qat chewers, which may have confounded the results. Nevertheless, analysis of the data after exclusion of these cases did not change the results significantly. There were also cigarette smokers among the two groups, but their distribution did not significantly differ between the two groups.

# Conclusion

The present study identified changes shifts in tongue microbiome in association with shammah use, namely enrichment of Rothia and Streptococcus species which are known to produce high levels of acetaldehyde. Further studies are needed to validate these findings and to explore their relevance to oral carcinogenesis and, possibly, the interaction between tobacco and alcohol.

# Methods

### *Study design and subjects*

The objective of this study was to explore the potential effect of chronic use of shammah, as an example of ST product, on the tongue microbiome in comparison to the tongue microbiome of shammah non-users. The study was conducted during the academic years 2018/2019 and 2019/2020. Participants were recruited to this cross-sectional study from among attendants of the dental clinics at the College of Dentistry, Jazan University. They had to be 20-40 years old and systemically healthy (as self-reported). Due to the difficulty recruiting females who use shammah, the study was limited to male subjects. Shammah users (SU) were defined as those who used shammah daily for at least one year without a period of cessation, while shammah non-users (SNU) were required to have no history of shammah use. Subjects with moderate to severe gingivitis or periodontitis, or who had a history of antibiotic, antifungal or steroids use, or periodontal treatment, including prophylaxis, in the last 3 months were excluded. Patient's demographic data were obtained using a structured interview questionnaire. Clinical examination included assessment of bleeding on probing, periodontal pocket depth and dental caries using Decay, Missing Filling index for Teeth (DMFT).

The study was approved by the Scientific Research Ethics Committee, Jazan University (REC39/2-432), and was conducted in compliance with the Helsinki Declaration on medical research involving human subjects. Informed consents were obtained from all the participants.

### *Tongue scrapping and DNA extraction*

The study subjects were instructed not to eat, drink, or smoke at least an hour before sample collection. Tongue scrapping samples were collected as follows. The tongue was isolated using sterilized cotton rolls and gauze, and the participant was asked to protrude his tongue forward as much as he could. The tongue was stabilized by holding the tip with a piece of sterile gauze and its surface was dried with another piece. Using a sterilized metal spatula, the dorsal surface of the tongue was scrapped with overlapping strokes starting posteriorly all the way to the tip. The collected scrapping was then transferred with a paper point into an Eppendorf tube containing 600 µl Tris-EDTA buffer (pH 8.0) and stored at – 20 °C.

Prior to DNA extraction, the samples were thawed, vortexed vigorously, and centrifuged at 14,000 rpm for one minute (Micro 120, Hettich Zentrifuge, Germany) to pellet the cells. The supernatant was decanted and the pellet washed once with 500 µl phosphate-buffered saline, suspended in 180 µl of lysozyme solution (20 mg/ml), and incubated overnight at 37 °C. DNA was then extracted using PureLink™

Genomic DNA Mini Kit (Invitrogen, USA) according the manufacturer's instructions, using an elution volume of 100 µl. The quantity of DNA was assessed by Jenway Genova Nano 3-in-1 Spectrophotmeter (Jenway®, UK). The extracts were stored at −20°C for subsequent analysis.

### *16S sequencing and bioinformatic analysis*

Library preparation and sequencing of the 16S rRNA gene were performed at the Australian Center for Ecogenomics (Brisbane, Australia) as described elsewhere [27]. In brief, the degenerate primers 27FYM [28] and 519R [29] were employed to amplify the V1-3 region in standard PCR conditions. The generated amplicons (~ 520 bp) were purified and, in a second PCR, tagged with 8-base barcodes. The libraries were then pooled in equimolar concentrations and sequenced on a MiSeq (Illumina, USA) using the v3 2x300 bp chemistry, with a minimum sequencing depth of 30,000 reads per sample.

The raw data were preprocessed, including merging of reads, primer-trimming, quality-filtration, alignment and chimera removal, as detailed previously [21]. The resultant high quality, merged reads were classified using our species-level taxonomy assignment algorithm, described in details elsewhere [21, 30]. Downstream analysis of microbial profiles including subsampling, generation of taxonomy plots/tables and rarefaction curves, and calculation of species richness, coverage, alpha diversity indices and beta diversity distance matrices, was performed with QIIME (Quantitative Insights Into Microbial Ecology) software package version 1.9.1 [31]. Principal component analysis (PCoA) was used to cluster samples based on microbial similarity. Differentially abundant taxa were identified with linear discriminant analysis (LDA) effect size (LEfSe) [32].

