HUMAN MICROBIOME

Characterization of Oral Bacterial Composition of Adult Smokeless Tobacco Users from Healthy Indians Using 16S rDNA Analysis

Ankita Srivastava[1] · SukhDev Mishra[2] · Digvijay Verma[1] (ORCID)

Abstract

The present investigation is aiming to report the oral bacterial composition of smokeless tobacco (SLT) users and to determine the influence of SLT products on the healthy Indian population. With the aid of the V3 hypervariable region of the 16S rRNA gene, a total of 8,080,889 high-quality reads were clustered into 15 phyla and 180 genera in the oral cavity of the SLT users. Comparative analysis revealed a more diverse microbiome where two phyla and sixteen genera were significantly different among the SLT users as compared to the control group (p-value < 0.05). The prevalence of *Fusobacteria*-, *Porphyromonas*-, *Desulfobulbus*-, *Enterococcus*-, and *Parvimonas*-like genera among SLT users indicates altered bacterial communities among SLT users. Besides, the depletion of health-compatible bacteria such as *Lactobacillus* and *Haemophilus* also suggests poor oral health. Here, the majority of the altered genera belong to Gram-negative anaerobes that have been reported for assisting biofilm formation that leads in the progression of several oral diseases. The PICRUSt analysis further supports the hypothesis where a significant increase in the count of the genes involved in the metabolism of nitrogen, amino acids, and nicotinate/nicotinamide was observed among tobacco chewers. Moreover, this study has a high significance in Indian prospects where the SLT consumers are prevalent but we are deficient in information on their oral microbiome.

Keywords Smokeless tobacco · Oral microbiome · Bacterial ecology · Next-generation sequencing

Introduction

Approximately 356 million smokeless tobacco (SLT) users have been estimated in 140 countries that accounts for one-fourth of the total tobacco consumers [1]. Smokeless tobacco is one of the deadliest reasons for 1.5 million disability-adjusted life years (DALYs) among 113 countries [2]. The prevalence of SLT users is more in lower and lower-middle-income countries like India [1]. The factsheet of the World Health Organization (WHO) states that SLT consumption accounts for 48% of cardiovascular diseases, 23% of chronic respiratory diseases, 14% CMNNDs (communicable, maternal, neonatal, and nutritional diseases), and 10% cancer [3]. The existence of approximately 4000 chemical constituents in smokeless tobacco and derived products makes it a risky chewing/smoking product [4]. Several of these chemicals are derivatives of toxicants and carcinogens that include tobacco-specific nitrosamines (TSNAs), *N*–nitrosamino acids, volatile aldehydes, polyaromatic hydrocarbons (PAHs), metals, and metalloids. Besides, smokeless tobacco also exhibits a huge microbial load of bacteria, fungi, molds, and other microorganisms [5]. Han et al (2015) report that bacterial load on various SLT ranges from 8.33×10^1 to 2.54×10^5 CFU/g SLT [6], where numerous opportunistic pathogens (*Eubacterium*, *Porphyromonas*, and *Prevotella*) have also been documented. Several bacterial residents of smokeless tobacco assist in the formation of tobacco-specific nitrosamines (TSNAs) [7] as well as acetaldehyde production, a carcinogen [8]. The International Agency for Research on Cancer (IARC) has identified five types of TSNAs and categorized them into group I carcinogens [9].

Oral cavity encounters first with SLT products along with their inhabitant microorganisms and thus alters the overall bacterial composition [10]. Dysbiotic microbiota of the oral cavity has been reported for the onset of several diseases including the cancers of the oropharynx [11, 12], where SLT

✉ Digvijay Verma
digvijay.udsc@gmail.com

1 Department of Microbiology, Babasaheb Bhimrao Ambedkar University, Lucknow, India

2 Department of Bio-Statistics and Data Management, ICMR-National Institute of Occupational Health, Ahmedabad, India

products may further influence the overall bacterial communities.

SLT users are prevalent in India and its consumption leads to one million deaths every year in India [3]. We have limited information on oral microbiota from the Indian community [13], where we have no concrete data on the oral microbiome of SLT consumers. Therefore, the present investigation has been carried out to determine the oral bacterial composition of SLT users from a set of populations of north India and to assess the alteration of bacterial communities due to the SLT consumption as compared to the control group. Our findings revealed significant alterations in the oral bacterial communities among SLT users at various taxon levels (2 phyla and sixteen genera; p-value <0.05) as compared to the non-tobacco chewers. Besides, the phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt) analysis also showed an apparent shift in the functions, especially the genes involved in the metabolism of nitrogen, amino acids, and nicotinate/nicotinamide. The investigation will pave the way to determine the core oral microbiome of SLT users and to understand the influence of smokeless tobacco on a healthy oral microbiome especially from a developing country like India, where the SLT consumers account for 66.6% (approx. 237.4 million) of the total SLT consumers (356 million) among 140 countries [1].

Materials and Methods

Sample Collection, DNA Extraction, and Sequencing

Oral wash samples were collected from SLT users (n=20; code: HTC) and SLT non-users (n =20; code: HI) after their prior consents. The SLT users were in tobacco consumption habits for more than 5 years having 15–25 g of SLT product intake in a week. The protocol of the study was approved by the Institutional Ethical Committee, Netaji Subhas University of Technology, New Delhi (NSUT), as well as the All India Institute of Medical Sciences (AIIMS), New Delhi. None of the participants (HI and HTC) was alcoholic and was not on any medication/antibiotic for the last 3 months at the time of sample collection (Table 1). The health status of the participants was detected as healthy in terms of their diabetic status, systolic BP, and BMI. Besides, none of the participants have any lesions in their oral cavities. Metagenomic DNA extraction, V3 hypervariable region amplification, sequencing, and processing of reads have been carried out as mentioned in our recent publication [12]. Briefly, the metagenomic DNA was extracted using prior treatment with proteinase-K (10 mg/ml) and lysozyme (10 mg/ml) followed by the Qiagen DNeasy Blood and Tissue Kit (Qiagen, USA). The amplicon library of the V3 hypervariable region was prepared and sequenced at the Illumina MiSeq 2500 platform (Illumina, San Diego,

USA) (C:\Users\pc\Desktop\HTC NGS sequences). The sequences having a Phred score value >30 were processed using different filters such as conserved region, spacer, mismatch, and consensus read to obtain the quality reads. ClustalO program [14] and UCHIME [15] of USEARCH [16] were other programs used for downstream processing of the sequences.

OTU Generation, Taxonomy Assignment, and OTU Clustering

The QIIME program was used for generating operational taxonomic units (OTUs) at the similarity cutoff of 0.97 from the pre-processed sequences (QIIME version 1.9.1). The taxonomy assignment was carried out using the reference of the Greengenes database with the assistance of the PyNAST program.

Analysis of Diversity Indices, Data Interpretation, and the PICRUSt Analysis

Various diversity indices (Chao1, Shannon, Simpson, PD whole_tree, and Good's estimator Shannon and Simpson indices) were calculated to analyze the diversity of the bacteria in each sample [17] (Table 2). The data have been interpreted for different outputs such as heatmap and correlation graphs using CAMEO analysis tool, (https://github.com/avishekdutta14/CAMEO), SPSS V26, NCSS V20.0.2, and STAMP V2.1.3 [18]. The functional prediction of the 16S rDNA genes was performed using the PICRUSt algorithm version 1.1.0 [19].

Interpretation of the Data and Statistical Analysis

The generated OTUs were analyzed for the overall bacterial composition in the healthy oral microbiome of SLT users and non-users at the various taxon level. An attempt was made to determine the influence of tobacco on overall bacterial composition as well as their predicted functions among SLT users as compared to the SLT non-users. The data were evaluated for the assumption of normality. In the case of departure from normality, the equality of means between the groups was compared using Mann Whitney U test at the level of phylum and genera. A General Linear Model procedure was applied to assess the impact of tobacco on microbial diversity of subjects adjusting for categorical predictors such as sex and diet including interaction terms. Principal coordinate analysis (PCoA) plots were generated using the first two principal coordinates and labeled according to smoking status using the Euclidean distance matrix. The statistically significant differences at the level of tier I, tier II, and tier III were performed using a two-sided G-test (w/Yates)+Fisher's test at a 95% confidence interval to analyze the PICRUSt output using STAMP.

