Original Article

Oral microbiome profiles in oral potentially malignant disorders and oral cancer – A diagnostic perspective

Priya N. Deo[1#], Revati S. Deshmukh[1], Akshay H. Gaike[2,3,4#], Anu Christopher[5], Mohak Gujare[2,6], Mitali Inamdar[2]

[1]Department of Oral and Maxillofacial Pathology and Oral Microbiology, Bharati Vidyapeeth Deemed to be University, Dental College and Hospital, Pune, Maharashtra, [2]National Centre for Cell Science, Pune, Maharashtra, [5]Department of Pathology, Bharati Hospital and Research Centre, Pune, Maharashtra, [3]Department of Zoology, Savitribai Phule Pune University, Pune, Maharashtra, India, Current Affiliations: [4]Department of Cellular, Computational and Integrative Biology, University of Trento, Trento, Italy, [6]Host-Microbe Interactomics, Wageningen University and Research, Wageningen, Netherlands

[#]These authors contributed equally.

Abstract

Background: Dysregulation of the oral microbiome has been correlated with many diseases, but oral microbiome in the etiopathogenesis of oral cancer remains a grey area and needs to be explored. It is imperative to understand the oral microbiome profiles so as to know the variations in the composition from normal to pre-cancer to cancer.

Aim: To profile the oral microbiome of normal, oral potentially malignant disorders (leukoplakia – Leu, oral submucous fibrosis – OSMF) and oral squamous cell carcinoma (OSCC) by Next-Generation Sequencing of the 16S ribosomal rRNA gene.

Material and Methods: This is an observational cross-sectional study. A total of 50 subjects were selected for this study, which included the the normal, Leukoplakia, OSMF, and OSCC groups. Bacterial genomic DNA was extracted, and 16S rRNA gene sequencing of the V4 region was carried out using the Illumina MiSeq system. Bio-informatics data analysis was carried out using the DADA2 pipeline and phyloseq R package, and the *t*-test was used for statistical analysis.

Results and Conclusion: Variations in the composition of the oral microbiome were identified across all study groups, and significant differences were noted in certain microbial taxa across normal, pre-cancer, and cancer. Certain bacterial taxa were detected only in OSCC. An increase in relative abundance of Gram-negative bacteria as well as an increasing trend in the abundance of periodontal taxa was observed in OSCC. This study generated a baseline data which may provide a guideline for future functional and integrative oral microbiome studies. Variations in oral microbiome composition may be used as biomarkers and provide signatures during the progression from normal to pre-cancer to cancer.

Keywords: 16S rRNA gene, next-generation sequencing, oral cancer, oral microbiome

Address for correspondence: Dr. Priya N. Deo, Department of Oral and Maxillofacial Pathology and Oral Microbiology, Bharati Vidyapeeth Deemed to be University, Dental College and Hospital, Pune, Maharashtra, India.
E-mail: priyanimishdeo@gmail.com
Akshay H. Gaike: https://orcid.org/0000-0002-5355-3018

Submitted: 30-May-2024, **Revised:** 12-Aug-2024, **Accepted:** 17-Jan-2025, **Published:** 28-Mar-2025

Access this article online

Quick Response Code:

Website: https://journals.lww.com/JPAT/

DOI: 10.4103/jomfp.jomfp_140_24

How to cite this article: Deo PN, Deshmukh RS, Gaike AH, Christopher A, Gujare M, Inamdar M. Oral microbiome profiles in oral potentially malignant disorders and oral cancer – A diagnostic perspective. J Oral Maxillofac Pathol 2025;29:87-97.

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is the most frequent oral cancer accounting for 90% of all oral malignancies. It is one of the leading causes of morbidity and mortality in the world.[1] The International Agency for Cancer Research has postulated that the prevalence of cancer in India will rise from 1 million in 2012 to more than 1.7 million in 2035.[2]

OSCC is a multi-factorial disease.[3] A genetic predisposition has been suggested to play a role in the aetiology of oral cancer. The majority of people who are exposed to the well-known risk factors like tobacco, alcohol, and betel quid do not develop oral cancer. Also, sporadically oral cancer cases are reported in young adults and non-tobacco and non-alcohol users.[4] Therefore, there is a need to investigate other potential risk factors that contribute to oral carcinogenesis, one of which is the microbiome. Many risk factors have been discussed, and many microbes individually also have been followed, but the vast oral microbiome still needs recognition.

The oral microbiome comprises of diverse microbial communities across different microenvironments including tooth and epithelial surfaces.[5] Each surface of the oral cavity is covered with a biofilm which constitutes a microbiome aggregate.[6]

Every individual has a unique microbiome, and a balanced microbiome is the key to maintaining health. Dysbiosis in the constitution of the microbiome has been related to many diseases including pre-cancer and cancer.[7]

Increasing research and significant evidence has emerged implicating bacteria in the aetiology of certain cancer types; for example, *Helicobacter pylori* is associated with gastric cancer, *Chlamydia trachomatis* is linked to cervical cancer, and *Salmonella typhi* to gall bladder cancer.[8]

The connection between microbiome and cancer is intricate and is influenced by factors such as host genetics, diet, oral hygiene, and environmental factors.[9]

Oral microbiome appears to influence OSCC via modulation of cell metabolism by altering nutrient and vitamin concentrations, thus promoting the production of cytokines which are involved in different pathological conditions.[10]

However, connections between oral microbiome and oral cancer are relatively unexplored, and a bacterial spectrum in oral potentially malignant disorders and OSCC is yet to be identified. This has triggered a considerable amount of research to explore and understand the possible involvement of the oral microbiome in the aetiopathogenesis of OSCC. There is a need to identify biomarkers with the notion that every patient is unique and needs personalized treatment.