## Abbreviations

SU
Shammah users
SNU
Sammah non-users
ST
Smokeless tobacco
DMFT
Decay, Missing Filling index for Teeth
FDR
False discovery rate
R. mucilaginosa
Rothia mucilabinosa
A. meyeri
Actinomyces meyeri
S. vestibularis
Streptococcus vestibularis

S. sanguinis
Streptococcus sanguinis
S. mitis
Streptococcus mitis
µl
Microliter
DNA
Deoxyribonucleic acid
rpm
Round per minute
°C
Degree Celsius
mg/ml
Milligram per milliliter
PCR
Polymerase chain reaction
bp
base pair
QIIME
Quantitative Insights Into Microbial Ecology
PCoA
Principal component analysis
LDA
linear discriminant analysis
LEfSe
linear discriminant analysis (LDA) effect size

# Declarations

### *Ethics approval and consent to participate*

The study was approved by the Scientific Research Ethics Committee, Jazan University (REC39/2-432), and was conducted in compliance with the Helsinki Declaration on medical research involving human subjects. Informed consents were obtained from all the participants.

### *Consent for publication*

Not applicable.

### *Availability of data and material*

The dataset supporting the conclusions of this article is available from NCBI's Sequence Read Archive, [PRJNA605810; http://www.ncbi.nlm.nih.gov/bioproject/605810]

### *Competing interests*

The authors declare that they have not any kinds of conflict of interest.


### *Funding*

The study was funded by The Deanship of Scientific Research, Jazan University through The Eighth Scientific Research Program (Number: JUP//000261).


### *Authors' contributions*

EH contributed to conception and design, data acquisition and analysis and drafted the manuscript. MA and HH contributed to data acquisition. AA contributed to conception and design. DB and TC contributed to data analysis. NA contributed to conception and design, data analysis and interpretation and critically revised the manuscript. All authors revised the draft and approved the final version of the manuscript.

### *Authors' information*


***Esam Halboub, Abdulwahhab H. Al-Amir, and Husham E. Homeida***, Department of Maxillofacial Surgery and Diagnostic Sciences, College of Dentistry, Jazan University, Jazan, Saudi Arabia. ***Mohammed Alakhali***, Department of Preventive Dentistry, College of Dentistry, Jazan University, Jazan, Saudi Arabia. ***Divyashri Baraniya*** and ***Nezar Noor Al-Hebshi***, Oral Microbiome Research Laboratory, Department of Oral Health Sciences, Maurice H. Kornberg School of Dentistry, Temple University, Philadelphia, PA, USA. ***Tsute Chen***, Department of Microbiology, Forsyth Institute, Cambridge, MA, USA.


# References


1. World Health Organization, *WHO report on the global tobacco epidemic: Offer help to quit tobacco use*. 2019, World Health Organization. Available from: https://www.who.int/tobacco/global_report/en//: Geneva, Switzerland.
2. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans and International Agency for Research on Cancer., *Smokeless tobacco and some tobacco-specific N-nitrosamines*, in *A review of human carcinogens: Personal habits and indoor combustions, Volume 100 E*. 2012, World Health Organization; distributed by WHO Press: Lyon, France; Geneva. p. 267-321.
3. Petti, S., M. Masood, and C. Scully, *The magnitude of tobacco smoking-betel quid chewing-alcohol drinking interaction effect on oral cancer in South-East Asia. A meta-analysis of observational studies.* PLoS One, 2013. **8**(11): e78999.
4. Khan, Z., J. Tonnies, and S. Muller, *Smokeless tobacco and oral cancer in South Asia: a systematic review with meta-analysis.* J Cancer Epidemiol, 2014. **2014**: 394696.