Table 1 Various parameters of the healthy smokeless tobacco users during this investigation

Sample ID*	Parameters[#] Gender	Food habit	Age	Health status* Place[$]	Tobacco	Diabetic status	Systolic BP mean	BMI
HTC-1	Male	Non-vegetarian	42	New Delhi	Khaini/Guthkha	No	132	22.3
HTC-2	Male	Non-vegetarian	45	New Delhi	Khaini	No	134	25.6
HTC-3	Male	Vegetarian	40	New Delhi	Loose leaves/Khaini	No	122	28.9
HTC-4	Male	Vegetarian	40	New Delhi	Loose leaves/Kuber	No	146	21.6
HTC-5	Male	Non-vegetarian	24	New Delhi	Khaini/Surti	No	118	27.7
HTC-6	Male	Non-vegetarian	24	New Delhi	Khaini/Guthkha	No	112	24.8
HTC-7	Male	Non-vegetarian	46	New Delhi	Khaini	No	138	29.1
HTC-8	Male	Vegetarian	40	New Delhi	Khaini	No	126	24.6
HTC-9	Male	Vegetarian	37	New Delhi	Gutkha/Khaini	No	134	26.9
HTC-10	Female	Non-vegetarian	52	New Delhi	Khaini/Loose leaves	Yes	116	27.3
HTC-11	Male	Vegetarian	41	New Delhi	Khaini	No	118	24.9
HTC-12	Female	Vegetarian	35	New Delhi	Loose leaves/Gutkha/Khaini	No	120	23.8
HTC-13	Male	Non-vegetarian	50	New Delhi	Khaini/Loose leaves	No	132	26.6
HTC-14	Male	Non-vegetarian	28	New Delhi	Loose leaves/Gutkha/Khaini	No	128	22.9
HTC-15	Male	Non-vegetarian	35	New Delhi	Khaini	No	124	22.6
HTC-16	Male	Non-vegetarian	58	New Delhi	Loose leaves/Gutkha/Kuber	No	128	21.9
HTC-17	Male	Vegetarian	60	New Delhi	Loose leaves/Gutkha/Khaini	No	132	27.3
HTC-18	Female	Vegetarian	54	New Delhi	Khaini/loose leaves/Guthkha	Yes	134	27.1
HTC-19	Female	Vegetarian	51	New Delhi	Khaini/Loose leaves	No	112	26.3
HTC-20	Male	Non-vegetarian	55	New Delhi	Loose leaves/Gutkha/Khaini	No	124	25.3

*Systolic BP and BMI of SLT non-users were in the range of 110–145 and 20–28 respectively. All SLT non-users (HI) were non-diabetic and devoid of any lesions in their oral cavities

[#] The age of SLT non-users (HI) was in the range of 21–57 having both vegetarian and non-vegetarian male and female participants

[$] Participants of both the groups belong to same demographic region. The all parameters of the SLT non-users (HI) are mentioned in Supplementary table 1 (Verma et al., 2020)

Results

FastQC Analysis and Pre-processing of the Reads

The paired-end amplicon sequencing from twenty different samples revealed 8,080,889 high-quality reads (~680 Mb) which were converted into 86,270 OTUs on passing through four different filters along with the removal of consensus reads and chimeric sequences, and singletons (Supplementary table 1, 2, 3, and 4). It could be due to the use of QIIME 1.9.1-UCLUST that usually generates a large number of counterfeit OTUs [20]. However, UCLUST offers the fastest runtime with reasonably good performance; therefore, it can be used for microbiome analysis [21].

Bacteriome of SLT Users

Overall Phyla On taxonomic assignment, the generated OTUs were assigned into 15 phyla, 23 classes, 46 orders, 99 families, and 180 genera. Overall analysis revealed that Proteobacteria (25.836%) dominates over Firmicutes (23.844%) among SLT users followed by Bacteroidetes (19.664%), Fusobacteria

(3.222%), Actinobacteria (5.392 %), Candidate division _GN02 (0.081 %), Spirochaetes (0.363 %), Candidate division _SR1 (0.3059 %), and Candidate division _TM7 (0.197 %). Other small contributors of the SLT user's oral cavity were Tennericutes (0.0789%), Synergistetes (0.033%), Cyanobacteria (0.262 %), Chloroflexi (0.0016%), Elusimicrobia (0.001%), and Acidobacteria (Fig. 1a).

Overall Genera The taxonomic assignment at the genus level revealed a total of 180 assigned genera. Overall analysis among twenty participants revealed that *Neisseria* (14.30%) was among the dominant genera followed by *Streptococcus* (12.20%) and *Prevotella* (9.493%). It was further extended by *Hemophilus* (4.403%), *Porphyromonas* (4.251%), *Rothia* (4.091%), *Granulicatella* (2.978%), *Actinobacillus* (2.831%), *Veillonella* (1.963%), and *Fusobacterium* (1.863%). The remaining genera exhibited their portion below 1% (Fig. 1b).

Diversity Indices Subject HTC-8 and HTC-12 showed the maximum richness or diversity having fourteen phyla followed by HTC-1, HTC-2, HTC-3, HTC-4, HTC-6, HTC-7, HTC-10,

Table 2 Diversity indices and estimators obtained among SLT users

Sample ID	OTU observed		Dominant phyla (%)	Dominant genera (%)	Diversity indices and estimators*			
	Phyla level	Genus level			Chao1	Shannon	Simpson	PD_whole tree
HTC-1	14	159	Proteobacteria (21.116)	*Streptococcus* (15.036)	9557.30	5.2053923	0.912	187.48
HTC-2	15	171	Firmicutes (16.514)	*Prevotella* (9.547)	8657.72	5.0484204	0.876	175.96
HTC-3	14	138	Firmicutes(34.4008)	*Streptococcus* (17.020)	4470.07	5.9524961	0.958	89.44
HTC-4	15	170	Proteobacteria (25.818)	*Neisseria* (12.528)	9574.38	5.9125061	0.945	192.68
HTC-5	13	141	Proteobacteria (29.830)	*[Prevotella]* (16.496)	7677.02	5.8913464	0.954	162.32
HTC-6	15	177	Firmicutes(31.493)	*Streptococcus* (16.617)	10425.03	5.8118277	0.944	214.26
HTC-7	15	173	Proteobacteria (43.487)	*Neisseria*(27.020)	10442.60	6.0041897	0.953	209.44
HTC-8	16	160	Firmicutes (26.487)	*Rothia* (20.244)	4136.56	5.277001	0.921	93.73
HTC-9	14	142	Proteobacteria (27.587)	*Haemophilus* (13.542)	8124.82	5.5267243	0.946	164.76
HTC-10	14	172	Proteobacteria (36.266)	*Klebsiella* (11.97817)	3600.83	6.0059864	0.942	108.05
HTC-11	14	159	Firmicutes (32.564)	*Streptococcus* (25.272)	7798.74	5.5154175	0.942	154.21
HTC-12	17	167	Firmicutes (34.407)	*Porphyromonas* (15.075)	5281.27	6.2574256	0.965	114.07
HTC-13	15	136	Proteobacteria (28.628)	*Neisseria* (24.667)	7084.04	5.549518	0.931	146.12
HTC-14	13	152	Proteobacteria (37.223)	*Neisseria* (20.315)	4596.40	5.6361764	0.940	95.51
HTC-15	15	182	Bacteroidetes (30.209)	*Prevotella* (27.490)	9278.87	5.6010132	0.945	206.62
HTC-16	15	141	Actinobacteria (33.041)	*Rothia* (30.934)	3382.18	4.4979157	0.898	79.72
HTC-17	15	183	Proteobacteria (38.116)	*Neisseria* (29.996)	9808.48	6.1823731	0.960	210.04
HTC-18	15	176	Firmicutes (40.916)	*Streptococcus* (12.403)	8901.64	6.9343595	0.976	197.39
HTC-19	15	139	Firmicutes (36.978)	*[Prevotella]* (14.224)	4427.81	5.6286031	0.942	92.89
HTC-20	15	127	Proteobacteria (48.950)	*Neisseria* (34.613)	5167.03	5.4568625	0.931	121.11

*Diversity indices and estimators of SLT non-users (HI) are mentioned in Table 1 (Verma et al., 2020)

HTC-13exhibited thirteen phyla, while subjec, HTC-15, HTC-16, HTC-17, and HTC-20 having thirteen phyla in each subject. A total of twelve phyla were observed in the subjects HTC-5, HTC-9, HTC-11, HTC-14, HTC-18, and HTC-19 (Supplementary file 1). The calculated diversity indices were also observed highest in HTC-18 and HTC-12 having the Shannon index of 6.93 and 6.25 respectively. The values of Simpson indices (SI) for the subject HTC-18 also showed highest in all subjects having a value of 0.9761331 followed by HTC-12 (0.965). It was followed by HTC-17 (SI: 0.960) and HTC-3 (SI: 0.958). While the subject HTC-2 showed the least richness as compared to the indices of other subjects, all the subjects showed Good estimator's values > 99. Comparative analysis between SLT and non-SLT users were further analysed for various diversity indices (Fig. 2).

