JOURNAL OF MEDICAL MICROBIOLOGY

RESEARCH ARTICLE

Karabudak *et al.*, *Journal of Medical Microbiology*

DOI 10.1099/jmm.0.001003

MICROBIOLOGY SOCIETY

Analysis of the effect of smoking on the buccal microbiome using next-generation sequencing technology

Sema Karabudak[1,†], Oguz Ari[1,†], Bengul Durmaz[2], Tuba Dal[3], Tugcan Basyigit[3], Mahmut Tayyar Kalcioglu[4] and Riza Durmaz[1,3,*]

Abstract

Purpose. This study aimed to investigate the effect of smoking on the buccal microbiome and to analyse the descriptive ability of each of the seven hypervariable regions in their 16S rRNA genes.

Methodology. Microbiome compositions of 40 buccal swab samples collected from smokers (*n* =20) and non-smokers (*n* =20) were determined using 16S rRNA sequencing. Seven different 16S rRNA hypervariable regions (V2, V3, V4, V6-7, V8 and V9) in each sample were amplified using the Ion Torrent 16S Metagenomics kit and were sequenced on the Ion S5 instrument.

Results. Seven hypervariable regions in the 16S rRNA gene were successfully sequenced for all samples tested. The data obtained with the V2 region was found to be informative but the consensus data generated according to a number of operational taxonomic unit reads gathered from all seven hypervariable regions gave the most accurate result. At the phylum level, no statistically significant difference was found between smokers and non-smokers whereas relative abundances of *Veillonella atypica*, *Streptococcus australis*, *Prevotella melaninogenica*, *Prevotella salivae* and *Rothia mucilaginosa* showed significant increases in the smoker group (*P*-adj=0.05). Alpha diversity results did not show a significant difference between the two groups; however, beta diversity analysis indicated that samples of smoker and non-smoker groups had a tendency to be clustered within themselves.

Conclusion. The results of the current study indicate that smoking is a factor influencing buccal microbiome composition. In addition, sequencing of all seven hypervariable regions yielded more accurate results than those with any of the single variable regions.

INTRODUCTION

The human oral cavity, which is the initial digestive organ that breaks down carbohydrates and dietary lipids, hosts more than 700 bacterial species [1, 2]. Along with other micro-organisms residing in the oral cavity, they constitute the oral microbiome, which can be classified into two parts, the core microbiome and variable microbiome. The core microbiome is a key element for maintaining oral health and has similar composition in every individual as specific dominant species can be observed in each of the oral habitats [3]. In contrast, the variable microbiome can differ due to various factors such as diet type, hygiene and delivery mode [4–6]. There are 11 different microbial habitats in the oral cavity such as the saliva, tongue, hard palate, dorsum, palatine tonsils, throat, buccal mucosa, keratinized gingiva, supra-gingival plaque, subgingival plaque, dentures and lips, with each containing its own core microbiome composition [3]. For instance, the dominant micro-organisms in saliva are *Prevotella melaninogenica* and *Streptococcus salivarius* [7]. In the buccal mucosa, *Streptococci* are predominant; however, *Haemophilus parainfluenzae, Simonsiella* spp., *Granulicatella* spp., *Gemella* spp., *Veillonella*

Received 29 November 2018; Accepted 03 May 2019; Published 14 June 2019

Author affiliations: [1]Ankara Yildirim Beyazit University, Central Research and Application Center, Ankara, Turkey; [2]Yuksek Ihtisas University, Faculty of Medicine, Department of Clinical Microbiology, Ankara, Turkey; [3]Ankara Yildirim Beyazit University, Faculty of Medicine, Department of Clinical Microbiology, Ankara, Turkey; [4]Istanbul Medeniyet University, Faculty of Medicine, Department of Otolaryngology, Istanbul, Turkey.

***Correspondence:** Riza Durmaz, rdurmaz@ybu.edu.tr

Keywords: smoking; next-generation sequencing; metagenomic analysis; buccal microbiome.

Abbreviations: DNA, deoxyribonucleic acid; FDR, false discovery rate; H_2S, hydrogen sulfide; OTU, operational taxonomic unit; P-adj, P adjusted; PBS, phosphate buffered saline; PC1, principal coordinate 1; PCoA, principal coordinates analysis; PCR, Polymerasechain reactions; pM, picomolar; QIIME, quantitative insights into Microbial Ecology; rRNA, ribosomal ribonucleic acid; SD, Standard Deviation; SFA, saturated fatty acids; spp, several species.

†These authors contributed equally to this work

spp. and *Prevotella* spp. have been detected in samples from the buccal mucosa [8].

As the composition of the oral microbiome is crucial for oral health, disturbance of the oral microbiome in favour of pathogenic bacteria causes periodontal diseases such as gingivitis and periodontitis [9]. As such, smoking may have serious adverse effects on the oral microbiome. Because cigarettes contain numerous substances that can be toxic for some beneficial bacterial species living in the oral cavity, smoking may ultimately lead to disease [10]. Furthermore, oxygen deprivation caused by smoking might result in the proliferation of anaerobic species [11], and bacteria present on the cigarette could be transferred to the oral cavity through lips before ignition, which in turn could alter the oral microbial diversity [12]. It is also suggested that increased acidity of saliva due to smoking may give rise to acidophilic micro-organisms including Bacteria and Archaea [13].

Next generation sequencing is a powerful tool for microbiome studies in terms of determining the microbiome composition using bacterial 16S rRNA gene sequencing. Utilizing this method has allowed analysis of a large number of samples within several hours. In general, there are two common approaches for 16S rRNA next-generation sequencing: Illumina and Ion Torrent. In the Illumina 16S metagenomic sequencing systems, two regions of the 16S rRNA gene (V3 and V4) are sequenced [14], whereas Ion Torrent Ion 16S Metagenomics Kit has the ability to amplify and sequence seven hypervariable regions (V2, V3, V4, V6-7, V8 and V9) [15]. There are significant differences among hypervariable regions on differentiation of bacteria at different taxonomic levels. For instance, *Staphylococcus aureus* and coagulase-negative *Staphylococcus* species were best distinguished by V1, while V2 and V3 were defined to be the most appropriate regions to separate most of the bacterial species except for closely related *Enterobacteriaceae*. As compared to other regions, V4, V5, V7 and V8 were found to have less power for discrimination of the bacteria at both species and genus level [16]. A current study showed that primers covering the V3/V4 region determined a higher number of taxa than those covering the V1/V2 region [17]. Since there is no single variable region that can differentiate among all bacteria, examining more than one, even all hypervariable regions are preferable.