Cancer-related changes in the oral microbiome have been studied by microbial culture and molecular methods. Advanced methods like Next Generation Sequencing (NGS) allow investigation of the microbial communities to unprecedented depths.

This study was designed with an aim to profile the oral microbiome of normal, oral potentially malignant disorders (leukoplakia – Leu, oral submucous fibrosis – OSMF) and OSCC by NGS of the 16S ribosomal rRNA gene followed by bioinformatics data analysis.

METHODOLOGY

This is an observational cross-sectional study. The protocol of this study was approved by the Institutional Research Committee and Institutional Ethics Committee, Bharati Vidyapeeth Deemed to be University, Dental College and Hospital, Pune, India. (Ref – Br V/19-03-2019).

Study population

A total of 50 subjects, between 40 to 60 years of age, including both males and females were selected for this study. The study groups consisted of normal subjects (healthy controls) reporting for routine extractions, scaling, etc with no other medical history and habit history (n = 10), subjects with leukoplakia with histo-pathologically confirmed dysplasia (n = 10), subjects clinically and histo-pathologically diagnosed with OSMF (n = 10), and subjects clinically and histo-pathologically diagnosed with OSCC (n = 20).

Subject recruitment

Subjects were recruited from Bharati Vidyapeeth (Deemed to be University), Dental College and Hospital, Pune and Bharati Hospital and Research Centre, Pune. DNA extraction and spectro-photometric analysis were performed at the Department of Pathology, Bharati Hospital and Research Centre, Pune. NGS and bioinformatics data analysis were carried out at the National Centre for Microbial Resource (NCMR), National Centre for Cell Science, Pune.

Written informed consent was acquired from every subject selected for the study. A thorough clinical examination was performed, and a structured history was obtained about demographics, medical history, diet, and habits.

Sample collection

An unstimulated saliva sample collection method was used. Subjects were refrained from eating or drinking for at least 30 minutes before the sampling procedure. Subjects were requested to collect saliva in their mouth for at least a minute and drool into a labelled 50 ml collection tube (Falcon sterile conical polypropylene tube with flat-top screw cap). Aseptic techniques were followed for the collection of the specimen. All samples were transported to the laboratory in appropriate containers on ice within 2 hours of collection and stored at -80^{0}C before further processing.

DNA extraction

Saliva samples were thawed at room temperature before DNA extraction. Bacterial genomic DNA was extracted from the samples using a QIAamp DNA mini kit (Qiagen, Germany, Catalogue number – 51304) according to the manufacturer's instructions. (Protocol from Qiagen).[11]

Polymerase chain reaction

DNA was quantified using Nanodrop (Thermo Scientific, USA), and the quality of DNA was checked by gel electrophoresis. The DNA samples were subjected to amplification of the 16S rRNA gene using V4 region-specific primers, forward [5'GTGCCAGCMGCCGCGGTAA3'] and reverse [5'GGACTACHVGGGTWTCTAAT3']. PCR was performed using the following conditions: initial denaturation at 95^0 C for 5 minutes, denaturation at 95^0 for 30 seconds, annealing at 55^0 C for 30 seconds, extension at 72^0 C for 30 seconds, and final extension at 72^0 C for 5 minutes. A total of 25 cycles were performed for amplification.[12]

Library preparation

After amplification, PCR products were cleaned by AMPure XP beads (catalogue no A63882, Beckman Coulter, Inc, USA) and subjected to library preparation using Nextra XT library preparation kit (Illumina, USA), followed by limited cycle PCR.

Quality control of the final DNA library

The final clean-up was performed using AMPure XP beads, and the obtained libraries were assessed for fragment distribution and quality control using Tapestation (Agilent Technologies, USA). The size of the final library on the Tapestation was around 420 bp.

16S rRNA gene amplicon sequencing

The quantified libraries were amplified and sequenced using the Illumina MiSeq platform which generated high-quality full-length reads of the V4 region of the 16S rRNA gene in a single 48-hour run. All the steps were followed as per the protocol for the Illumina MiSeq system.[13]

Bio-informatics data analysis

After sequencing on the Illumina MiSeq platform, the raw data were obtained as paired-end FASTQ files. 16S rRNA sequencing data analysis was carried out using the Divisive Amplicon Denoising Algorithm (DADA2) pipeline (https://github.com/benjjneb/dada2),[14,15] and phyloseq R package (Version 4.0) was used for statistical analysis. SILVA reference database was used for taxonomic assignment.[15] Representative sequences were assigned at different taxonomic levels (Phylum, Family and Genus) with a 97% cut-off value. The *t*-test was used for statistical analysis. A P value of less than 0.05 was considered statistically significant.

RESULTS

The demographic characteristics of the study subjects are presented in Tables 1 and 2.

Illumina generated raw data reads were demultiplexed, and barcode adaptor sequences were removed. Raw sequences were processed based on the DADA2 workflow.[16] The low-quality sequences were filtered and trimmed, and the remaining reads were subsequently denoised and merged, and chimeras were removed.