Inter-individual Differences at the Level of Phylum and Genera Comparative analysis of the top ten phyla among twenty SLT users revealed that Proteobacteria were dominant in half of the participants (HTC-1, HTC-4, HTC-5, HTC-7, HTC-9, HTC-10, HTC-13, HTC-14, HTC-17, and HTC-20), where HTC-20 exhibited the maximum OTUs (48.950 %) of Proteobacteria, while HTC-8 showed a minimum of 7.49% OTUs. Similarly,

Firmicutes were observed as dominant phyla among eight participants (HTC-2, HTC-3, HTC-6, HTC-8, HTC-11, HTC-12, HTC-18, and HTC-19). Bacteroidetes were dominant only in one SLT user, i.e., HTC-15. However, it was of second most abundant phylum in more than half of the participants (HTC-2, HTC-5, HTC-6, HTC-7, HTC-9, HTC-10, HTC-11, HTC-12, HTC-17, HTC-18, HTC-19, and HTC-20) (Supplementary file 1).

Inter-individual differences at the genus level resulted in high variation. Six subjects (HTC-4, HTC-7, HTC-13, HTC-14, HTC-17, HTC-20) were populated by *Neisseria*, while *Streptococcus* showed its dominance in five participants (HTC-1, HTC-3, HTC-6, HTC-11, HTC-18). Two oral cavities were dominated by the genera *Prevotella* (HTC-2 and HTC-15) and *Rothia* (HTC-8 and HTC-16); however, *Prevotella* was found to be the second contributor in seven subjects (HTC-4, HTC-10, HTC-13, HTC-17, HTC-18, HTC-19, and HTC-20) (Supplementary file 2).

Shared Genera Of the one hundred eighty assigned genera, thirty-eight genera showed their occurrence among all the twenty SLT users (Supplementary file 3). The top twenty genera include *Neisseria*, *Streptococcus*, *Prevotella*, *Haemophilus*, *Porphyromonas*, *Rothia*, *Granulicatella*, *Actinobacillus*,
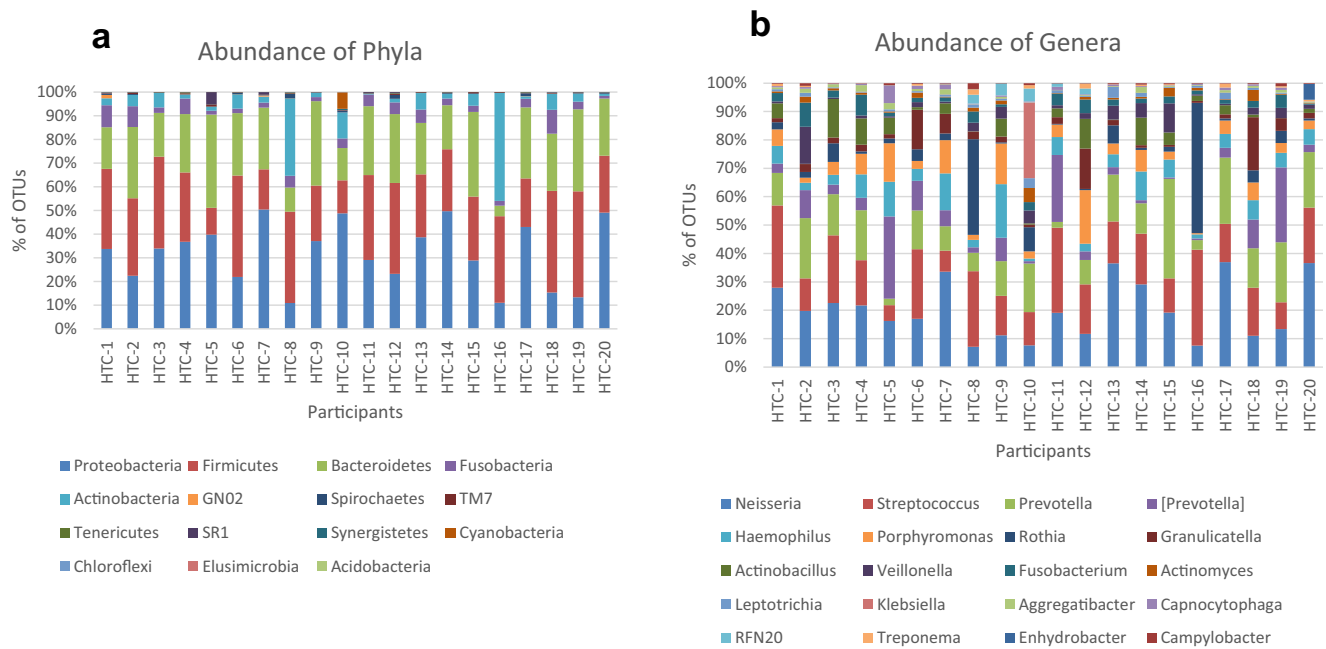
Fig. 1 The relative abundance of phyla (**a**) and top twenty genera, (**b**) present in the oral cavity of each SLT user.

Veillonella, *Fusobacterium*, *Actinomyces*, *Leptotrichia*, *Aggregatibacter*, *Capnocytophaga*, *Treponema*, *Campylobacter*, *Corynebacterium*, *Parvimonas*, *Selenomonas*, and *Mogibacterium* (Fig. 1b). *Peptostreptococcus*, *Bacteroides*, *Candidate Division_TG5*, *Cardiobacterium*, *Catonella*, *Microbacterium*, *Streptomyces*, *Paenibacillus*, and *Bifidobacterium* were the other highly shared genera among 85 to 95% of the SLT users. It was followed by *Lactobacillus*, *Peptococcus*, and *Candidatus division* SHD-231 that were shared among 80% of the total SLT oral cavities. Approximately seventy-five genera were common in less than 80% of the oral cavities. Approximately twenty-five genera were least shared and observed among two participants only. *Halomonas*, *Candidatus_Tammella*, *Roseomonas*, *Marinilactibacillus*, *Roseburia*, *Arcobacter*, *Sneathia*, and *Xanthobacter* are the name of few such genera.

Saccharopolyspora, *Acetobacter*, *Bilophila*, *Limnobacter*, *Brevibacillus*, and *Tessaracoccus* are among the thirty genera that were not shared and identified at the individual level. On the Spearman correlation analysis, heatmaps of shared genera showed a comparatively weak correlation among SLT users (Fig. 3a) as compared to the control group (Fig. 3b).

Comparative Analysis of SLT Users and Non-users

Phyla Level Among SLT users, the total number of assigned phyla was reduced to 15 as compared to the 20 bacterial phyla in SLT non-users [12]. These depleted phyla were Chlorobi, Thermi, Gemmatimonadetes, Nitrospirae, and Candidate division_OD1 as compared to the phyla detected in the oral microbiome of SLT non-users.
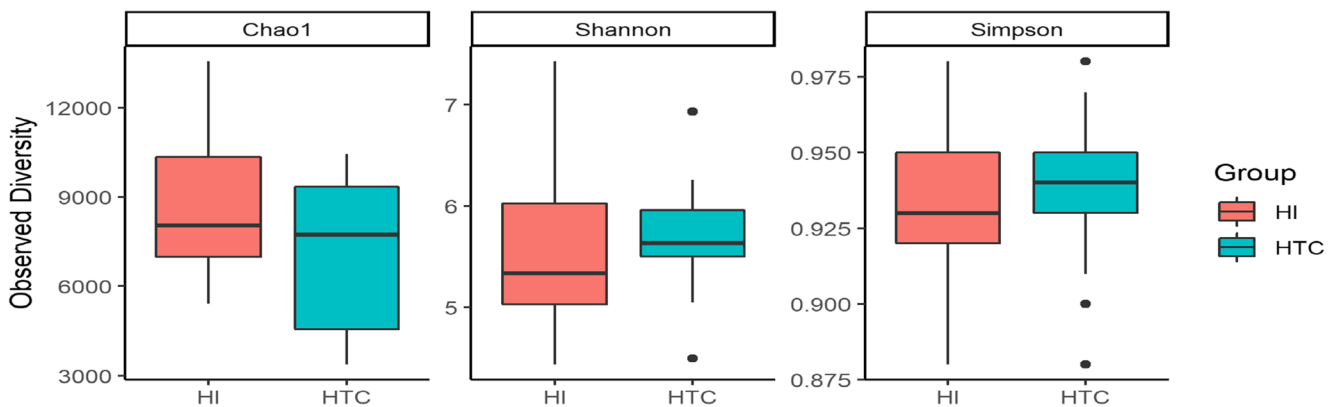
Fig. 2 Box plots illustrating alpha diversity indices (Chao1, Simpson index, and Shannon index) in bacterial microbiomes of SLT users (HTC) and non-users (HI) samples. Median values and inter-quartile ranges have been indicated in the plots