5. Tomar, S.L., S.S. Hecht, I. Jaspers, R.L. Gregory, and I. Stepanov, *Oral Health Effects of Combusted and Smokeless Tobacco Products.* Adv Dent Res, 2019. **30**(1): 4-10.
6. Jiang, X., J. Wu, J. Wang, and R. Huang, *Tobacco and oral squamous cell carcinoma: A review of carcinogenic pathways.* Tob Induc Dis, 2019. **17**: 29.
7. Al-Hebshi, N.N., W.S. Borgnakke, and N.W. Johnson, *The Microbiome of Oral Squamous Cell Carcinomas: a Functional Perspective.* Current Oral Health Reports, 2019. **6**: 145–160.
8. Perera, M., N.N. Al-Hebshi, D.J. Speicher, I. Perera, and N.W. Johnson, *Emerging role of bacteria in oral carcinogenesis: a review with special reference to perio-pathogenic bacteria.* J Oral Microbiol, 2016. **8**: 32762.
9. Wu, J., B.A. Peters, C. Dominianni, Y. Zhang, Z. Pei, L. Yang, et al., *Cigarette smoking and the oral microbiome in a large study of American adults.* ISME J, 2016. **10**(10): 2435-46.
10. Yang, Y., W. Zheng, Q.Y. Cai, M.J. Shrubsole, Z. Pei, R. Brucker, et al., *Cigarette smoking and oral microbiota in low-income and African-American populations.* J Epidemiol Community Health, 2019. **73**(12): 1108-1115.
11. Jin, J., L. Guo, L. VonTungeln, M. Vanlandingham, C.E. Cerniglia, and H. Chen, *Smokeless tobacco impacts oral microbiota in a Syrian Golden hamster cheek pouch carcinogenesis model.* Anaerobe, 2018. **52**: 29-42.
12. Patil, S., A.W.H. Alamir, G. Arakeri, K.H. Awan, S.H. Bhandi, A. Aljabab, et al., *The relationship of shammah (Arabian snuff) chewing to the risk of oral cancer and oral potentially malignant disorders.* J Oral Pathol Med, 2019. **48**(6): 425-432.
13. Quadri, M.F.A., S.K. Tadakamadla, and T. John, *Smokeless tobacco and oral cancer in the Middle East and North Africa: A systematic review and meta-analysis.* Tob Induc Dis, 2019. **17**: 56.
14. Ren, W., Z. Xun, Z. Wang, Q. Zhang, X. Liu, H. Zheng, et al., *Tongue Coating and the Salivary Microbial Communities Vary in Children with Halitosis.* Sci Rep, 2016. **6**: 24481.
15. Nasher, A.T., N.N. Al-Hebshi, E.E. Al-Moayad, and A.M. Suleiman, *Viral infection and oral habits as risk factors for oral squamous cell carcinoma in Yemen: a case-control study.* Oral Surg Oral Med Oral Pathol Oral Radiol, 2014. **118**(5): 566-572 e1.
16. Amer, A., S. Galvin, C.M. Healy, and G.P. Moran, *The Microbiome of Potentially Malignant Oral Leukoplakia Exhibits Enrichment for Fusobacterium, Leptotrichia, Campylobacter, and Rothia Species.* Front Microbiol, 2017. **8**: 2391.
17. Abdrazak, A., W. Aine, N.N. Al-Hebshi, C.M. Healy, and G.P. Moran, *Acetaldehyde production by Rothia mucilaginosa isolates from patients with oral leukoplakia.* Journal of Oral Microbiology, 2020. **Accepted Manuscript**.
18. Homann, N., *Alcohol and upper gastrointestinal tract cancer: the role of local acetaldehyde production.* Addict Biol, 2001. **6**(4): 309-323.
19. Meurman, J.H., *Oral microbiota and cancer.* J Oral Microbiol, 2010. **2**.