A total number of 50,30,863 raw reads were generated from 47 saliva samples after sequencing on the Illumina MiSeq platform, and 38,02,436 high-quality sequences were retained for further downstream analysis. The total

Table 1: Age and Gender Distribution in all study groups

Study groups	Normal	Leukoplakia	OSMF	OSCC
No of Subjects	10	10	10	20
Age (Mean)	45.1	49	45.1	53.8
Gender				
Male	07	07	08	17
Female	03	03	02	03

Table 2: Type, Duration, and Frequency of Habit in all study groups

	No of subjects (Out of 40)	Percentage
Type of Habit		
Tobacco	27	67.5%
Gutkha	08	20%
Cigarette	03	7.5%
Mishri	02	5%
Duration		
<5 years	06	15%
5 to 10 years	18	45%
10 to 15 years	06	15%
15 to 30 years	10	25%
Frequency		
2-3 times a day (Low)	07	17.5%
3-4 times a day (Moderate)	19	47.5%
5 times and above (High)	14	35.0%
2-3 times a day (Low)	07	17.5%

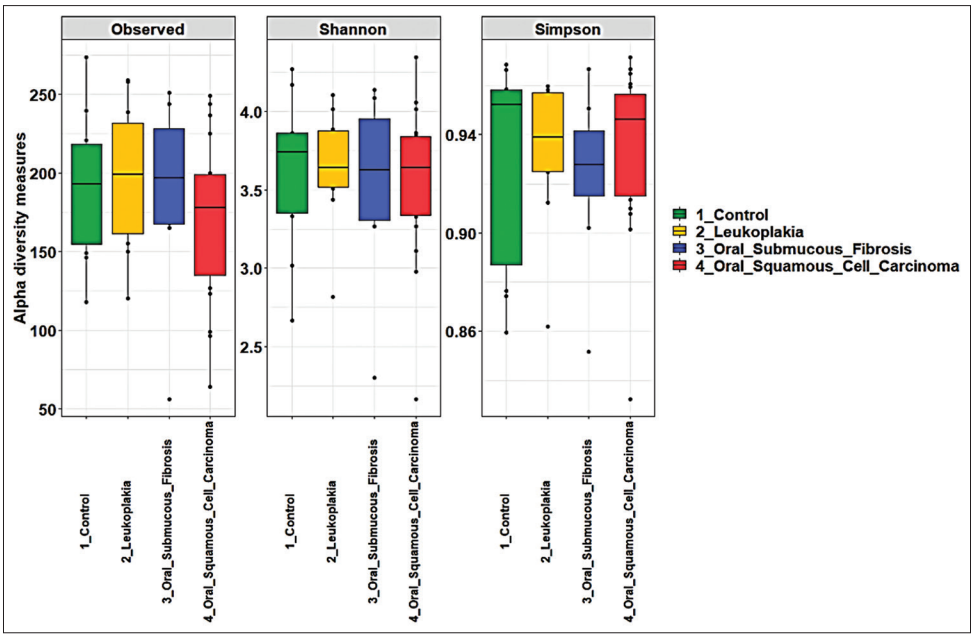

Figure 1: Alpha diversity analysis – Box-plot of Alpha diversity analysis of the oral microbiome across all four study groups (normal – control, oral potentially malignant disorders – leukoplakia, OSMF, and OSCC). The boxes denote the interquartile ranges (IQR), and the black line represents the median. Dots represent the outlier samples

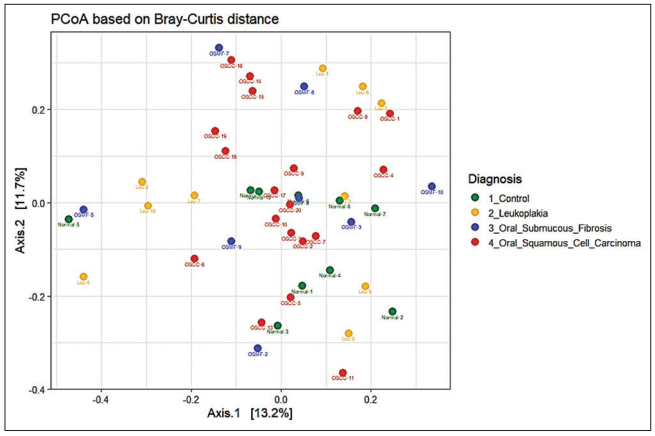

Figure 2: Beta diversity analysis of the oral microbiome across all four study groups (normal – control, oral potentially malignant disorders – leukoplakia, OSMF, and OSCC) – Principal Coordinate Analysis (PCoA) based on data matrix using the Bray-Curtis distance metrics

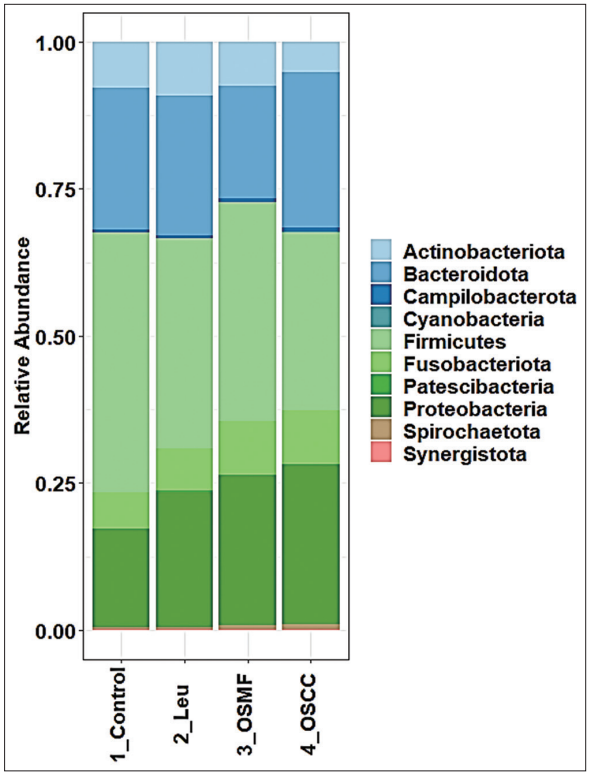

Figure 3: Stacked box plots showing the relative abundance of microbial taxa at the phylum level. Top ten phyla across all study groups – normal (control), leukoplakia (Leu), oral submucous fibrosis (OSMF), and oral squamous cell carcinoma (OSCC)

number of Amplicon Sequence Variants (ASVs) obtained from the process was 1572. Each ASV obtained from the process was compared for taxonomic assignment. The sequence statistics for bioinformatics data analysis is given in Supplementary Table S1.