a

Heat Map of the Spearman Correlation Matrix

Variables

Spearman Correlation
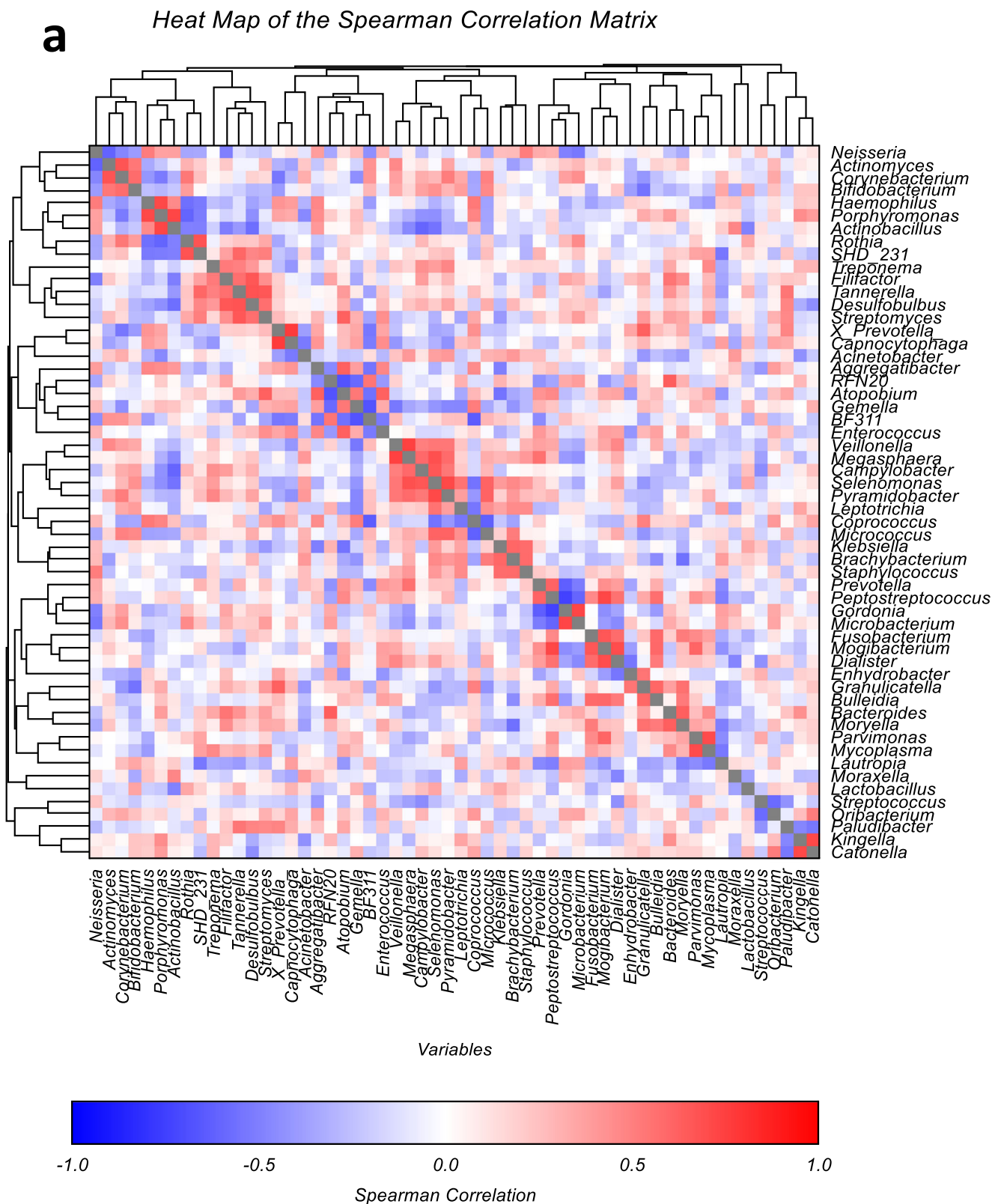
Fig. 3 Spearman correlation heatmaps for (a) HTC and (b) HI were generated using un-weighted pair group (average group) clustering method, including only those genera that were shared in both the groups and/or identified as significantly different in between the groups using NCSS software version 20.0.2. HI groups exhibit comparatively strong positive correlation with each other as compared to the HTC group
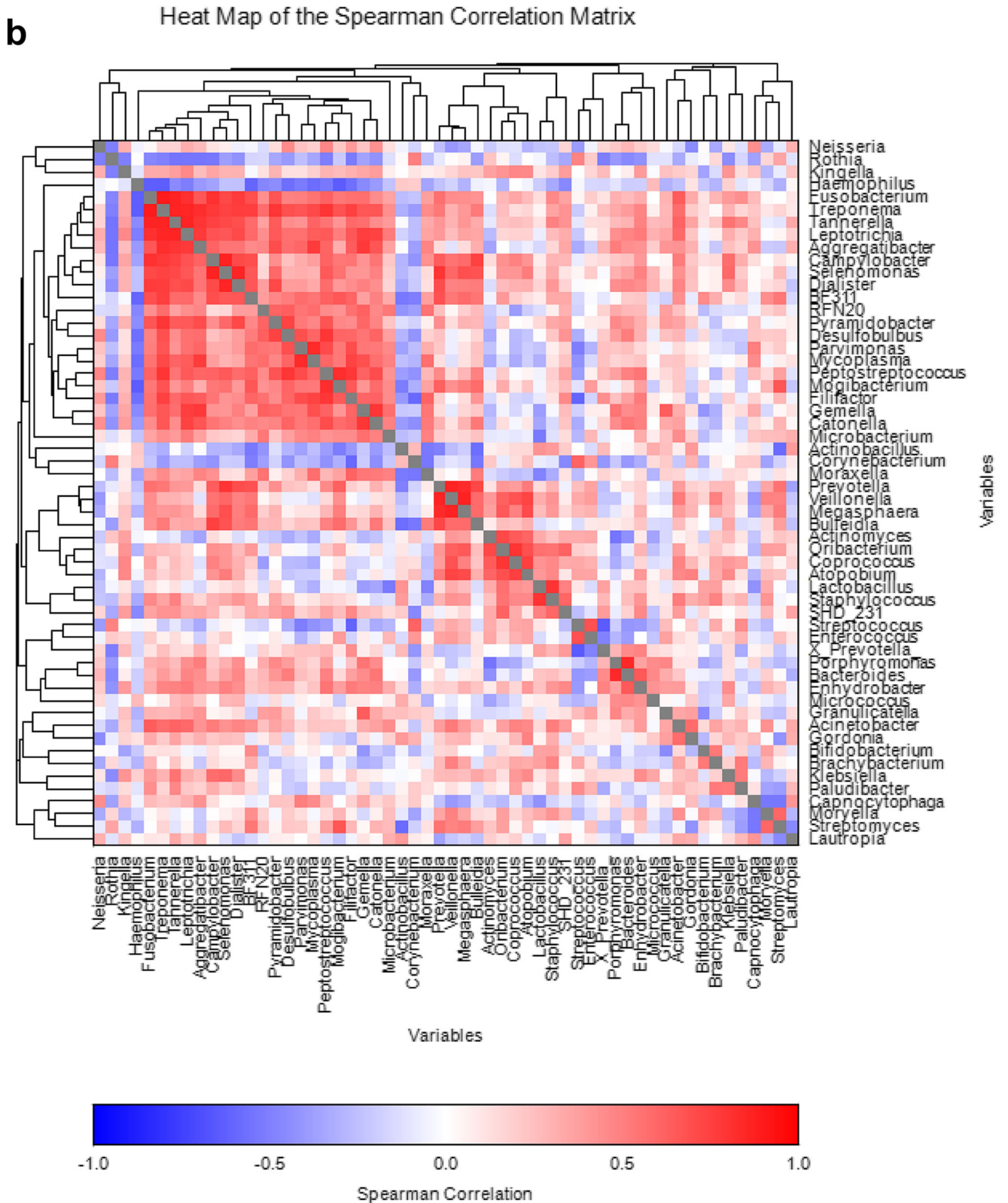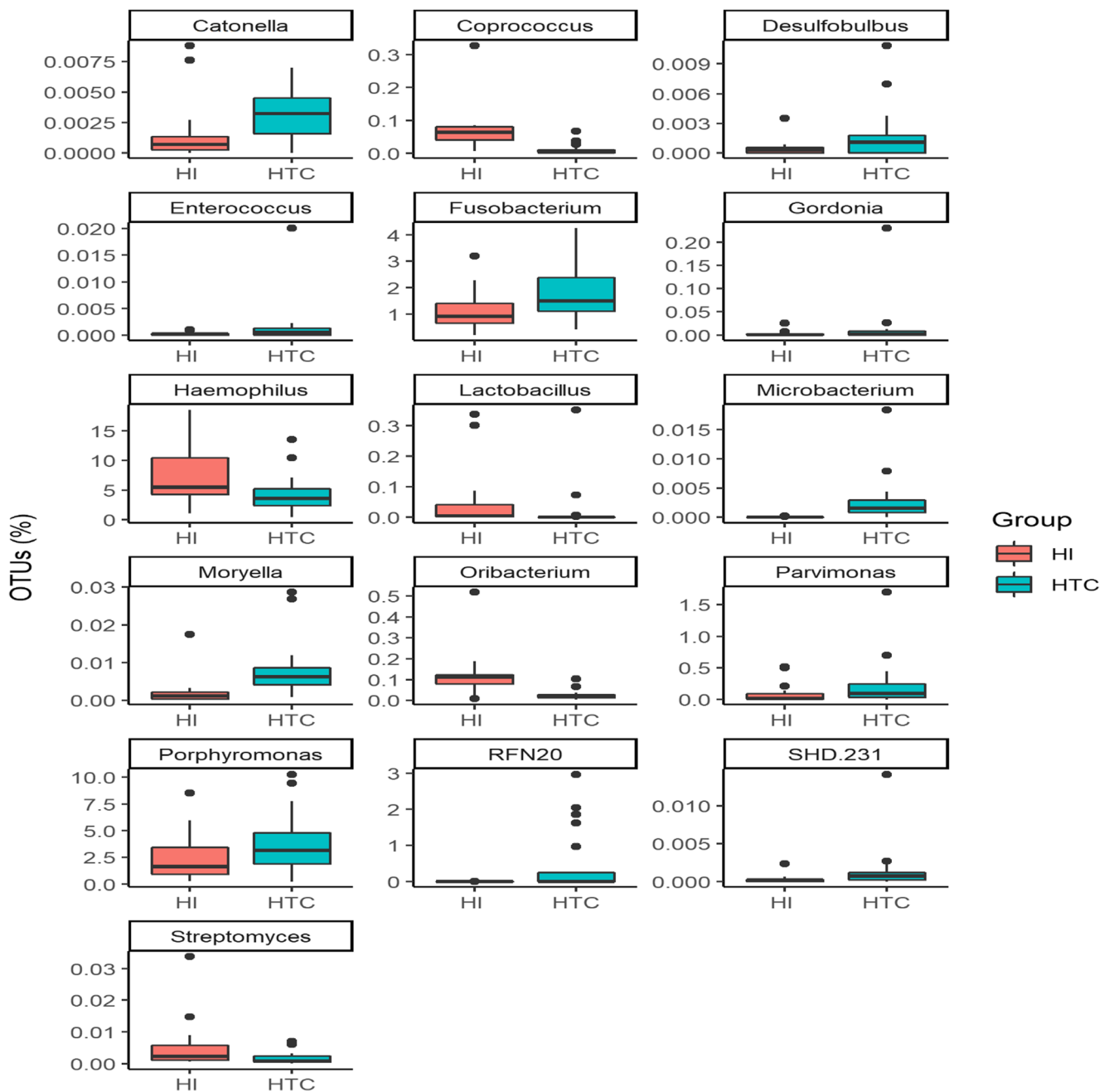
Fig. 3 continued