In order to extend our knowledge on the effects of smoking on buccal microbiome composition, we studied buccal swab samples from both smoker and non-smoker volunteers followed by bacterial DNA extraction and 16S rRNA gene sequencing using the Ion Torrent system. We also aimed to evaluate the performance of each of the seven hypervariable regions in 16S rRNA gene for the determination of bacterial diversity.

METHODS

Study population

Power analysis at 99 % confidence interval was used to define the number of samples. In total. 40 volunteers were chosen, including 20 smokers and 20 non-smokers that fulfilled the following inclusion criteria for good health: no use of antibiotics in the past 3 months, no respiratory infections such as tonsillitis, pharyngitis, sinusitis, rhinitis, and no oral aphthous lesions on buccal mucosa, and absence of dental problems. All volunteers were moderate smokers (15–24 cigarettes per day) who were within the normal range of body mass index and had a Mediterranean diet. Both groups included male and female volunteers aged between 22 and 51 years (mean value=28 years). The male to female sex ratio in smokers and non-smokers was 1.2 (11/9) and 0.7 (8/12), respectively (P=0.799). The mean age of smokers and non-smokers was 30 years (22–51 years) and 26.2 years (22–40 years), respectively (P=0.355). Buccal swabs from both sides of the cheek were collected from each volunteer by the same specialist using a common protocol. All volunteers were instructed not to eat and wash mouth in the 2 h prior to sample collection. After sampling, each buccal swab was dipped into tubes containing 400 µl of PBS solution and pressed against the inner wall of the tube to transfer the micro-organisms from the cotton swab to the solution. The tubes were then labelled and used for further DNA extraction.

In addition to 40 buccal samples, a microbial community control (ZymoBIOMICSTM, CA, USA), which includes known ratios of DNA from a mixture of different species of bacteria (*Bacillus subtilis*, *Listeria monocytogenes*, *Staphylococcus aureus*, *Enterococcus faecalis*, *Lactobacillus fermentum*, *Salmonella enterica*, *Pseudomonas aeruginosa* and *Escherichia coli*) was examined in the present study. Microbial community standard sample was kindly gifted by ARDI Medikal İstanbul-Turkey.

E. coli DNA (30 µg ml^{-1}) provided with the Ion Torrent 16S Metagenomics Kit was used as an internal control.

DNA extraction

DNA extraction was performed on 200 µl of samples according to the 'DNA Purification from Buccal Swabs' protocol from the Thermo Scientific GeneJET Genomic DNA Purification Kit [18]. During DNA isolation, sterility conditions were tightly applied to prevent environmental contamination. DNA concentrations were measured using the NanoDrop 2000 C (Thermo Scientific, DE, USA) and 10 ng of each DNA sample was then used for amplification of the 16S rRNA gene.

Amplification of 16S rRNA hypervariable regions and library preparation

The Ion Torrent 16S Metagenomics Kit (Thermo Fisher Scientific, Warrington, England) was used to amplify 16S rRNA genes of all the samples. Following the amplification conditions indicated in the manufacturer's protocol, PCR was performed in two pools, each including specific primer sets. Pool 1 contained primers targeting the V2-4-8 regions whereas pool 2 contained those targeting the V3-6, 7–9 regions. Amplicon lengths of V2, V3, V4, V6-7, V8 and V9 hypervariable regions were 250, 215, 288, 260, 295 and 209, respectively [19]. After PCR amplification, equal volumes (20 µl) of PCR products from pools 1 and 2 of each sample were combined into a single PCR tube. Then 40 µl of combined

PCR product were purified using the Agencourt AmpureXP kit according to the manufacturer's instructions (Beckman Coulter, Brea, CA, USA). After DNA quantification using the Qubit Fluorometer with the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Warrington, England), the DNA library was prepared using Ion Plus Fragment Library Kit (Thermo Fisher Scientific, Austin, TX, USA) and Ion Xpress Barcodes Adapters 1–16 KitTM (Life Technologies, Carlsbad, CA, USA) according to the manufacturer's protocol. The DNA library was purified with Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA). After a five-cycle PCR reaction and purification with Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA), library concentrations were determined using the Qubit Fluorometer. Each library was diluted to a concentration of 16 pM and equal volumes of each library were pooled.

Template preparation and next-generation sequencing

The Ion 520 and Ion 530 Kit (Thermo Fisher Scientific, Austin, TX, USA) for the Ion Chef and Ion S5 workflow was used in this study. The Ion ChefTM instrument clonally amplified the pooled library on nanosized ionosphere particles by emulsion PCR. Bead enrichment and chip loading were also conducted using the Ion ChefTM Instrument. The massively parallel sequencing reaction was achieved on the Ion S5TM Next-Generation Sequencing system with 950 flows.

Data processing and quality assessment

Analysis of results was performed using the Ion Reporter Software on the Ion 16S Metagenomics Kit analysis module. In the scope of analysis, the reads were mapped to two reference 16S rRNA databases namely Greengenes v13.5 and MicroSEQ ID v3.0. [20]. Raw sequence data were processed to eliminate low-quality reads. The following default parameters of Ion Reporter Software were applied to hit high confidence target species. After trimming of the primer sequences from both ends, we filtered out reads shorter than 200 bases. Moreover, reads with alignment coverage less than 90.0 % were discarded. Reads having a number of unique reads greater than 500 were considered as valid. Chimeric sequences were automatically identified and removed by Ion Reporter software [21].

OTU clustering

In order to cluster the sequences, operational taxonomic unit (OTU) analysis was performed using Ion Reporter Software, which runs Quantitative Insights into Microbial Ecology (QIIME). Genus and species cut-off values were accepted at 97 and 99 %, respectively in QIIME.

Diversity analysis

Alpha and beta diversity graphics created by QIIME software were exported from the Ion Reporter Software. For alpha diversity analysis the observed species and the Chao1 and Shannon indexes were generated to analyse species diversity within samples. Chao1 and Shannon are two non-parametric indexes that estimate the richness and diversity of the species, respectively. Observed species calculates the amount of OTUs observed in each sample [22]. According to these three parameters, box and whisker plots were created based on the species level at 99 % similarity. In the beta diversity analysis, cluster analysis was performed using three-dimensional principal coordinates analysis (PCoA) with Bray–Curtis dissimilarity index.

Statistical analysis

Relative abundances were calculated by normalizing the OTU table of raw counts provided by Ion Reporter Software. The taxa were clustered at phylum, family, genus and species levels. Non-parametric Mann–Whitney test was used to determine statistically significant differences in the relative abundances between smoker and non-smoker groups for the phylum, genus and species comprising 1 % or more of the total abundance across samples. We also used the Mann–Whitney test to compare alpha diversity between the smoker and non-smoker groups. The Benjamini–Hochberg false discovery rate (FDR) protocol was performed after performing the Mann–Whitney test.