Alpha diversity analysis

The alpha diversity metrics calculated were observed for ASVs using the Shannon index for species richness and evenness and the Simpson index for evenness. This study showed an increase in the observed number of ASVs in leukoplakia and OSMF and a decrease in OSCC though not significant. According to Shannon and Simpson indices, there was a decrease in alpha diversity in leukoplakia, OSMF, and OSCC compared to the normal [Figure 1].

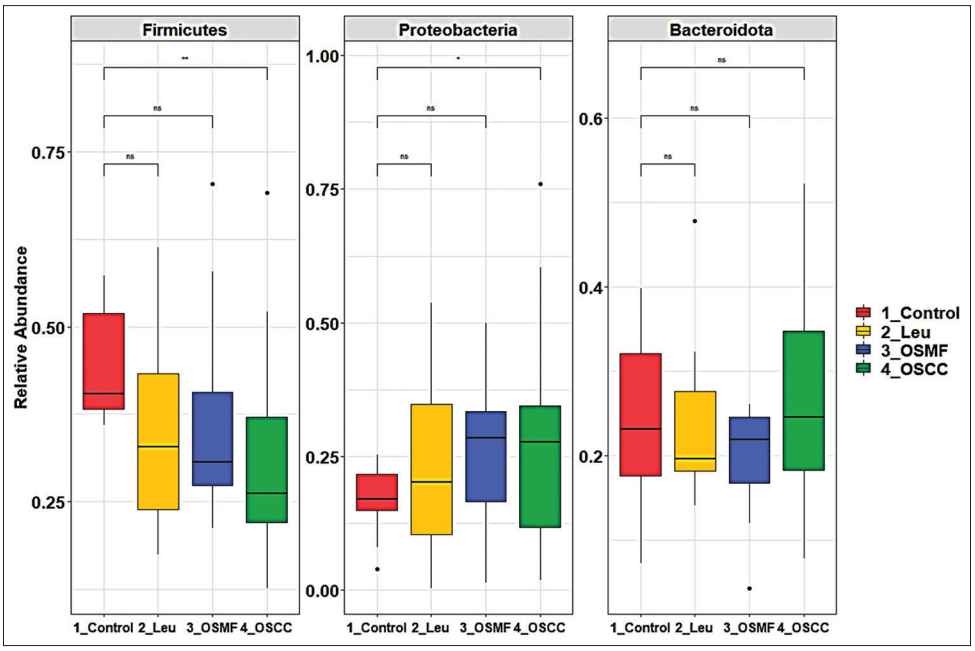

Figure 4: Comparison among major phyla across all the study groups – normal (control), leukoplakia (Leu), oral submucous fibrosis (OSMF), and oral squamous cell carcinoma (OSCC)

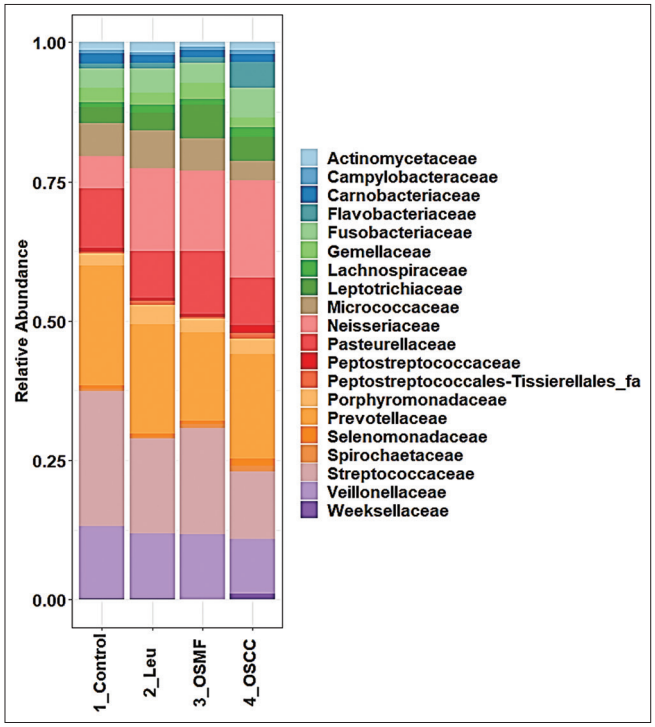

Figure 5: Stacked box plots showing relative abundance of microbial taxa at family level across all the study groups – Normal (Control), Leukoplakia (Leu), Oral Submucous Fibrosis (OSMF) and Oral Squamous Cell Carcinoma (OSCC)

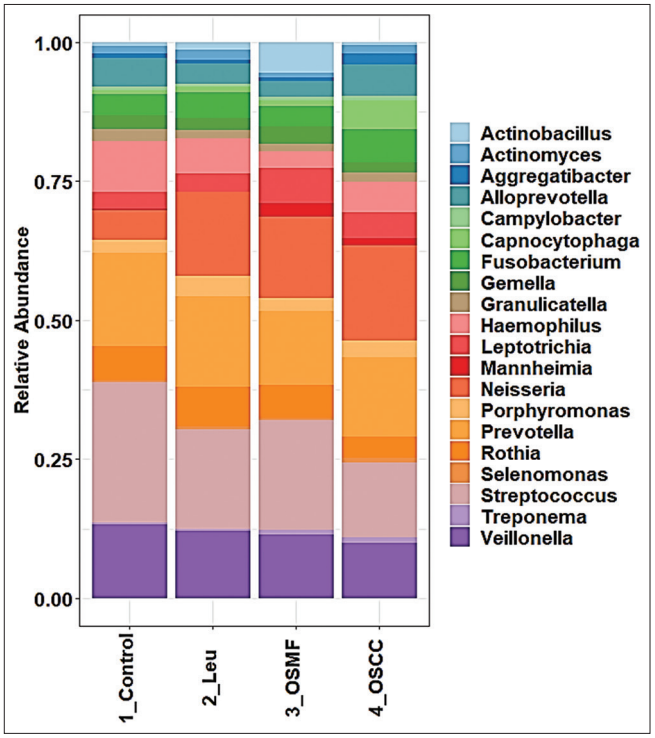

Figure 6: Stacked box plots showing relative abundance of microbial taxa at genus level. Top twenty bacterial genera across all study groups - Normal (Control), Leukoplakia (Leu), Oral Submucous Fibrosis (OSMF) and Oral Squamous Cell Carcinoma (OSCC)