20. Perera, M., N.N. Al-Hebshi, I. Perera, D. Ipe, G.C. Ulett, D.J. Speicher, et al., *Inflammatory Bacteriome and Oral Squamous Cell Carcinoma.* J Dent Res, 2018. **97**(6): 725-732.
21. Al-Hebshi, N.N., A.T. Nasher, M.Y. Maryoud, H.E. Homeida, T. Chen, A.M. Idris, et al., *Inflammatory bacteriome featuring Fusobacterium nucleatum and Pseudomonas aeruginosa identified in association with oral squamous cell carcinoma.* Sci Rep, 2017. **7**(1): 1834.
22. Kurkivuori, J., V. Salaspuro, P. Kaihovaara, K. Kari, R. Rautemaa, L. Gronroos, et al., *Acetaldehyde production from ethanol by oral streptococci.* Oral Oncol, 2007. **43**(2): 181-6.
23. Pavlova, S.I., L. Jin, S.R. Gasparovich, and L. Tao, *Multiple alcohol dehydrogenases but no functional acetaldehyde dehydrogenase causing excessive acetaldehyde production from ethanol by oral streptococci.* Microbiology, 2013. **159**(Pt 7): 1437-1446.
24. Hashibe, M., P. Brennan, S.C. Chuang, S. Boccia, X. Castellsague, C. Chen, et al., *Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium.* Cancer Epidemiol Biomarkers Prev, 2009. **18**(2): 541-50.
25. Seerangaiyan, K., A.J. van Winkelhoff, H.J.M. Harmsen, J.W.A. Rossen, and E.G. Winkel, *The tongue microbiome in healthy subjects and patients with intra-oral halitosis.* J Breath Res, 2017. **11**(3): 036010.
26. Ye, W., Y. Zhang, M. He, C. Zhu, and X.P. Feng, *Relationship of tongue coating microbiome on volatile sulfur compounds in healthy and halitosis adults.* J Breath Res, 2019. **14**(1): 016005.
27. Al-Hebshi, N.N., F.A. Alharbi, M. Mahri, and T. Chen, *Differences in the bacteriome of smokeless tobacco products with different oral carcinogenicity: compositional and predicted functional analysis.* Genes, 2017: in press.
28. Frank, J.A., C.I. Reich, S. Sharma, J.S. Weisbaum, B.A. Wilson, and G.J. Olsen, *Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes.* Appl Environ Microbiol, 2008. **74**(8): 2461-70.
29. Lane, D.J., B. Pace, G.J. Olsen, D.A. Stahl, M.L. Sogin, and N.R. Pace, *Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses.* Proc Natl Acad Sci U S A, 1985. **82**(20): 6955-9.
30. Al-Hebshi, N.N., A.T. Nasher, A.M. Idris, and T. Chen, *Robust species taxonomy assignment algorithm for 16S rRNA NGS reads: application to oral carcinoma samples.* J Oral Microbiol, 2015. **7**: 28934.
31. Caporaso, J.G., J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, et al., *QIIME allows analysis of high-throughput community sequencing data.* Nat Methods, 2010. **7**(5): 335-6.
32. Segata, N., J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W.S. Garrett, et al., *Metagenomic biomarker discovery and explanation.* Genome Biology, 2011. **12**(6).
33. Parks, D.H., G.W. Tyson, P. Hugenholtz, and R.G. Beiko, *STAMP: statistical analysis of taxonomic and functional profiles.* Bioinformatics, 2014. **30**(21): 3123-4.