RESULTS

Comparison of the efficiency of seven hypervariable regions

We evaluated the performance of each hypervariable region in terms of mapped read counts and in determining microbial composition. Amplicons varying between 200–300 bp in length were sequenced in a total of 19 855 430 reads. A total of 15 661 083 valid reads were obtained after data processing. Valid reads were mapped to Greengenes v13.5 and MicroSEQ ID v3.0 databases, resulting in a total of 11 095 228 mapped reads. The mean value of mapped reads was 277 380. Mapped read counts were not homologous for each hypervariable region. For the majority of samples, V3 region showed the highest number of reads (36.3 %), followed by V4 (16.7 %), V8 (16 %), V6-7 (15.3 %), V2 (8.7 %) and V9 (6.7 %) regions.

Each hypervariable region was analysed for its ability to determine the buccal microbiome composition according to genera. Genera with relative abundances less than 1 % were excluded in this analysis. Variations were detected in terms of microbiome diversity and relative abundance between different regions. Neither region was found to be capable of identifying all detected genera nor showed homology with consensus data. The hypervariable V2, V3, V4, V6-7, V8, V9 regions and consensus of all seven regions detected 16, 11, 10, 10, 9, 5 and 21 genera, respectively. Particularly, the V3 region was not informative for detecting *Fusobacterium* even though it possessed the highest mapped read percentage. *Haemophilus* was the only genus that could be distinguished by all hypervariable regions. The genus *Rothia*, *Prevotella*, *Gemella*, *Streptococcus*, *Veillonella* and *Neisseria* could be identified by both V2 and V3 regions whereas the V4 region could not identify *Rothia* and *Veillonella* among these genera. Remarkably, *Mannheimia*, *Thioalkalimicrobium* and *Zooshikella* were exclusively detected in the V9 region only (Fig. 1).

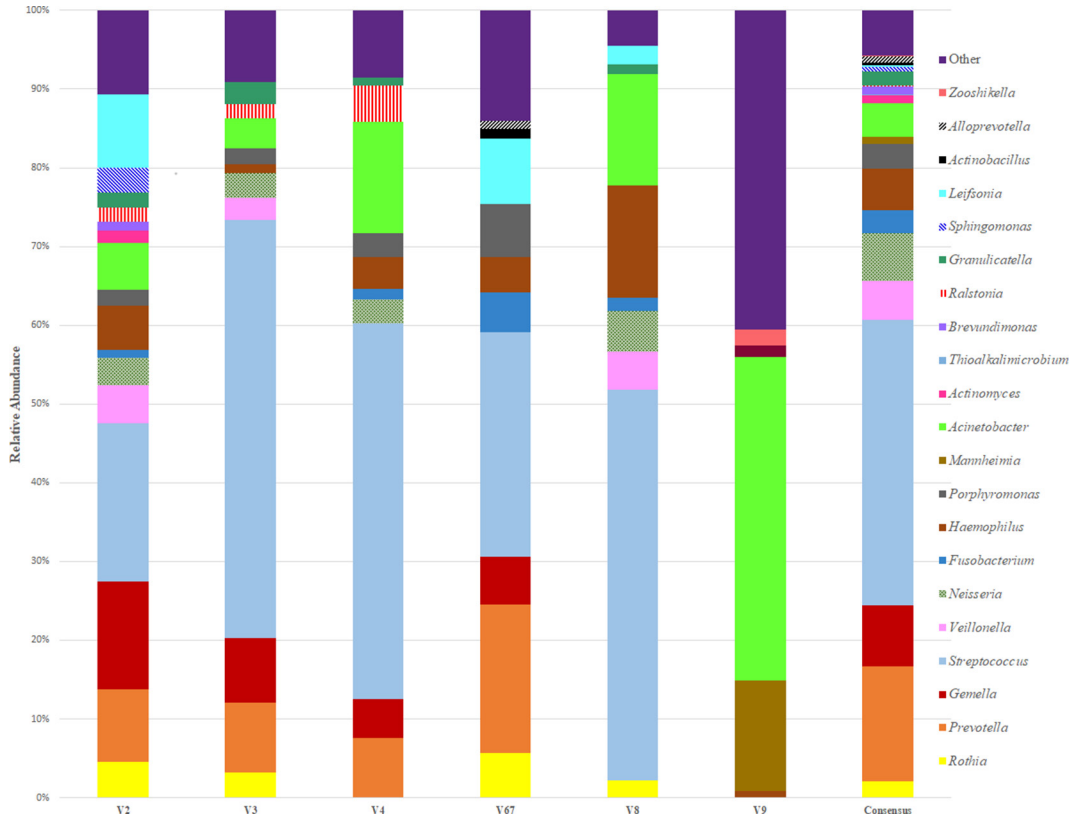

Fig. 1. Relative abundance of different genera according to mapped reads for each hypervariable region of 16S rRNA gene. Consensus data shows the overall microbiome distribution for each sample according to mapped reads generated from seven hypervariable regions. [*Y* axis represents relative abundance (%) value and *X* axis shows each hypervariable region and consensus.]

Microbiome composition in the tested samples

Before analysing the buccal sample results, we evaluated the relative abundances of the bacteria in the microbial community standard sample. Our results showed close homology with the microbial composition in the standard sample especially at genus level, except *E. coli* that could not be identified. Detecting no taxa other than those present in the standard sample confirmed the absence of any external contamination in our study.

We analysed mapped and valid read numbers for each sample in order to verify that the results were not affected by specific individuals. The numbers of valid and mapped reads per individual varied from 67 672 to 11 736 64 and from 44 425 to 878 228, respectively. While the mean value for valid reads was 391 527 with a standard deviation of 310 193, it was 277 380 (SD: 228 335) for mapped reads (Table 1). The numbers of valid and mapped reads were also compared between smoker and non-smokers, and no significant difference was recorded between two groups (Fig. 2).

Five dominant phyla including Firmicutes, Bacteroides, Actinobacteria, Fusobacteria and Proteobacteria were found in all buccal samples, and they accounted for 99.76 % of the total tags. Comparison of relative abundances of these phyla between smoker and non-smoker groups showed no statistically significant differences (all *P*-adj values >0.099) (Fig. 3A). According to a comparative analysis of the relative abundances of different genera, *Veillonella* and *Actinomyces* showed statistically significant increase in smokers with *P*-adj values of 0.039 (Fig. 3B).