Beta diversity analysis–

The principal co-ordinate analysis (PCoA) was used to evaluate the extent of the similarity amongst the bacterial communities. A dot represents each sample. Visualizing the Bray Curtis distance metrics between samples in a PCoA plot, samples that are more distant from one another are dissimilar. The inter-individual variations were found to be greater in leukoplakia, OSMF, and OSCC compared to normal. The normal group showed a close cluster of subjects compared to other disease groups which appear more separated.

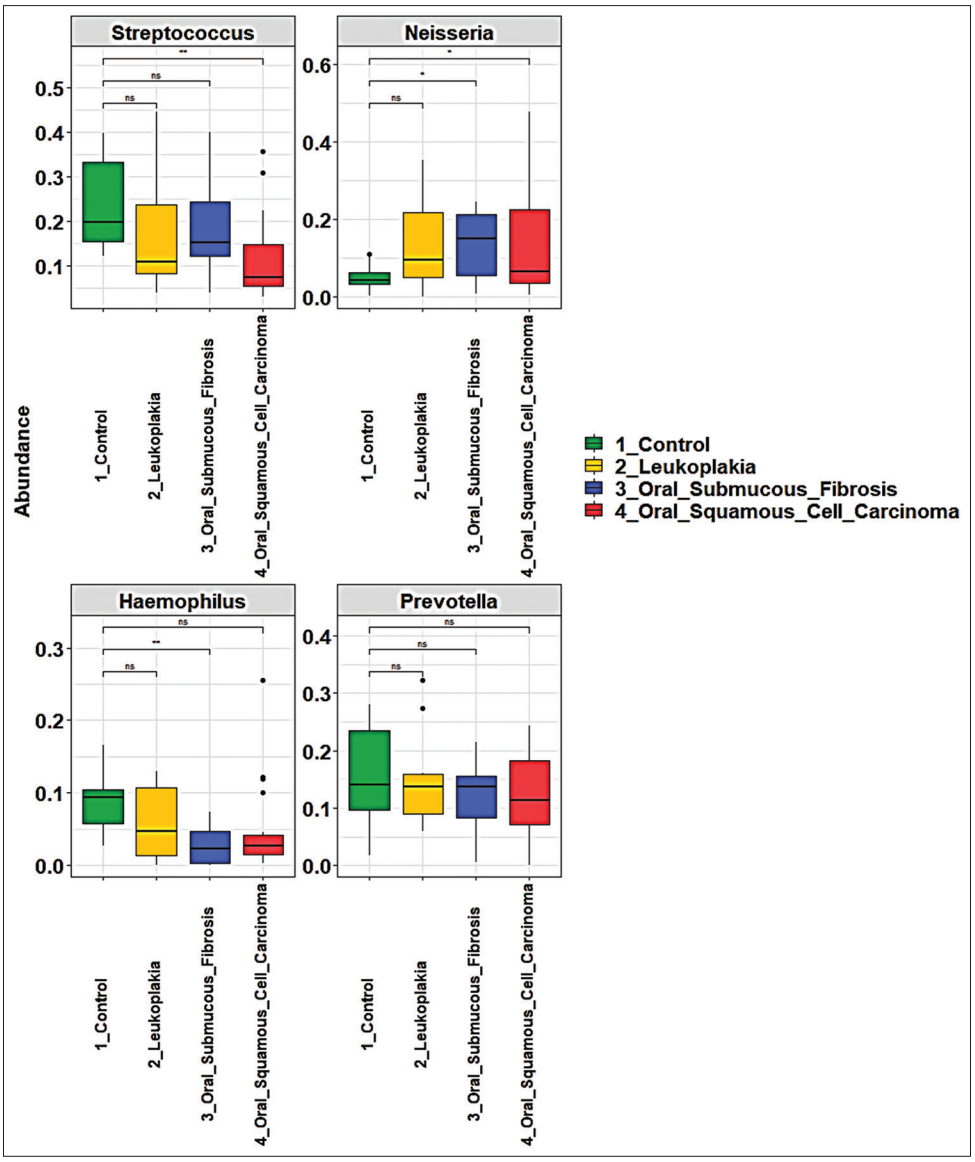

Figure 7: Comparison of differentially abundant bacterial genera across all four study groups

Therefore, leukoplakia, OSMF, and OSCC showed a more diverse microbiome in comparison to the normal [Figure 2].

Relative abundance of bacteria at phylum, family, and genus levels

The identified ASVs were classified into 21 phyla, 100 families, and 197 genera.

At the phylum level, *Synergistota, Spirochaetota, Proteobacteria, Patescibacteria, Fusobacteriota, Firmicutes, Cyanobacteria, Campilobacterota, Bacteroidota,* and *Actinobacteriota* [Figure 3] were detected. *Firmicutes* (43.94%), *Bacteroidota* (23.70%), and *Proteobacteria* (16.87%) were identified as the most abundant phyla. *Firmicutes* showed a decreasing trend in leukoplakia and OSMF and a significant decrease in OSCC (32.17%) compared to normal (43.94%), *P* value < 0.01. *Bacteroidota* also showed a decreasing trend in leukoplakia and OSMF compared to the normal. However, the relative abundance of *Proteobacteria* was found to be increased in leukoplakia (21.15%), OSMF (25.18%), and OSCC (26.24%) in comparison to the normal (16.87%). The difference between normal and OSCC was found to be statistically significant, *P* value < 0.05 [Figure 4].