Fig. 4 The OTU (%) distribution of statistically significant genera between SLT users (HTC) and non-users (HI). Data was normalized by taking center log ratio (clr)

Genera Level Here, the total assigned genera were increased among SLT users as compared to the control group. Among SLT users, the total assigned genera were 180 than that of 162 among the SLT non-users (Supplementary table 5). Moreover, 122 genera were common in between the groups, where 58 genera were unique in the SLT group. Similarly, forty genera of SLT non-users were missing among SLT users.

Statistical Analysis Of the fifteen common phyla, Fusobacteria and Tenericutes were detected as significantly different (p-

value < 0.05) among both groups. The p-values and FDR-adjusted p-values are shown in Supplementary table 5. Genera level analysis revealed a total of sixteen genera, viz., *Fusobacterium*, *Porphyromonas*, *Parvimonas*, RFN20, *Oribacterium*, *Lactobacillus*, *Gordonia*, *Coprococcus*, *Moryella*, *Catonella*, *Microbacterium*, *Streptomyces*, *Desulfobulbus*, *Enterococcus*, SHD.231, and, *Heamophilus* that were significantly different in between the two groups (Fig. 4). SLT users showed more prevalence of *Fusobcaterium*, *Porphyromonas*, RFN20, *Parvimonas*,
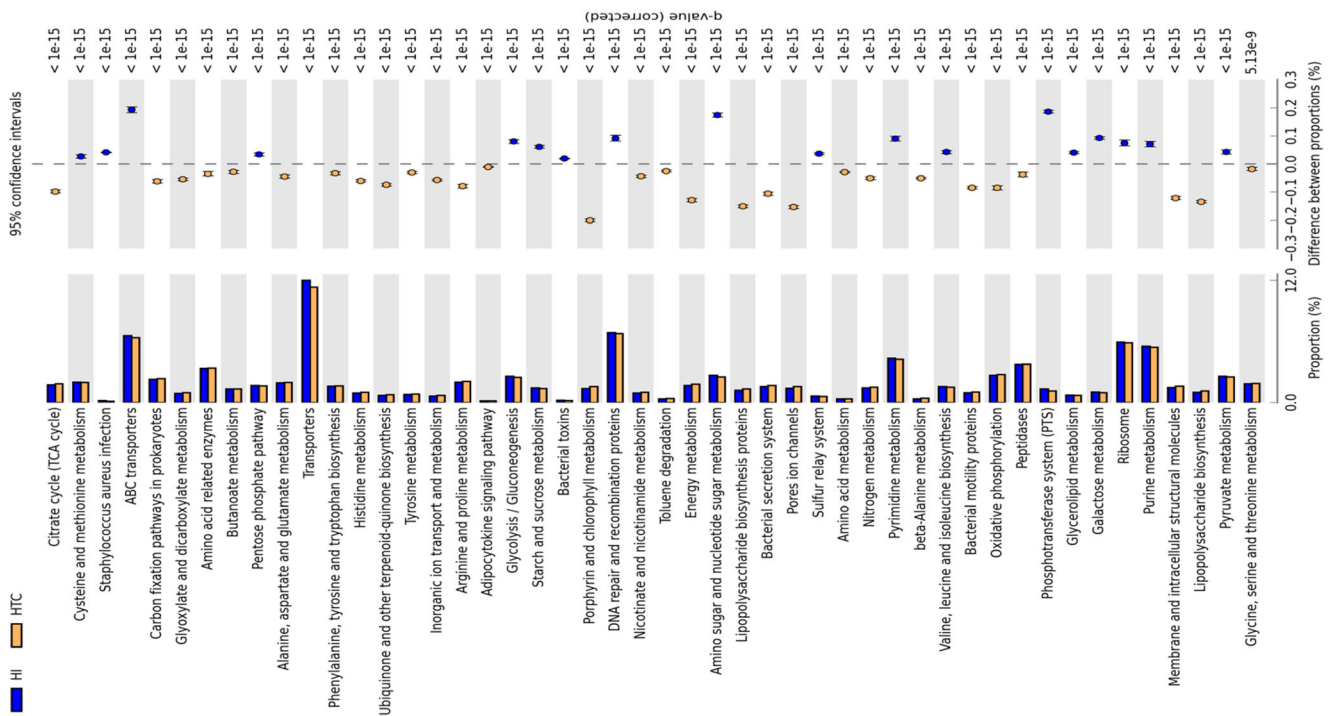
Fig. 5 The PICRUSt analysis at tier III reveals significant differences among the genes involved in toluene/nitrotoluene degradation as well as nicotinate/nicotinamide degradation

Gordonia, *Moryella*, *Catonella*, *Microbacterium*, D*esulfobulbus*, *Enterococcus*, and SHD.231, while other remaining genera were relatively higher in SLT non-users (Supplementary table 5). The *p*-values and FDR-adjusted *p*-alues are shown in Supplementary table 5.