In total, 13 different genera and 25 species comprising ≥1 % of the total abundance across all samples were used for the comparison of smokers and non-smokers. After adjusting *P*-values using the Benjamini–Hochberg FDR procedure, only the relative abundance of *Veillonella* and *Actinomyces* was found to have a significant increase in smokers compared to non-smokers (*P*-adj=0.039). Similarly, relative abundances of *Veillonella atypica, Streptococcus australis*, *Prevotella salivae*, *Prevotella melaninogenica* and *Rothia mucilaginosa* showed significant increases in the smoker group (*P*-adj=0.05) (Table 2).

Alpha diversity analysis showed no statistically significant difference in the number of observed species between smokers and non-smokers (*P*=0.265). Similarly, there was no significant difference in community richness and diversity between the two groups as indicated by the *P*-values of Chao1 and Shannon indexes such as 0.254 and 0.06, respectively (Fig. 4).

Table 1. Numbers of valid and mapped reads for each sample of the study group

Sample no.	Valid read	Mapped read
Sample 1	231 751	151 937
Sample 2	350 288	232 208
Sample 3	418 439	281 068
Sample 4	793 058	580 246
Sample 5	104 1881	726 410
Sample 6	10 161 49	734 899
Sample 7	10 049 75	771 942
Sample 8	765 547	584 053
Sample 9	764 671	541 988
Sample 10	10 067 61	675 151
Sample 11	11 736 64	878 228
Sample 12	830 316	583 811
Sample 13	143 241	89 473
Sample 14	138 383	84 517
Sample 15	318 484	196 692
Sample 18	286 262	173 551
Sample 21	317 939	237 573
Sample 22	243 510	193 909
Sample 24	187 042	143 235
Sample 25	234 578	183 820
Sample 26	324 799	263 232
Sample 27	430 081	280 678
Sample 29	325 716	247 569
Sample 30	283 966	232 216
Sample 31	296 479	203 663
Sample 32	252 985	197 422
Sample 33	161 675	113 175
Sample 34	123 641	71 330
Sample 35	179 466	130 448
Sample 36	328 893	249 258
Sample 37	105 054	74 585
Sample 38	110 212	74 679
Sample 39	124 026	73 160
Sample 40	198 284	122 150
Sample 41	144 115	104 265
Sample 42	67 672	44 425
Sample 45	156 579	92 931
Sample 51	278 625	175 786
Sample 52	200 297	116 639
Sample 53	301 579	182 906

Mean values for valid and mapped reads were 391 527 and 277 380, respectively. For valid reads, minimum and maximum values were 67 672 and 11 736 64 while they were 44 425 and 878 228 for mapped reads, respectively. Standard deviation was calculated as 310 193 for valid reads and 228 335 for mapped reads.

Principal coordinate analysis (PCoA) based on Bray–Curtis dissimilarity was performed to visualize the relationship among the microbiome communities in the 40 samples. Based on a weighted Bray–Curtis PCoA, except six samples, the remaining 14 samples taken from non-smokers were grouped to the bottom of the graph along PC1, whereas 14 of the samples taken from smokers were grouped to the top of the graph along PC1 (Fig. 5).

DISCUSSION

Massive parallel sequencing using next-generation sequencing technology has boosted whole genome sequencing and has proved a powerful approach for metagenomic analysis [23, 24]. The 16S Metagenomics sequencing kit of Ion Torrent has an advantage that allowed us to amplify and sequence seven hypervariable regions of the 16S rRNA gene. Using this kit, the current study provides data about the composition of the buccal microbiome in both smoker and non-smoker individuals. Importantly, this is the first study to evaluate the performance of each of the seven hypervariable regions of the 16S rRNA gene in determining the diversity of buccal microbiome. As expected, each hypervariable region resulted in varying mapped reads and microbiome composition. Although the V3 region had the highest percentage of mapped reads, it was found to be less informative compared to the V2, which was the most informative region, detecting 16 of 21 genera found in the consensus of all seven hypervariable regions. Each of V4, V6-7 and V8 regions was able to detect only half of the genera found by all hypervariable regions. On the other hand, V9 was the less informative region identifying only one-fourth of all genera detected. Since no hypervariable region represented the microbiome composition that was similar to those provided by the consensus data, sequencing seven hypervariable regions would be the best interest for comprehensive studies.

After testing the discriminative abilities of hypervariable regions, we performed several comparative analyses between the smoker and non-smoker groups. Alpha diversity analysis with Chao1 and Shannon indexes showed no significant difference in species richness and microbiome diversity between the samples from smokers and non-smokers (all *P*-values>0.06). Chao1 is an abundance-based estimator of

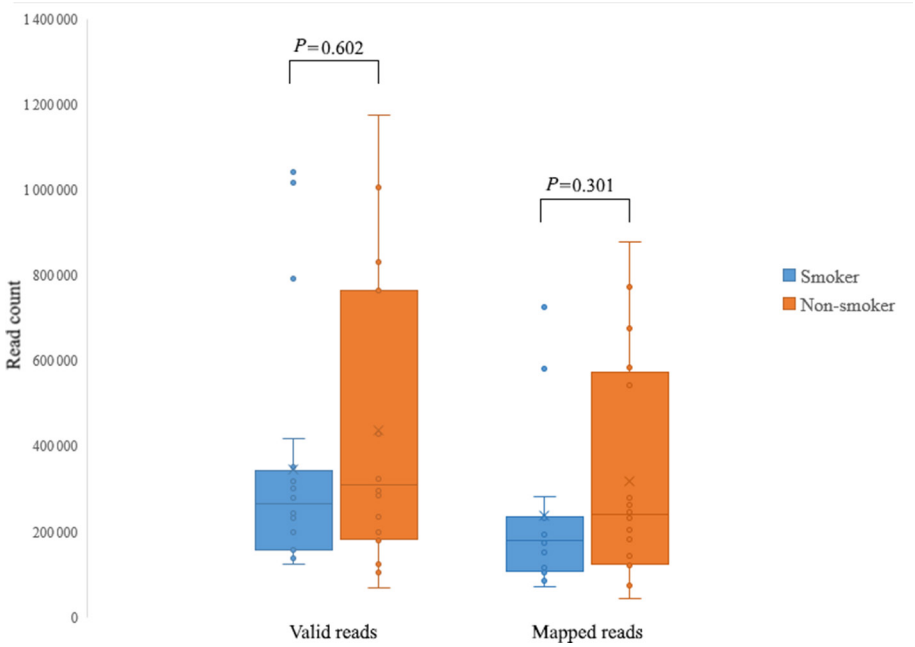

Fig. 2. Box and whisker plot of the valid and mapped reads for smoker and non-smoker groups. No statistically significant differences were found for both valid and mapped reads with *P*-values of 0.602 and 0.301, respectively.

species richness [25], whereas Shannon index represents species diversity [26]. In contrast, beta diversity analysis revealed that samples from both smoker and non-smoker groups tended to cluster within themselves.