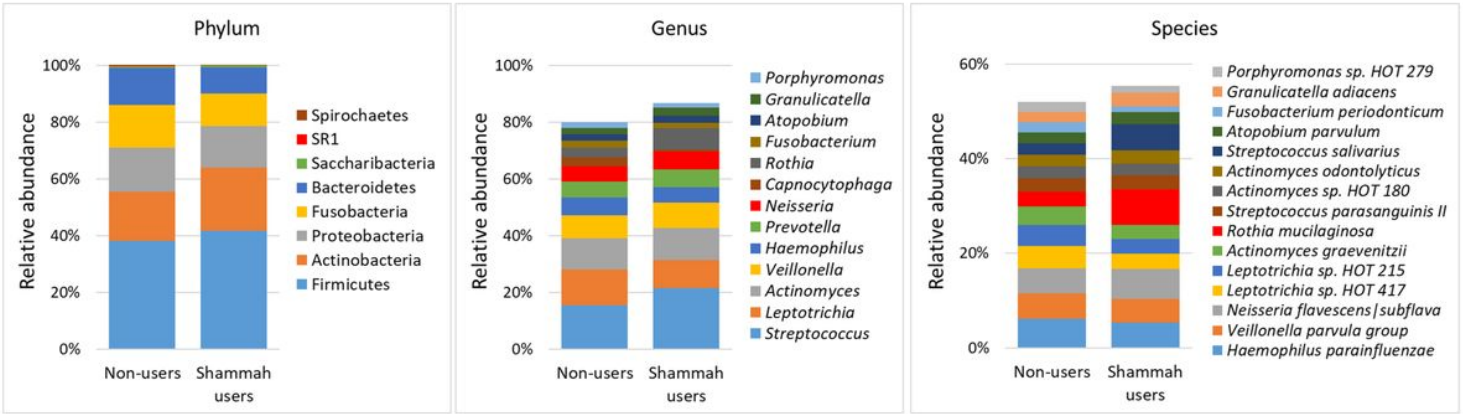
# Table

**Table 1.** Characteristics of the study groups described as mean±SD or number (%) as appropriate

| Variable | Shammah Users (n = 29) | Shammah non-users (n = 23) | P value* |
|---|---|---|---|
| Age | 27.34±6.9 | 27.7±7.19 | 0.859 |
| Education | | | 0.443 |
| Illiterate | 5 (17.2) | 2 (8.7) | |
| Primary | 5 (17.2) | 5 (21.7) | |
| Secondary | 16 (51.7) | 9 (39.1) | |
| University | 4 (13.8) | 7 (30.4) | |
| Type of Shammah | | NA | NA |
| White | 12 (41.4) | | |
| Black | 17 (58.6) | | |
| Duration of shammah use (years) | 8.66±7.11 | NA | NA |
| Frequency of shammah use per day | 11.86±4.43 | NA | NA |
| Qat Chewing | | | |
| Yes | 19 (65.5) | 2 (8.7) | < 0.001 |
| No | 10 (34.5) | 21 (91.3) | |
| Frequency of qat chewing per week† | 2.11±0.81 | 1.5±0.71 | 0.324 |
| Smoking | | | |
| Yes | 7 (24.1) | 3 (13) | 0.482 |
| No | 22 (75.9) | 20 (87) | |
| Cigarette/Day | 7.43±3.46 | 5.33±4.04 | 4.25 |
| DMFT | 5.41±4.38 | 5.78±4.11 | 0.758 |
| Decay | 4.45±3.67 | 3.57±2.9 | 0.350 |
| Missing | 0.52±0.91 | 0.61±0.89 | 0.718 |
| Filling | 0.76±1.46 | 1.61±2.62 | 0.173 |

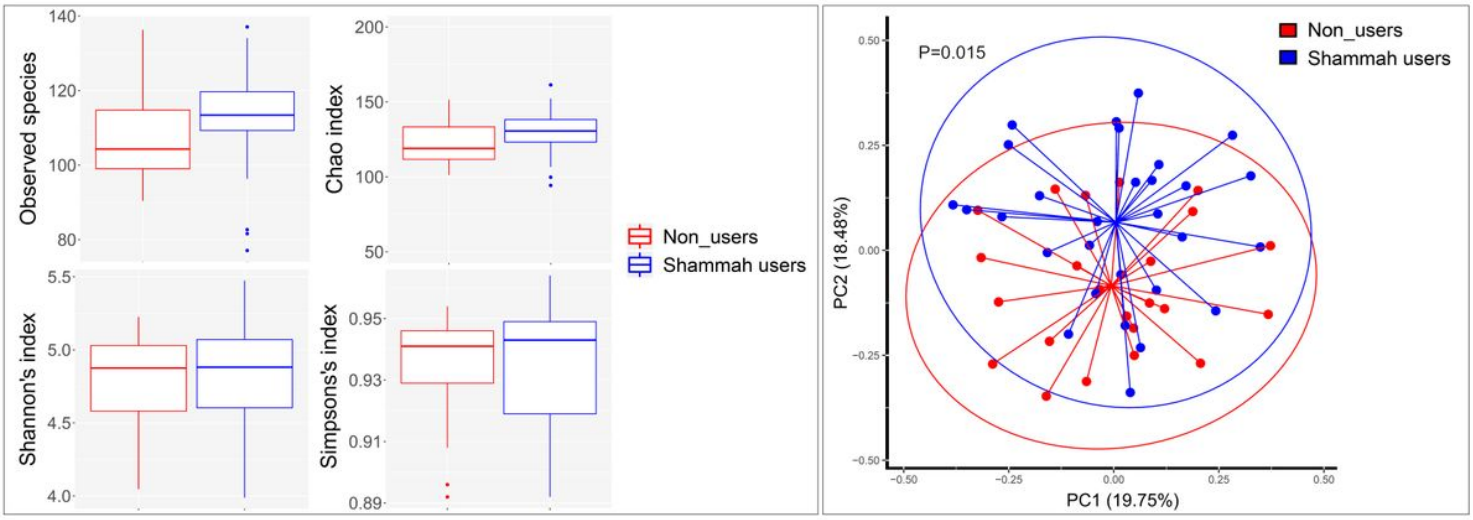
* Chi-squared or Student's t-test as appropriate. †: n = 19 and 2 for Shammah Users and Shammah Non-Users, respectively. : n = 7 and 3 for Shammah Users and Shammah Non-Users, respectively.