Effect of Tobacco Consumption at the Level of Gender and Diet Only one genus *Atopobium* (*p*-value <0.001) was identified as significantly different at the level of gender, while *Mycoplasma* (*p*-value 0.027) was significantly different at the level of diet.

PICRUSt Analysis The relative abundance of predicted genes revealed that the majority of the OTUs belong to the genes involved in metabolic pathways, i.e., 47.14% followed by the genes of genetic information/processing (21.63%), environmental information processing (12.74%), and human diseases (0.98%). Tier I showed a significant increase in the count of genes involved in cellular processes, environmental information processing, human diseases, and metabolism, among HTC users as compared to the control group (Benjamini–Hochberg FDR-adjusted value <0.05, *q* value −1e−15) (Supplementary fig 1a). Tier II represents the extension of tier I and its analysis revealed a significant abundance for the genes that belong to amino acid metabolism, cellular processes, and signaling, metabolism of cofactors and vitamins, xenobiotic biodegradation, and metabolism in HTC users, while SLT non-users showed an abundance of genes for carbohydrate metabolism, replication, and repair, lipid metabolism, transcription, translation, and nucleotide metabolism in tier II analysis (*q* value: −1e−15) (Supplementary fig. 1b). Tier III further expands the tier I– and tier II–level genes. Overall, tier III analysis showed an enhancement in the number of genes involved in nitrogen metabolism, amino acid metabolism, and nicotinate/nicotinamide metabolism, as well as toluene degradation in SLT users (Fig. 5). Besides, SLT users also showed more abundance for the genes involved in bacterial secretion systems, bacterial motility proteins, and other transporters, while the genes of starch and sugar metabolism, glycolysis/gluconeogenesis, DNA repair and recombination proteins, ABC transporter, and phosphotransferase system (PTS) were more abundant in SLT non-users (Benjamini–Hochberg FDR-adjusted *q* value= 1e−15).

Discussion

India bears a high burden of smokeless tobacco consumers due to the occurrence of a wide variety of SLT products that are easy to access to every age group [1]. Tobacco consumption is one of the major reasons for preventable deaths globally, where India accounts for approximately one million deaths every year due to tobacco consumption in any form [3].

Our study encompasses the first extensive characterization of the oral microbiome of SLT users from the Indian population. Here, we also report the effect of smokeless tobacco on a

healthy oral microbiome and observed that alterations occur at various taxon levels and so forth their predicted functions also deviate accordingly. Various forms of tobacco products have been reported to influence the microbiome of healthy oral cavities [10, 22] but such information is lacking from the Indian community. We observed that the count of total genera was 180 in SLT users as compared to 162 in the control group. Of these 180 genera, 122 genera were shared in both the groups, where 58 were unique to SLT users only. It may be due to the introduction of inhabitant tobacco bacteria that compete with the autochthonous bacteria of the oral cavity for nutrients and other limiting substrates and consequently enhance the overall bacterial composition [23, 24]. The depletion of 40 genera in SLT users as compared to the control group is also noteworthy and envisages that several toxicants, heavy metals, alkaloids, and other xenobiotic compounds of SLT may be the possible reasons after it [5]. In a traditional cultivation approach, SLT has been reported to alter (promote/ inhibit) the growth of eleven isolated oral bacteria, where SLT promoted the growth of seven strains and inhibited 4 strains that exhibit a role in biofilm formation [22]. The Middle Eastern smokeless tobacco product Dokha has also shown a significant association with oral microbiome dysbiosis among the UAE population [25].

The statistical analysis uncovered that the beta diversity between the groups was consistent (p-value 0.7) (Supplementary fig. 2). However, differently significant candidates were identified at the level of the various taxa. Two phyla (Fusobacteria, and Tenericutes) were identified as significantly different between the two groups (p-value <0.05) (Supplementary table 5). Elevated count of Fusobacteria among SLT users indicates poor oral health. Fusobacteria belongs to the highly common anaerobic Gram-negative bacteria of the oral cavity that act as a connecting link between early and late colonizers of dental plaque [26]. Due to several pathogenic characteristics like production of adhesins, co-aggregation with other bacteria, susceptibility towards several potent antibiotics, incorporation of outer membrane protein (OMPs), and their abundance during poor dental hygiene makes it a potential candidate for the onset of several periodontal diseases [27]. It further suggests that tobacco consumers are more prone to Fuosobacterial diseases. A marginal decrease in the count of Tenericutes (p-value: 0.049) in SLT users is the subject of further research, especially in oral health. Mycoplasma, a highly studied genus of the phylum Tenericute, has been associated with the pathogenesis of severe respiratory and urogenital tract infections [28]. The count of Tenericutes was comparatively higher in the patients with periodontitis as compared to the healthy controls [29].

Statistical analysis revealed only sixteen significantly distinct genera in between these two groups (p-values <0.05). Besides, *Fusobacterium* as discussed above, other genera such as *Porphyromonas*, *Parvimonas*, *Gordonia*, *Moryella*, *Catonella*, *Microbacterium*, *Desulfobulbus*, *Enterococcus*,

and SHD.231 were also significantly elevated in SLT users. *Fusobacterium*, *Porphyromonas*, *Haemophilus*, *Enterococcus*, and *Lactobacillus* have been identified as highly altered genera during tobacco-associated oral dysbiosis also [30]. The association of *Fusobacterium* and *Porphyromonas* in dental biofilm formation is well established and paves the way for several oral diseases [26, 27]. *Parvimonas* (formerly known as the member of *Micromonas*) belong to Gram-positive anaerobic cocci (GPAC) and shares properties with *Peptostreptococcus* [31] which further affirm our findings that these antecedent bacteria may act as an early colonizer in biofilm formation and allow the adhesion of *Fusobacterium*- and *Porphyromonas*-like Gram-negative bacteria [26, 27]. *Parvimonas micra* is the only reported genus of this taxon to date [32, 33] and also has been reported as a prominent oral pathogen [31]. Similarly, the genus *Catonella* and *Moryella* also exhibit only one species *Catonella morbi* [34] and *Moryella indoligenes* [35] respectively which cause endodontic infections and periodontitis. *Moryella* is a weak saccharolytic bacteria [35] that produce acid during carbohydrate fermentation and may create an acidic environment favorable for adherence of other strong saccharolytes into the oral cavity and elevates their count during oral illness [36, 37]. In our case, the count of *Moryella* was significantly higher in SLT users (p-value 0.02) which suggests further exploring its role in oral health. Elevated count of *Enterococcus* (a lactic acid bacteria) may also suggest a similar function in SLT user's oral cavity. Besides, members of *Enterococcus* have been recognized as the transient constituents of the oral microbiome advocate its appearance in the oral cavity due to external sources [38]. Tsuzukibashi et al (2015) suggest that *Microbacterim* does not constitute the human normal oral microflora and exhibit in denture wearing oral cavities [39]. *Gordonia* and *Microbacterium* count were another significantly abundant methylotrophic actinobacteria in SLT users that further suggest them as transient genera in the healthy human oral cavity [39, 40]. The members of these genera have been recognized as environmental strains having a role in nitrogen metabolism [41, 42]. Therefore, these genera may play a key role in the transformation of nicotine to tobacco-specific nitrosamines [43, 44]. Such findings encourage us to identify the inhabitant bacteria of smokeless tobacco products for enhancing our information on transient bacteria. Increased *Desulfobulbus* count in SLT users corroborates our findings with the previous reports, where enhanced *Desulfobulbus* count has been identified in the patients of periodontal diseases [45, 46]. Anaerobic and chemoorganothrophic *Desulfobulbus* potentially transform sulfur into sulfate/ sulfide and plays a crucial role in dental caries [47]. The prevalence of methylotrophic bacteria and *Desulfobulbus* in SLT user's oral cavity can also be linked for halitosis.