The taxa such as *Streptococcaceae, Neisseriaceae, Veillonellaceae,* and *Prevotellaceae* were the most abundantly detected families [Figure 5]. Family *Neisseriaceae* showed an increase in the relative abundance in the leukoplakia, OSMF, and OSCC compared to the normal.

Twenty differentially abundant genera were identified across all study groups [Figure 6]. Among them, the top ten genera found were *Streptococcus, Veillonella, Prevotella, Neisseria, Rothia,*

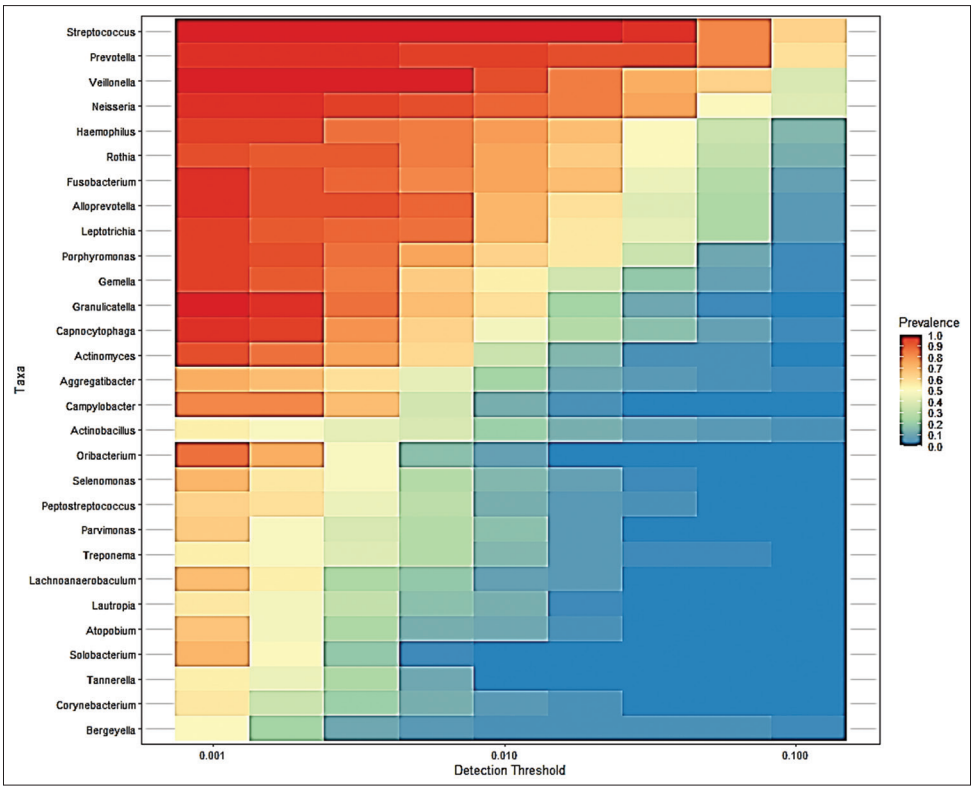

Figure 8: Heat map of core microbiome across all study groups (Cell colour indicates the relative prevalence of taxa at a specific detection threshold ranging from 0.0 to 1.0)

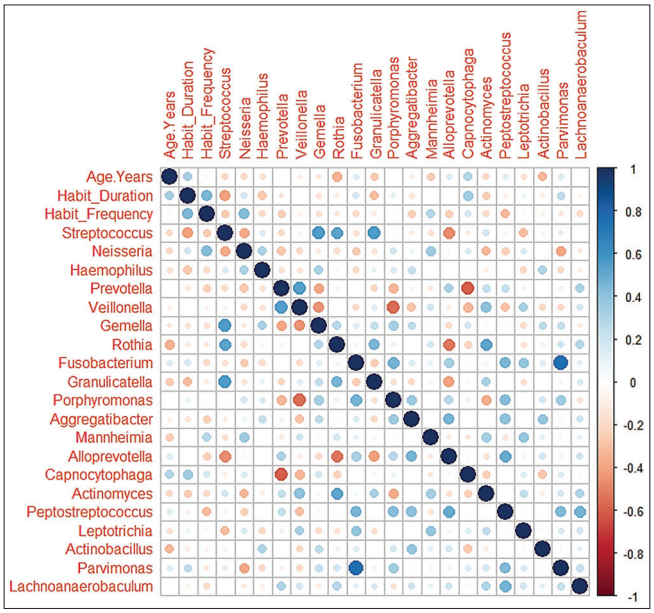

Figure 9: Spearman correlation analysis among the top abundant genera across all study groups. The Spearman correlation values ranges from -1 (red) to +1 (blue) for relative abundance

Fusobacteria, Leptotrichia, Gemella, Alloprevotella, and Actinobacillus. Streptococcus genus showed a statistically significant decrease in the OSCC (11.71%) compared to the normal (23.72%), *P* value < 0.01. The most abundantly found genera in OSCC was *Neisseria* (15.05%), followed by *Prevotella* (12.46%), *Streptococcus* (11.71%), and *Veillonella* (8.73%). The genus *Neisseria* showed a statistically significant increase in OSCC in comparison to the normal, *P* value < 0.05. There was a significant difference in the genus *Haemophilus* between the normal (8.65%) and OSMF (2.82%), *P* value < 0.01 [Figure 7]. Variations found in abundantly detected genera are shown in Supplementary Figures S1 and S2. There was an increase

in the relative abundance of bacterial genera belonging to the red complex responsible for periodontal disease namely *Porphyromonas, Treponema,* and *Capnocytophaga*. Genus *Porphyromonas* showed an increase in the relative abundance of leukoplakia (3.41%) and OSCC (2.66%) compared to the normal (2.13%). There was an increase in the relative abundance of *Treponema* in OSMF (0.67%) and OSCC (0.88%) as compared to the normal (0.46%) and an increase in the relative abundance in the genus *Capnocytophaga* in leukoplakia (0.95%), OSMF (1.01%), and OSCC (4.45%) compared to the normal (0.73%).