# Figures



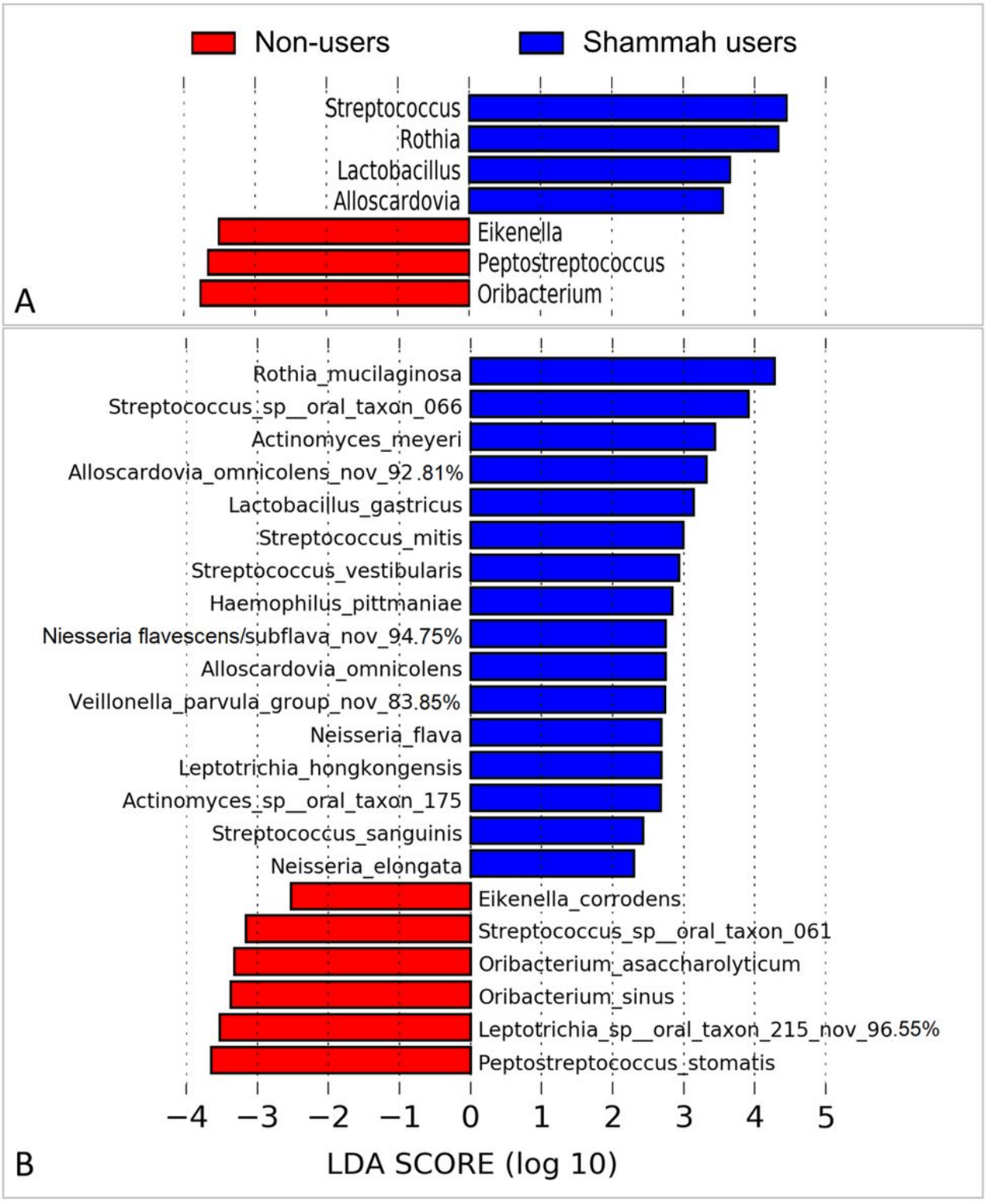

## Figure 2

Microbiological profiles. DNA extracted from tongue scrapings was sequenced for the V1-V3 region of the 16S rRNA gene using paired-end chemistry. The generated reads were merged, quality-filtered and classified to the species level using a BLASTn-based algorithm. The stacked bars show the average relative abundances of (a) all phyla, (b) top genera and (c) top species identified in the study groups.



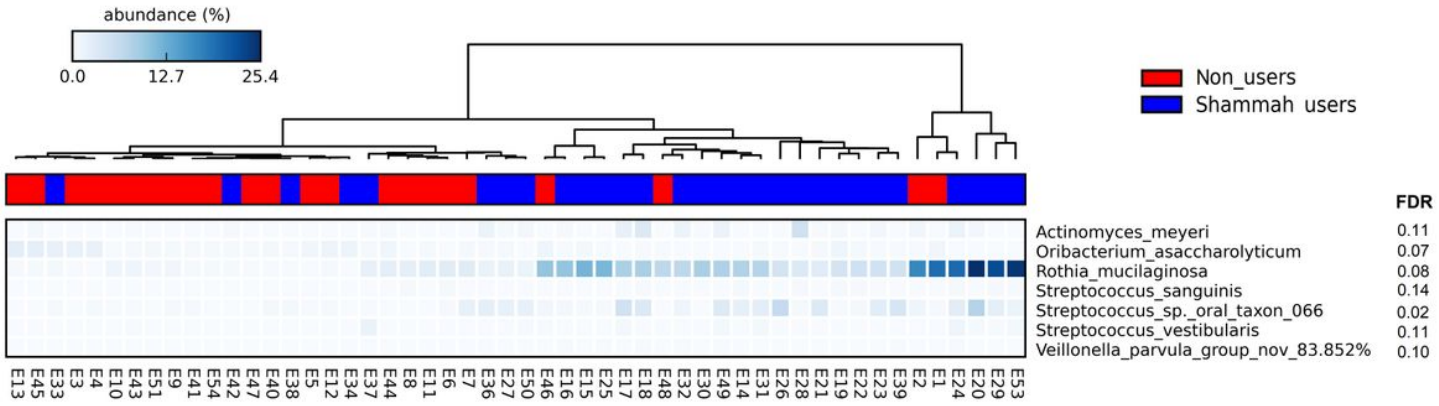

## Figure 4

Species richness and diversity. Taxonomic profiles were rarified and used to calculate observed richness, expected richness (Chao index), alpha diversity indices (Shannon's and Simpson's) and distance matrices employing standard QIIME scripts. Left: Box and whisker plots of species richness and aloha diversity in each group. Differences were not significant by Mann–Whitney U test. Right: clustering of samples with PCoA based on Bary-Curtis distance matrix. Significance of separation was assessed with ANOSIM. Plots were generated with R Package.