We obtained a lower count of *Lactobacillus* and *Haemophilus* in SLT users. It may be due to the inhibitory effect of SLT

products on several other oral isolates [7, 48]; therefore, it needs to be examined carefully using standard strains by traditional cultivation approaches. In an interesting study, *Lactobacillus* count was reduced in presence of a higher concentration of nitrate in the medium [49]; therefore, we may hypothesize that the oral cavity of tobacco consumers exhibits an abundance of nitrate/nitrite that can be estimated in terms of TSNAs (2.3–27 μg/g SLT powder in India; 50). *Lactobacillus* and *Haemophilus* have been reported as health compatible commensals and thus their reduction signals for poor oral health [24]. Other significantly depleted genera in SLT users were *Oribacterium*, *Caprococcus*, and *Streptomyces*. The elevated count of *Oribacterium* has been reported in oral cancers [51, 52]. *Streptomyces* is chiefly a soil dweller while *Coprococcus* is more prominent in fecal samples; therefore, their occurrence in the oral cavity may be considered as conjectural [53, 54].

Analysis at the level of gender and food habits further validates our findings that food habits (vegetarians and non-vegetarians) and gender do not affect the overall bacterial composition if the population belongs to the same demographic regions [13]. Only two genera *Atopobium* and *Mycoplasma* were statistically different at the gender and diet level respectively. *Atopobium*, a Gram-positive H_2S-producing rod is used as a biomarker for the diagnosis of OSCC (oral squamous cell carcinoma) and plays a crucial role in bacterial vaginosis and colorectal cancer [55]. However, its significant abundance at the gender and diet level needs to be further examined.

The statistical analysis of predicted genes in between the groups revealed significant findings at tiers I, II, and III using STAMP version 1.2.3. These levels depict the hierarchical functional pathways of the Kyoto Encyclopedia of Gene and Genome (KEGG) that were obtained by using the output of the PICRUSt analysis. Differences between mean proportion showed a significant increase in the genes involved in metabolism, cellular processes, environmental information processing, and human diseases among HTC users. It indicates that tobacco and its inhabitant bacteria showing its influence in deciding the overall oral bacterial composition and therefore associated pathways are also being affected [56]. Tier II analysis uncovered extended analysis, where mean abundance proportion was relatively higher for the genes involved in the amino acid metabolism, xenobiotic biodegradation, and metabolism activity in HTC users. Oral bacteria encounters first with tobacco flakes that contain several recalcitrant compounds and toxicants [12, 50], and therefore bacteria may shift their metabolic activities for degrading such xenobiotic compounds [57]. A higher proportion of genes of toluene/nitrotoluene degradation, nicotinate/nicotinamide degradation, and nitrogen metabolism among SLT users at the tier III level further affirms the influence of tobacco on the overall bacterial metabolic activity as compared to the healthy participants. We hypothesize here that an increase in genes of nitrogen metabolisms indicates the conversion of several nitrogen derivatives and their transformation into tobacco-specific nitrosamines on reaction with nicotine into the oral cavity of SLT users.

Conclusion

The present investigation is the first information on the oral microbiome of SLT users from the Indian population that enhances our understanding of the oral bacterial composition of SLT users. Overall, we conclude that smokeless tobacco significantly alters (enrich/deplete) the oral microbiome. Enrichment of the SLT's oral cavity with several pathogenic and anaerobic bacteria suggests the onset of various oral diseases. The information of significantly altered genera could be explored further for making biomarkers for an early diagnosis of tobacco-induced oral disease. However, further research is required in this direction with a large sample size by considering varied socio-economic factors to reveal a better picture. An attempt should also be made to explore the inhabitant microorganisms of various commercial tobacco products to establish a correlation with the oral bacteria. Shotgun metagenomics could be a better tool to uncover the overall dynamics of the oral bacteria in terms of their functional characteristics, where we can directly retrieve the information of the genes involved in nitrogen metabolism (especially the nitrogen reductase gene). This study has significance as India is among the largest consumers of SLT users, where we have no concrete information on the oral bacterial composition of SLT users/SLT products.

Declarations

Ethics approval The protocol of the study was approved by the Institutional Ethical Committee, Netaji Subhas University of Technology, New Delhi (NSUT), as well as the All India Institute of Medical Sciences (AIIMS), New Delhi (Ref No. IEC/NP 166/2014/ RP-14/2014).

Consent to participate A signed consent was taken from all the participants.

Consent for publication On acceptance of the manuscript, the copyrights will be transferred to the publisher.

Conflict of interest The authors declare no competing interests.

References

1. Sinha DN, Gupta PC, Kumar A, Bhartiya D, Agarwal N, Sharma S, Singh H, Parascandola M, Mehrotra R (2018) The poorest of poor suffer the greatest burden from smokeless tobacco use: a study from 140 countries. Nicotine Tob. Res. 20:1529–1532

2. Siddiqi K, Islam Z, Khan Z, Siddiqui F, Mishu M et al (2019) Identification of policy priorities to address the burden of smokeless tobacco in Pakistan: a multi method analysis. Nicotine Tob. Res. 29: ntz163

3. WHO Factsheet India 2018. Received from regional office of South East Asia.

4. Rodgman A, Perfetti TA (2009) The chemical components identified in tobacco and tobacco smoke prior to 1954: a chronology of classical chemistry. Beitr. Tabakforsch. Int. 23:277–333

5. Vishwakarma A, Verma D (2020) Exploring the microbiome of smokeless tobacco. In: Chaudhary, Raj, Verma, Akhter (eds) Microorganisms for Sustainable Environment and Health. Elsevier, Amsterdam, pp 167–178

6. Han J, Sanad YM, Deck J, Sutherland JB, Li Z, Walters MJ (2016) Bacterial populations associated with smokeless tobacco products. Appl. Environ. Microbiol. 82:6273–6283

7. Law AD, Fisher C, Jack A, Moe LA (2016) Tobacco, microbes, and carcinogens: correlation between tobacco cure conditions, tobacco-specific nitrosamine content, and cured leaf microbial community. Microb. Ecol. 72:120–129

8. Halboub E, Al-Akhali MS, Alamir AH, Homeida HE, Baraniya D et al (2020) Tongue microbiome of smokeless tobacco users. BMC Microbiol. 20:201

9. National cancer institute and centers for disease control and prevention (2014) Smokeless tobacco and public health: a global perspective. U.S. Department of health and human services, Centers for disease control and prevention and National Institutes of Health, National cancer institute. NIH Publication No, Bethesda, MD, pp 14–7983

10. Miluna S, Rostoka D, Skadioo I, Reinis A, Priedite V, Koka R, Lauva D, Krojea J (2017) The oral microbiome of smokeless tobacco users in Latvia. Proc. Latv. Acad. Sci. 71:33–37

11. Ganesan SM, Joshi V, Fellows M, Dabdoub SM, Nagaraja HN (2017) A tale of two risks: smoking diabetes and the subgingival microbiome. ISME J 11:2075–2089

12. Bornigen D, Ren B, Pickard R, Li J, Ozer E, Hartmann EM, Xiao W, Tickle T, Rider J, Gevers D, Franzosa EA (2017) Alterations in oral bacterial communities are associated with risk factors for oral and oropharyngeal cancer. Sci. Rep. 7:17686

13. Verma D, Srivastava A, Garg PK, Akhter Y, Dubey AK, Mishra SD, Deo SVS (2020) Taxonomic profiling and functional characterization of the healthy human oral bacterial microbiome from the north Indian urban sub-population. Arch. Microbiol. https://doi.org/10.1007/s00203-020-02084-7

14. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Sys Biol 7:539

15. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200

16. Edgar R (2013) UPARSE: Highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10:996–998

17. Bik EM, Eckburg PB, Gill SR, Nelson KE, Purdom EA (2006) Molecular analysis of the bacterial microbiota in the human stomach. Proc. Natl. Acad. Sci. U. S. A. 103:732–737

18. Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: Statistical analysis of taxonomic and functional profiles. Bioinformatics 30:3123–3124

19. Langille MG, Zaneveld J, Caporaso JG (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat. Biotechnol. 31:814–821

20. Prodan A, Tremaroli V, Brolin H, Zwinderman AH, Nieuwdorp M, Levin E (2020) Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. PLoS One 15:e0227434

21. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome 6:90

22. Liu M, Jin J, Pan H, Feng J, Cerniglia CE, Yang M, Chen H (2016) Effect of smokeless tobacco products on human oral bacteria growth and viability. Anaerobe 42:152–161

23. Camelo-Castillo AJ, Mira A, Pico A et al (2015) Subgingival microbiota in health compared to periodontitis and the influence of smoking. Front. Microbiol. 6:119

24. Mason MR, Preshaw PM, Nagaraja HN, Dabdoub SM, Rahman A, Kumar PS (2015) The subgingival microbiome of clinically healthy current and never smokers. ISME J 9:268–272

25. Valles Y, Inman CK, Peters BA (2018) Types of tobacco consumption and the oral microbiome in the United Arab Emirates Healthy Future (UAEHFS) Pilot Study. Sci. Rep. 8:11327

26. Verma D, Garg PK, Dubey AK (2018) Insights into the human oral microbiome. Arch. Microbiol. 200:525–540

27. Jenkinson HF, Lamont RJ (2005) Oral microbial communities in sickness and in health. Trends Microbiol. 13:589–595

28. Gupta S, Gupta R, Sinha DN, Mehrotra R (2018) Relationship between type of smokeless tobacco and risk of cancer: a systematic review. Indian J. Med. Res. 148:56–76

29. Lundmark A, Hu Y, Huss M, Johannsen G, Andersson AF, Yucel-Lindberg T (2019) Identification of salivary microbiota and its association with host inflammatory mediators in periodontitis. Front. Cell, Microbiol 9:216

30. Huang C, Shi G (2019) Smoking and microbiome in oral, airway, gut and some systemic diseases. J. Transl. Med. 17:225

31. Ho D, Ang G, Er C (2018) An unusual presentation of *Parvimonas micra* infective endocarditis. Cureus 10:e3447

32. Tindall B, Euzeby J (2006) Proposal of *Parvimonas* gen. nov. and *Quatrionicoccus* gen. nov. as replacements for the illegitimate, prokaryotic, generic names *Micromonas* Murdoch and Shah 2000 and *Quadricoccus* Maszenan et al 2002, respectively. Int. J. Syst. Evol. Microbiol. 56:2711–2713

33. Murphy EC, Frick IM (2013) Gram-positive anaerobic cocci – commensals and opportunistic pathogens. FEMS Microbiol. Rev. 37:520–553

34. Siqueira Jr JF, Rocas IN (2006) *Catonella morbi* and *Granulicatella adiacens*: new species in endodontic infections. Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod. 102:259–264

35. Carlier JP, Kouas G, Han XY (2007) *Moryella indoligenes* gen. nov., sp. nov., an anaerobic bacterium isolated from clinical specimens. Int. J. Syst. Evol. Microbiol. 57:725–729

36. Costalonga M, Herzberg MC (2014) The oral microbiome and the immunobiology of periodontal disease and caries. Immunol. Lett. 162:22–38

37. Chen X, Winckler B, Lu M, Cheng H, Yuan Z, Yang Y (2015) Oral microbiota and risk for esophageal squamous cell carcinoma in a high- risk area of China. PLoS One 10:e0143603

38. Komiyama EY, Lepesqueur LSS, Yassuda CG, Samaranayake LP, Parahitiyawa NB, Balducci I (2016) *Enterococcus* species in the

oral cavity: prevalence, virulence factors and antimicrobial susceptibility. PLoS One 11:e0163001

39. Hung WL, Wade WG, Boden R, Kelly DP, Wood AP (2011) Facultative methylotrophs from the human oral cavity and methylotrophy in strains of *Gordonia*, *Leifsonia*, and *Microbacterium*. Arch. Microbiol. 193:407–417

40. Tsuzukibashi O, Uchibori S, Kobayashi T, Saito M, Umezawa K, Ohta M, Shinozaki-Kuwahara N (2015) A selective medium for the isolation of *Microbacterium* species in oral cavities. J. Microbiol. Methods 116:60–65

41. Romanowska I, Kwapisz E, Mitka M, Bielecki S (2010) Isolation and preliminary characterization of a respiratory nitrate reductase from hydrocarbon-degrading bacterium *Gordonia alkanivorans* S7. J. Ind. Microbiol. Biotechnol. 37:625–629

42. Zhang D, Li W, Huang X, Wen Q, Liu M (2013) Removal of ammonium in surface water at low temperature by a newly isolated *Microbacterium* sp. strain SFA13. Bioresour. Technol. 137:147–152

43. Fisher MT, Bennett CB, Hayes A, Kargalioglu Y, Knox BL, Xu D, Muhammad-Kah R, Gaworski CL (2012) Sources of and technical approaches for the abatement of tobacco specific nitrosamine for mationin moist smokeless tobacco products. Food Chem. Toxicol. 50:942–948

44. Wei X, Deng X, Cai D, Ji Z, Wang C, Yu J, Li J, Chen S (2014) Decreased tobacco-specific nitrosamines by microbial treatment with *Bacillus amyloliquefaciens* DA9 during the air-curing process of burley tobacco. J. Agric. Food Chem. 62:12701–12706

45. Junemann S, Prior K, Szczepanowski R (2012) Bacterial community shift in treated periodontitis patients revealed by Ion Torrent 16S rRNA gene amplicon sequencing. PLoS One 7:e41606

46. Chen WP, Chang SH, Tang CY, Liou ML, Tsai SJ, Lin YL (2018) Composition analysis and feature selection of the oral microbiota associated with periodontal disease. Biomed. Res. Int. 2018:3130607

47. Greabu M, Totan A, Miricescu D, Radulescu R, Virlan J, Calenic B (2015) Hydrogen sulfide, oxidative stress and periodontal diseases: a concise review. Antioxid (Basel) 5:3

48. Al-hebshi N, Alharbi F, Mahri M, Chen T (2017) Differences in the bacteriome of smokeless tobacco products with different oral carcinogenicity: compositional and predicted functional analysis. Genes 8:106

49. Hyde ER, Andrade F, Vaksman Z, Parthasarathy K, Jiang H, Parthasarathy DK (2014) Metagenomic analysis of nitrate-reducing bacteria in the oral cavity: implications for nitric oxide homeostasis. PLoS One 9:e88645

50. Nasrin S, Chen G, Watson CJW, Lazarus P (2020) Comparison of tobacco-specific nitrosamine levels in smokeless tobacco products: high levels in products from Bangladesh. PLoS One 15:e0233111

51. Guerrero-Preston R, Godoy-Vitorino F, Jedlicka A, Rodríguez-Hilario A, Gonzalez H, Bondy J (2016) 16S rRNA amplicon sequencing identifies microbiota associated with oral cancer, human papilloma virus infection and surgical treatment. Oncotarget 7:51320–51334

52. Lim Y, Fukuma N, Totsika M, Kenny L, Morrison M, Punyadeera C (2018) The performance of an oral microbiome biomarker panel in predicting oral cavity and oropharyngeal cancers. Front. Cell. Infect. Microbiol. 8:267

53. Segata N, Haake SK, Mannon P (2012) Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol. 13R42:1–18

54. Bolourian A, Mojtahedi Z (2018) *Streptomyces*, shared microbiome member of soil and gut, as 'old friends' against colon cancer. FEMS Microbiol. Ecol. 1:94

55. Li Y, Tan X, Zhao X, Xu Z, Dai W, Duan W, Huang S, Zhang E, Liu J, Zhang S, Yin R, Shi X, Lu Z, Pan Y (2020) Composition and function of oral microbiota between gingival squamous cell carcinoma and periodontitis. Oral Oncol. 107:104710

56. Bhushan B, Yadav AP, Singh SB (2019) Diversity and functional analysis of salivary microflora of Indian Antarctic expeditionaries. J. Oral Microbiol. 11:1581513

57. Wu J, Peters BA, Dominianni C, Zhang Y, Pei Z (2016) Cigarette smoking and the oral microbiome in a large study of American adults. ISME J 10:2435–2446