Although the human oral microbiota includes many different micro-organisms, Actinobacteria, Bacteroidetes, Firmicutes, Fusobacteria and Proteobacteria are considered as the most significant phyla [23]. In agreement with the published literature, these five phyla were found in all 40 buccal samples with a total relative abundance of 99.76 %. A previous publication, which studied oral wash samples, reported that the relative abundance of Proteobacteria was

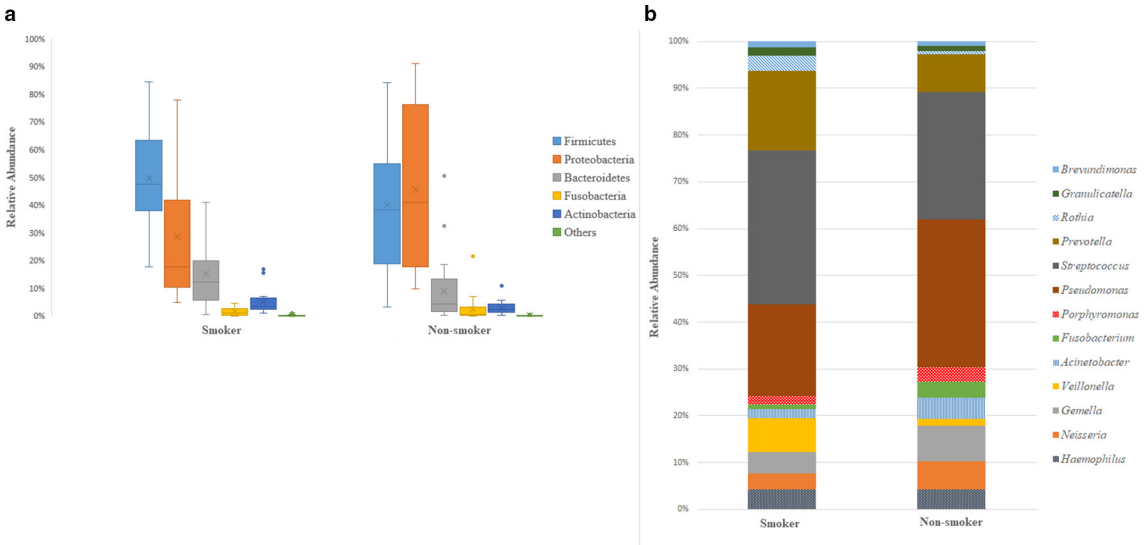

Fig. 3. Box and whisker plot of the relative abundance of dominant phyla in smokers and non-smokers (a), stacked taxonomic bar chart at genus level (b). At the phylum level, no statistically significant difference was found between smokers and non-smokers whereas relative abundance of genera *Veillonella* and *Actinomyces* significantly increased in smokers (*P*-adj=0.039).

Table 2. Comparison of statistically significant differences between smoker and non-smoker groups at genus and species levels

Mean relative abundance				
Genus	**Smoker**	**Non-smoker**	***P*-value**	***P*-adj**
Actinomyces	1,62 %	0,59 %	**0.003**	**0.039**
Veillonella	8,29 %	2,00 %	**0.003**	**0.039**
Neisseria	3,85 %	7,87%	**0.012**	0.073
Rothia	3,67 %	0,85 %	**0.017**	0.073
Prevotella	19,10 %	10,68%	**0.046**	0.149
Acinetobacter	2,23 %	5,88 %	0.121	0.314
Porphyromonas	1,98 %	4,21 %	0.165	0.323
Streptococcus	37,06 %	35,77 %	0.174	0.323
Gemella	5,07 %	9,95 %	0.565	0.758
Granulicatella	2,15 %	1,48 %	0.565	0.758
Haemophilus	4,81 %	5,72 %	0.583	0.758
Fusobacterium	1,05 %	4,48 %	0.602	0.758
Brevundimonas	1,10 %	0,95 %	0.758	0.758
Species	**Smoker**	**Non-smoker**	***P*-value**	***P*-adj**
Streptococcus australis	2,40 %	0,20%	**0.003**	**0.05**
Prevotella salivae	3,28 %	1,13 %	**0.003**	**0.05**
Prevotella melaninogenica	18,43 %	10,82 %	**0.004**	**0.05**
Rothia mucilaginosa	4,56 %	1,05 %	**0.004**	**0.05**
Veillonella atypica	7,85 %	1,03 %	**0.006**	**0.05**
Prevotella histicola	11,34 %	6,00 %	**0.015**	0.093
Streptococcus sinensis	2,46 %	0,15%	**0.028**	0.14
Streptococcus salivarius	5,66 %	0,37 %	**0.028**	0.14
Veillonella dispar	2,70 %	0,32 %	**0.046**	0.191
Prevotella veroralis	2,12 %	1,20 %	0.091	0.325
Haemophilus parainfluenzae	9,63 %	14,89 %	0.174	0.527
Gemella haemolysans	6,78 %	19,69 %	0.192	0.527
Neisseria flavescens	1,79 %	6,39 %	0.211	0.527
Granulicatella adiacens	2,56 %	2,24 %	0.253	0.575
Streptococcus pseudopneumoniae	1,90 %	1,98 %	0.414	0.862
Streptococcus mitis	2,45 %	4,68 %	0.565	0.989
Neisseria subflava	1,63 %	1,22%	0.583	0.989
Mannheimia varigena	2,49 %	2,92 %	0.698	0.989
Rothia dentocariosa	3,09 %	1,41 %	0.841	0.989
Fusobacterium nucleatum	0,79 %	4,46 %	0.883	0.989
Porphyromonas gingivalis	0,27 %	3,02 %	0.904	0.989
Granulicatella elegans	2,49 %	2,97 %	0.925	0.989
Prevotella denticola	0,25 %	3,94 %	0.925	0.989

Continued

Table 2. Continued

Species	Smoker	Non-smoker	*P*-value	*P*-adj
Fusobacterium periodonticum	0,82 %	3,65 %	0.989	0.989
Brevundimonas aurantiaca	1,56 %	1,80 %	0.989	0.989

The non-parametric Mann–Whitney test resulted in significant differences for five genera and nine species but after false discovery analysis performed with Benjamini–Hochberg method two genera (*Veillonella* and *Actinomyces*) and five species (*S. australis, P. salivae, P. melaninogenica, R. mucilaginosa* and *V. atypica*) showed statistically significant differences. All were labelled as bold.