The core oral microbiome detected across all study groups

Bacterial genera prevalent in more than 50% of the study population with more than 0.1% abundance (detection threshold) were considered as a part of the core microbiome. Among 197 genera detected, 29 genera were selected based on the fact that they were present in at least 50% of subjects [Figure 8]. Six genera were found in high abundance across all study groups, namely, *Streptococcus, Prevotella, Veillonella, Neisseria, Haemophilus,* and *Rothia*.

In this study, around 40 bacterial genera were found unique to OSCC and not detected in other groups [Supplementary Table S2].

Spearman correlation

Spearman correlation analysis estimates the microbial interaction between pairs of taxa.[17] Spearman correlation among top abundant genera for the analysis of the oral microbiome was performed including the demographic characteristics. The colour intensity and the size of the circle show the strength of the correlation. The Spearman correlation coefficient ranges from -1 (red) to +1 (blue) [Figure 9]. Genus *Parvimonas* was positively correlated with *Fusobacterium* (0.8), *Porphyromonas* (0.4), and *Peptostreptococcus* (0.6); however, it was negatively correlated with *Neisseria* (-0.6). The relative abundance of *Neisseria* showed a positive correlation with an increase in habit frequency (0.6). When the habit duration is decreased, the relative abundance of *Streptococcus* was decreased (-0.6) [Figure 9].

DISCUSSION

Microbiota plays a major role in the pathophysiology of many diseases. Bacteria are known to be linked with the aetiopathogenesis of a number of diseases including cancer. The World Health Organization (WHO) has categorised *Helicobacter pylori* as a Class I carcinogen.[18] However, the literature on oral microbiome as markers for oral cancer is still lacking and a significant consensus has not been reached. Oral cancer treatment modalities have shown considerable improvements, but the mortality rate of OSCC is still around 43% and the 5-year survival rate is 56%. This is due to the fact that majority of the oral cancers are diagnosed at late stages. The late diagnosis may be linked to a lack of consistent oral cancer prevention programs, the inability to detect early stages of oral cancer in asymptomatic subjects, and the lack of consistent biological markers for diagnostic purposes.[19]

This study profiled and compared the oral microbiome of normal, leukoplakia, OSMF, and OSCC. The V4 region of the 16S rRNA gene was selected for sequencing and analysis. The results of the diversity analysis (alpha and beta diversity) of the present study are similar to the previous reports by Pushalkar *et al*.,[20] Guerrero-Preston *et al*.,[21] Shin *et al*.,[22] Panda *et al*.,[23] Lim *et al*.,[24] and Perera *et al*.[25] A study by Wang *et al.* in 2017[26] showed minor variations in alpha diversity in the OSCC tissue and non-tumour tissues. A report by Li *et al.* in 2021[27] showed that β diversity of oral cancer samples showed marked differences from the normal and precancerous lesions. Their analysis revealed that samples of normal and precancerous lesions were mixed; however, oral cancer samples were more specific and different. The existing literature shows considerable differences in the diversity and richness between healthy and tumour samples. These variations could be due to the type of sample collected as there are complex niches in the oral cavity that harbour different micro-flora.

Phyla *Firmicutes, Proteobacteria,* and *Bacteroidota* were the most predominant phyla found in this study. Relative abundance of *Firmicutes* was decreased in leukoplakia (36.84%) and OSMF (37.51%) and significantly decreased in OSCC (32.17%) in comparison to the normal (43.94%). These results are in accordance with Banerjee *et al*., 2017,[28] Zhao *et al*., 2017,[6] Amer *et al*., 2020,[29] Gopinath *et al*., 2020,[30] and Schmidt *et al*., 2014.[31] However, according to reports by Zhang *et al.* in 2019[32] and Pushalkar *et al*., 2012,[33] *Firmicutes* was the dominant phylum in OSCC samples. *Proteobacteria* was increased in oral potentially malignant disorders (leukoplakia and OSMF) and OSCC compared to normal. The difference between the normal (16.87%) and OSCC (26.24%) was statistically significant. This is in accordance with the study conducted by Banerjee *et al.* in 2017.[28] *Proteobacteria* is a major Gram-negative phylum that includes *Escherichia, Salmonella,* and *Vibrio*. *Bacteroidota* is a major Gram-negative phylum which includes *Prevotella* and *Capnocytophaga*. *Capnocytophaga* was increased in OSCC (25.33%) relative to the normal (23.70%). This suggests that there is an increase in the relative abundance of Gram-negative bacteria in oral pre-cancer and oral cancer.

The most abundant genera identified in our study were *Streptococcus, Veillonella, Prevotella, Neisseria, Rothia, Fusobacteria, Leptotrichia, Gemella, Alloprevotella,* and *A*ctinobacillus. The genus Streptococcus belongs to Phylum *Firmicutes* and is the most predominant microorganism in a healthy oral cavity. This study showed a significant decrease in the relative abundance of *Streptococcus* in OSCC (11.71%) compared to the normal (23.72%). This result is in accordance with studies conducted by Su *et al.* in 2021[34] and Schmidt *et al.* in 2014.[31] *Neisseria* showed an increase in leukoplakia and OSMF and a significant increase in OSCC, *P* value < 0.05 compared to normal. A study by Zhang *et al.* in 2020[35] concluded that *Neisseria* was stable in both OSCC and control accounting for 20% in both groups. According to an article by Muto *et al.* in 2000,[36] *Neisseria* could be a source of carcinogenic acetaldehyde and therefore may play a potential role in alcohol-related carcinogenesis.