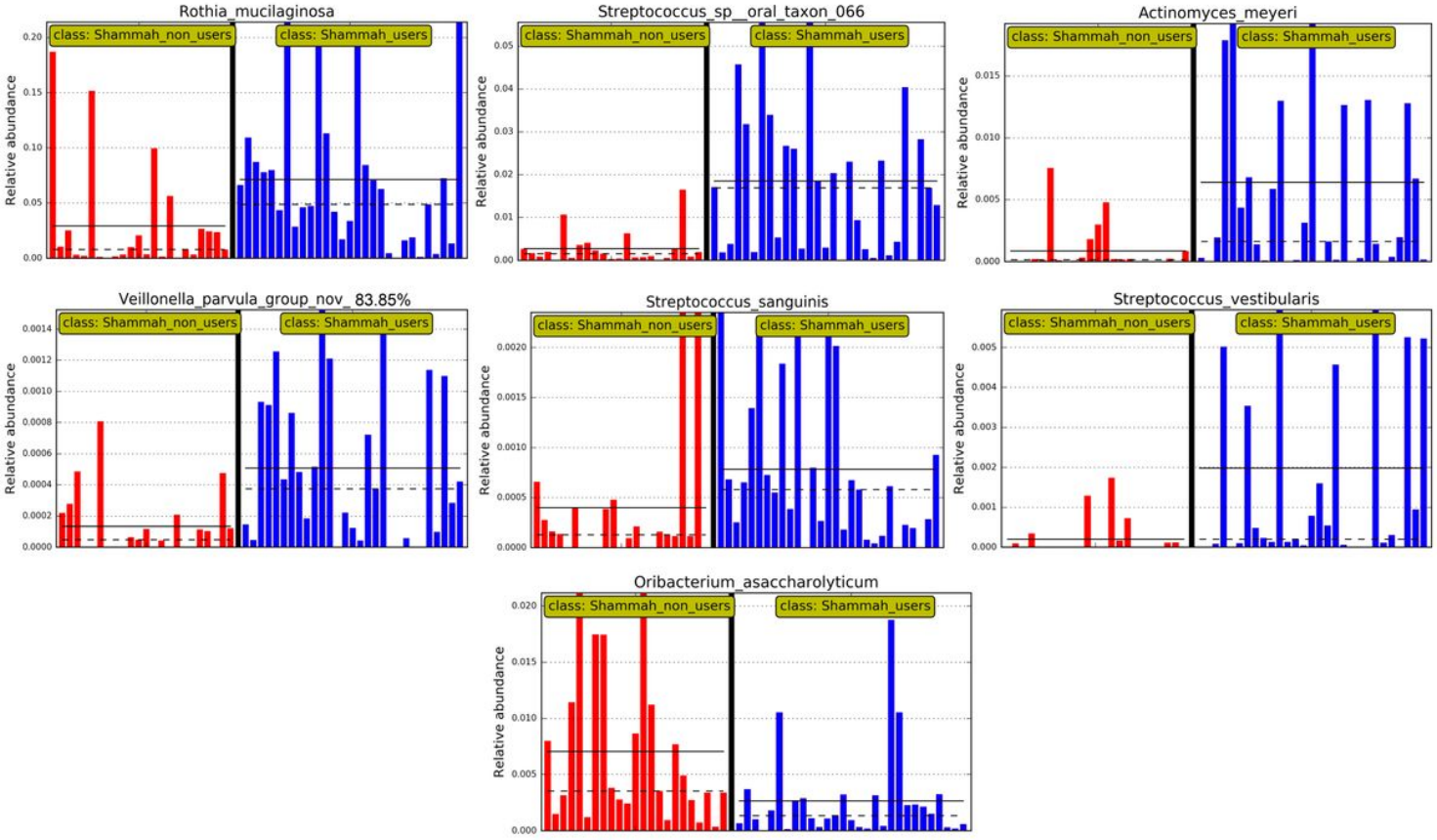

## Figure 5

Differentially abundant taxa. (a) Genera and (b) species that showed significant differences in relative abundance between the two study groups as identified by linear discriminant analysis (LDA) effect size analysis (LEfSe).




## Figure 8

Hierarchical centroid clustering. Samples were clustered based on the relative abundances of differentially abundant species with false discovery rate (FDR) ≤ 0.2. Clustering and plotting were performed with STAMP (statistical analysis of taxonomic and functional profiles) [33].




## Figure 9

Per sample abundance plots. Relative abundances of differentially abundant species with false discovery rate (FDR) ≤ 0.2 in individual samples.

# Supplementary Files

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

- S2.AbundancesandfrequenciesPhylumlevel.xlsx
- S1.Readsstatistics.xlsx
- S4.Abundancesandfrequenciesspecieslevel.xlsx
- SF1.LEfSeanalysisqatchewersexcluded.tif
- S3.AbundancesandfrequenciesGenuslevel.xlsx
- S4.Abundancesandfrequenciesspecieslevel.xlsx
- S1.Readsstatistics.xlsx
- S2.AbundancesandfrequenciesPhylumlevel.xlsx
- S3.AbundancesandfrequenciesGenuslevel.xlsx
- SF1.LEfSeanalysisqatchewersexcluded.tif