decreased in smoker individuals [10]. We found a similar decrease in the relative abundance of Proteobacteria in smokers, but the result was not statistically significant (*P*-adj=0.099). As for Actinobacteria, although there was a remarkable increase in its relative abundance in the smoker group compared to that in non-smokers, this difference was not statistically significant after Benjamini–Hochberg correction (*P*-adj=0.099). This result was also in agreement with the results of earlier studies performed on oral rinse samples [10] and subgingival plague samples [11]. It was previously reported that phylum Firmicutes was enriched in the oral microbiota of smokers [10]. However, no significant difference was observed in our study. One reason for these differences might be the difference on sample types, instead

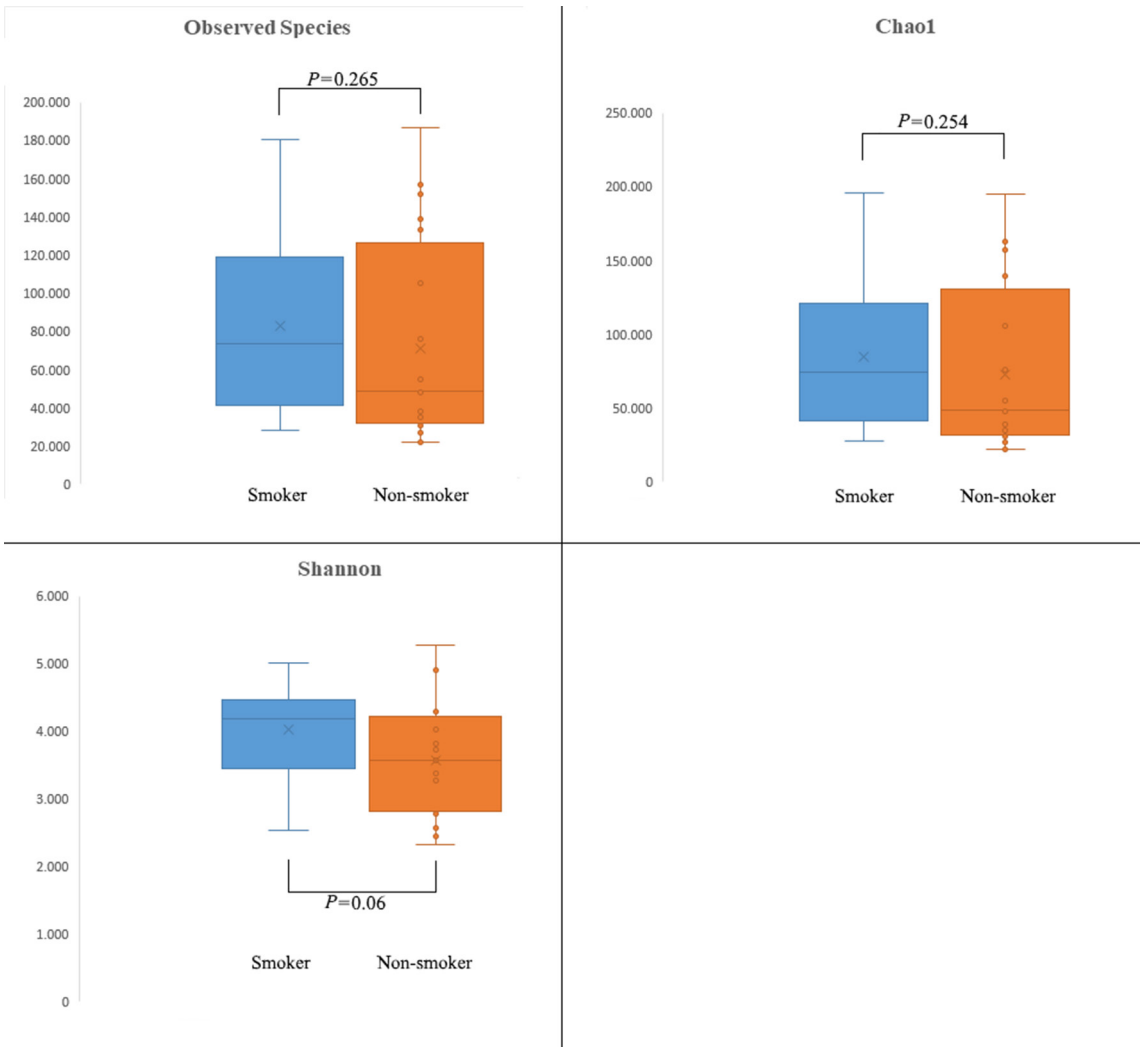

Fig. 4. Alpha diversity analysis at species level: Chao1, observed species and Shannon diversity indexes between smokers (blue) and non-smokers (red). No statistically significant difference was found between two groups (*P*=0.05).

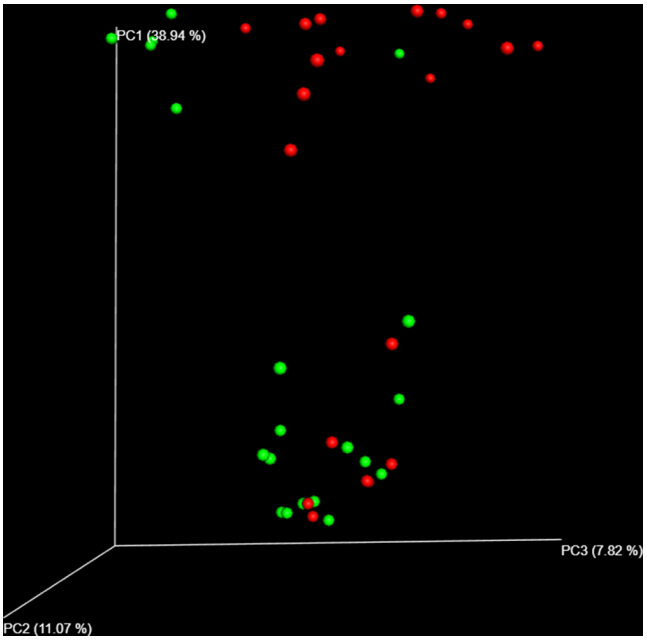

Fig. 5. Beta diversity analysis at species level: principal coordinate analysis (PCoA) based on Bray–Curtis dissimilarity between buccal microbiome of smokers (red) and non-smokers (green).

of buccal swabs used in our study, oral wash samples were analysed in the other study [10]. Furthermore, use of different sequencing technologies might be another reason for these differences. Instead of the Ion Torrent Metagenomics kit including seven hypervariable regions used in the current study, Wu *et al*. [10] utilized the Illumina kit searching only two hypervariable regions (V3 and V4) in the 16S rRNA gene.

We observed statistically significant increases in the relative abundances of genus *Actinomyces* and *Veillonella* (*P*-adj=0.039) and species *V. atypica*, *P. melaninogenica* and *R. mucilaginosa* (*P*-adj=0.05) in the smoker group. These species were previously found to be involved in oral nitrate reduction, a reaction that cannot be performed by mammalian cells [27, 28]. Under anaerobic conditions they use nitrate as a terminal respiratory electron acceptor to produce nitrite, which has a toxic effect when converted to carcinogenic nitrosamines and pro-inflammatory nitric oxide [27, 29]. Oral tobacco usage increases the intake of nitrate and alkaloids that can form carcinogenic nitrosamines. The nitrate-reducing metabolic activity of these bacteria may thus have important consequences on health [30]. Increased abundance of *V. atypica*, *P. melaninogenica* and *R. mucilaginosa* in our study suggests that increased nitrate intake due to smoking may have created an environment favouring nitrate-reducing species in the oral cavity. We also showed increase in the relative abundances of other nitrate-reducing bacteria such as *R. dentocariosa* and *V. dispar* in smokers. Thus, our study indicates a possible relationship between cancer and smoking-related changes in the buccal microbiota composition.

H2S-producing bacteria including species of the genera *Veillonella*, *Actinomyces* and *Prevotella* in the tongue biofilm were previously defined as a source of oral malodour [31]. Statistically significant increase in the genera *Veillonella* and *Actinomyces* and also species *V. atypica*, *P. melaninogenica* and *P. salivae* in smokers might be one of the reasons for oral malodour in smokers.

In addition, the facultative anaerobe *Streptococcus australis* species, which has tolerance for acidic environments, was enriched in smokers (*P*-adj=0.05), supporting the hypothesis that smoking cigarettes increases the acidity of saliva [13]. In contrast, our analysis revealed that the anaerobic species *Fusobacterium nucleatum* remarkably decreased in smokers even though previous studies have reported both increased [11] and decreased [29] abundances of *F. nucleatum* in the oral cavity.

Oral microbiome can be affected by various factors [24, 32]. In this regard, diet type and intake of different macronutrients plays an important role in the characteristics of oral microbiota [33]. In addition, saturated fatty acids (SFA) and vitamin C, B and E intake were found to be positively correlated with the richness and diversity of oral microbiota, specifically for Fusobacteria [34]. Socioeconomic status is also an important factor affecting the oral microbiota [35]. Furthermore, oral hygiene and high body mass index have significant effect on prevalence and diversity of oral microbiome [36]. While oral microbiome shows considerable variation between subjects, many studies analysing the time-dependent variation in oral microbiome in orally healthy individuals revealed a long-term stability or little change within subjects [24, 37–40]. Another important factor that affects oral microbiome is smoking [10, 11, 41, 42]. Although we were not able to get detailed information about all possible confounding factors, both smokers and non-smokers subjected in the current study had similar sex and age groups, oral health conditions, body mass indexes and Mediterranean diet type. For these reasons, it can be assumed that differences observed on the buccal microbiome between two groups was most probably related to smoking.

In conclusion, the current study confirmed that Firmicutes, Bacteroidetes, Actinobacteria, Fusobacteria and Proteobacteria were the five major phyla encountered in both smoker and non-smoker subjects. It is likely that smoking changes the buccal microbiome at different taxonomic levels. At the phylum level, no statistically significant difference was found between smokers and non-smokers whereas *Veillonella* and *Actinomyces* showed higher abundance in smokers compared to non-smokers. Moreover, relative abundances of *V. atypica*, *P. melaninogenica*, *P. salivae*, *R. mucilaginosa* and *S. australis* showed statistically significant increase in the smoker group. Although alpha diversity analysis did not indicate any notable differences between smokers and non-smokers, beta diversity analysis with Bray–Curtis dissimilarity index suggested that smoking status may be a contributing factor for community composition in the buccal microbiome. The V2 region was found to be the informative hypervariable region for

determining microbial composition at genus level. However, if all seven regions can be analysed, their consensus data provided the most comprehensive results.

Funding information
This study was supported by Ankara Yildirim Beyazit University Research Foundation, Project No. 4806.

Author contributions
Dr Rıza Durmaz: Management and coordination responsibility for the research activity planning and execution. Preparation of the published work, specifically critical review of the manuscript. Formulation and evaluation of research goals and aims. Sema Karabudak: Conducting a research and investigation process, specifically performing the experiments and data collection. Writing the initial draft, evaluation of the experimental data. Oguz Ari: Conducting a research and investigation process, specifically performing the experiments and data analysis. Writing the initial draft, evaluation of the experimental data. Dr Bengul Durmaz: Performing experiments, planning research activity and critical review of the manuscript. Dr Tuba Dal: Performing experiments and review of the manuscript. Tugcan Basyiyit: Analysis of the raw data and writing the initial draft. Dr Mahmut Tayyar Kalcioglu: Management and coordination responsibility for the research activity planning, selection of the study groups.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
This study was approved by the ethics committee of Inonu University, Medical Faculty, Malatya/Turkey (Protocol code: 2011/124). Verbal or written consent of the smoker and non-smoker volunteers enrolled in this study was obtained before sample collection.

References

1. Peterson J, Garges S, Giovanni M, McInnes P, Wang L *et al*. The NIH human microbiome project. *Genome Res* 2009;19:2317–2323.
2. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner AC *et al*. The human oral microbiome. *J Bacteriol* 2010;192:5002–5017.
3. Jia G, Zhi A, Lai PFH, Wang G, Xia Y *et al*. The oral microbiota – a mechanistic role for systemic diseases. *Br Dent J* 2018;224:447–455.
4. Bradshaw DJ, Marsh PD. Analysis of pH-driven disruption of oral microbial communities *in vitro*. *Caries Res* 1998;32:456–462.
5. Sachdeo A, Haffajee AD, Socransky SS. Biofilms in the edentulous oral cavity. *J Prosthodont* 2008;17:348–356.
6. Sampaio-Maia B, Caldas IM, Pereira ML, Perez-Mongiovi D, Araujo R. The oral microbiome in health and its implication in oral and systemic diseases. *Adv Appl Microbiol* 2016;97:171–210.
7. Papaioannou W, Gizani S, Haffajee AD, Quirynen M, Mamai-Homata E *et al*. The microbiota on different oral surfaces in healthy children. *Oral Microbiol Immunol* 2009;24:183–189.
8. Marsh PD, Martin MV, Lewis MAO, Williams D. *Oral Microbiology*, 5th ed. Edinburgh: Elsevier; 2009.
9. Mattila KJ, Pussinen PJ, Paju S. Dental infections and cardiovascular diseases: a review. *J Periodontol* 2005;76:2085–2088.
10. Wu J, Peters BA, Dominianni C, Zhang Y, Pei Z *et al*. Cigarette smoking and the oral microbiome in a large study of American adults. *ISME J* 2016;10:2435–2446.
11. Mason MR, Preshaw PM, Nagaraja HN, Dabdoub SM, Rahman A *et al*. The subgingival microbiome of clinically healthy current and never smokers. *ISME J* 2015;9:268–272.
12. Sapkota AR, Berger S, Vogel TM. Human pathogens abundant in the bacterial metagenome of cigarettes. *Environ Health Perspect* 2010;118:351–356.
13. Kanwar A, Sah K, Grover N, Chandra S, Singh R. Long-term effect of tobacco on resting whole mouth salivary flow rate and pH: an institutional based comparative study. *European J Gen Dent* 2013;2:296–299.
14. 16S metagenomic sequencing library preparation-preparing 16S ribosomal RNA gene amplicons for the Illumina MiSeq system. https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf
15. Ion torrent ion 16S™ metagenomics kit user guide; 2018. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0010800_Ion_16S_Metagenomics_Kit_QR.pdf
16. Chakravorty S, Helb D, Burday M, Connell N, Alland D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J Microbiol Meth* 2007;69:330–339.
17. Graspeuntner S, Loeper N, Künzel S, Baines JF, Rupp J. Selection of validated hypervariable regions is crucial in 16S-based microbiota studies of the female genital tract. *Sci Rep* 2018;8:9678.
18. Thermo scientific GeneJET genomic DNA purification kit; 2016. https://www.thermofisher.com/order/catalog/product/K0721
19. Chaitali Parikh EB, Stephen J. Rapid 16S next generation sequencing for bacterial identification in polymicrobial samples: Thermo Fisher Scientific; 2018. https://www.slideshare.net/ThermoFisher/c-parikh-asm2015posterfinal
20. Ion Reporter™ 5.0 software: Thermo Fisher Scientific. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/IonReporter_ v50_Help.pdf
21. Saeb ATM, Al-Rubeaan KA, Aldosary K, Udaya Raja GK, Mani B *et al*. Relative reduction of biological and phylogenetic diversity of the oral microbiota of diabetes and pre-diabetes patients. *Microb Pathog* 2019;128:215–229.
22. Zhang J, Sun QL, Zeng ZG, Chen S, Sun L. Microbial diversity in the deep-sea sediments of Iheya North and Iheya ridge, Okinawa Trough. *Microbiol Res* 2015;177:43–52.
23. Zarco MF, Vess TJ, Ginsburg GS. The oral microbiome in health and disease and the potential impact on personalized dental medicine. *Oral Dis* 2012;18:109–120.
24. Lazarevic V, Whiteson K, Hernandez D, Francois P, Schrenzel J. Study of inter- and intra-individual variations in the salivary microbiota. *BMC Genomics* 2010;11:523.
25. Chao A. Non-parametric estimation of the classes in a population. *Scand J Stat* 1984;11:265–270.
26. Morris EK, Caruso T, Buscot F, Fischer M, Hancock C *et al*. Choosing and using diversity indices: insights for ecological applications from the German biodiversity exploratories. *Ecol Evol* 2014;4:3514–3524.
27. Doel JJ, Benjamin N, Hector MP, Rogers M, Allaker RP. Evaluation of bacterial nitrate reduction in the human oral cavity. *Eur J Oral Sci* 2005;113:14–19.
28. Liddle L, Burleigh MC, Monaghan C, Muggeridge DJ, Sculthorpe N *et al*. Variability in nitrate-reducing oral bacteria and nitric oxide metabolites in biological fluids following dietary nitrate administration: an assessment of the critical difference. *Nitric Oxide* 2019;83:1–10.
29. Karasneh JA, Al Habashneh RA, Marzouka NAS, Thornhill MH. Effect of cigarette smoking on subgingival bacteria in healthy subjects and patients with chronic periodontitis. *BMC Oral Health* 2017;17:64.
30. Kato I, Vasquez AA, Moyerbrailean G, Land S, Sun J *et al*. Oral microbiome and history of smoking and colorectal cancer. *J Epidemiol Res* 2016;2:92–101.
31. Washio J, Sato T, Koseki T, Takahashi N. Hydrogen sulfide-producing bacteria in tongue biofilm and their relationship with oral malodour. *J Med Microbiol* 2005;54:889–895.
32. Crielaard W, Zaura E, Schuller AA, Huse SM, Montijn RC *et al*. Exploring the oral microbiota of children at various developmental stages of their dentition in the relation to their oral health. *BMC Med Genomics* 2011;4:22.
33. Tanner AC, Kent RL Jr, Holgerson PL, Hughes CV, Loo CY *et al*. Microbiota of severe early childhood caries before and after therapy. *J Dent Res* 2011;90:1298–1305.

34. Kato I, Vasquez A, Moyerbrailean G, Land S, Djuric Z *et al*. Nutritional correlates of human oral microbiome. *J Am Coll Nutr* 2017;36:88–98.
35. Belstrøm D, Holmstrup P, Nielsen CH, Kirkby N, Twetman S *et al*. Bacterial profiles of saliva in relation to diet, lifestyle factors, and socioeconomic status. *J Oral Microbiol* 2014;6:23609.
36. Yamashita Y, Takeshita T. The oral microbiome and human health. *J Oral Sci* 2017;59:201–206.
37. Belstrøm D, Holmstrup P, Bardow A, Kokaras A, Fiehn NE *et al*. Temporal stability of the salivary microbiota in oral health. *Plos One* 2016;11:e0147472–e.
38. Utter DR, Mark Welch JL, Borisy GG. Individuality, stability, and variability of the plaque microbiome. *Front Microbiol* 2016;7:564.
39. Rasiah IA, Wong L, Anderson SA, Sissons CH. Variation in bacterial DGGE patterns from human saliva: over time, between individuals and in corresponding dental plaque microcosms. *Arch Oral Biol* 2005;50:779–787.
40. Zaura E, Nicu EA, Krom BP, Keijser BJF. Acquiring and maintaining a normal oral microbiome: current perspective. *Front Cell Infect Microbiol* 2014;4:85.
41. Moon JH, Lee JH, Lee JY. Subgingival microbiome in smokers and non-smokers in Korean chronic periodontitis patients. *Mol Oral Microbiol* 2015;30:227–241.
42. Brook I. The impact of smoking on oral and nasopharyngeal bacterial flora. *J Dent Res* 2011;90:704–710.