Acetaldehyde is a proximal metabolite, classified as a Group 1 carcinogen by the WHO's International Agency for Research on Cancer. Elevated levels of acetaldehyde in whole mouth fluid have been linked with increased risk for cancers of the upper aero-digestive tract.[37]

This study showed a statistically significant difference in the genus *Hemophilus* in normal and OSMF. The genus *Capnocytophaga* was also found to be increased in OSCC(4.45%) compared to the normal (0.73%) [Figure 6]. According to a report by Mager *et al*.,[38] the level of *Capnocytophaga gingivalis* was significantly elevated in OSCC. This genus is a putative periodontal pathogen and was abundantly seen associated with recurrent cases of OSCC.[39] Our study showed an increase in the relative abundance of periodontal pathogens, namely, *Porphyromonas* and *Treponema* in OSCC. Many reports have shown the relationship between periodontal bacteria and OSCC.[40] Periodontal disease is considered as a risk factor for oral cancer. Inflammation, a hallmark of cancer, plays a major role in periodontitis and mediates carcinogenesis. Lipopolysaccharide synthesis, bacterial chemotaxis, and flagella assembly by bacteria are associated with inflammation in oral cancer.[35] Lipopolysaccharide endotoxin binds to the TLR receptor of leucocytes and stimulates the production of cytokines such as IL1-β, IL-6, and TNF–α. The upregulation of these cytokines alters different molecular and metabolic pathways which modulate cell metabolism and proliferation.[10]

In this study, around 40 bacterial genera were identified in relatively low abundance and unique to OSCC and not found in normal and oral potentially malignant disorders. Further research regarding the potential role of these bacteria in low abundance should be considered.

The six genera found in most abundance were *Streptococcus, Prevotella, Veillonella, Neisseria, Hemophilus,* and *Rothia*. *Streptococcus* is a facultative anaerobe, and *Prevotella* is an obligate anaerobic bacteria. *Neisseria* is aerobic and a primary colonizer of the oral cavity. The greater abundance of aerobic, facultative, and obligate anaerobes suggests that oxygen sensitivity plays a role in structuring the bacterial diversity of the oral cavity. The diversity of the core microbiome may provide a snapshot of the homeostasis, and deviations from this core may be linked to diseases.[41]

This study interpreted variations in the composition of the oral microbiome at phylum, family, and genus levels across all study groups and generated a baseline data. Current literature shows strong evidence suggesting the association between oral microbiome and oral cancer.

The present study may set a foundation for future functional microbiome research focussing on host–microbiome interactions. Further studies are required to confirm the association between oral microbiome and oral cancer. Microbial signatures may serve as biomarkers for early screening, assessing the risk of disease, diagnostics, and treatment planning.

CONCLUSION

This study concludes that there are variations in the composition of the oral microbiome in oral pre-cancer and cancer. Oral microbiome diversity was found to be reduced in pre-cancer and cancer, while inter-individual variations were increased. These variations in the composition of the oral microbiome may provide signatures during the progression from normal to pre-cancer to cancer, which may be used as biomarkers. Bacterial taxa unique to OSCC may be the keystone pathogens and, therefore, should not be ignored. To understand which particular taxa may contribute to tumorigenesis, an integrative omics approach is essential. Longitudinal studies focussing on the assessment of the microbiome before and after treatment should be considered. The sample size which includes different ethnicities, races, and geographical locations should be preferred so that results may be generalised. The findings of this study may provide new directives for further research to study the linkages between oral microbiome and oral pre-cancer and cancer.

Data availability

The raw data generated is available on NCBI- Sequence Read Archive (SRA) under the Bioproject ID: PRJNA1058875

Acknowledgement

Priya Nimish Deo extends her sincere gratitude and would like to acknowledge the valuable guidance and support of Dr. Yogesh Shouche (Honorary Scientist, National Centre for Cell Science, Pune) without which this study would not have been possible.

The authors would like to acknowledge the support of Bharati Vidyapeeth Deemed to be University, Dental College and Hospital, Pune, Bharati Hospital and Research Centre, Pune and the National Centre for Microbial Resource, National Centre for Cell Science, Pune.

Priya Nimish Deo would like express sincere gratitude to Dr. Ramesh Paranjape, In-charge Director and Distinguished Professor, Dr. PKBSRC, KLE Academy of Higher Education and Research, Belagavi who provided valuable inputs during the early phase of the study.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Kanavos P. The rising burden of cancer in the developing world. Ann Oncol 2006;17(Suppl 8):viii15-23.
2. Srivastava R, Wazir SS, Jyoti B, Kushwah S, Pradhan D, Priyadarshi P. Perception and outcome of oral cancer awareness among clinical undergraduate dental students of Tertiary health care centre at Kanpur city: A cross-sectional study. Natl J Maxillofac Surg 2020;11:89-93.
3. Gokul S, Sapna G. Immunosurveillance and oral cancer. J Tumor 2014;2:136-41.
4. Manchanda K. Understanding oral cancer at molecular level: The new frontier. Ann Trop Med Public Health 2015;8:88-93.
5. Benjamin WJ, Wang K, Zarins K, Bellile E, Blostein F, Argirion I, *et al*. Oral microbiome community composition in head and neck squamous cell carcinoma. Cancers (Basel) 2023;15:2549.
6. Zhao H, Chu M, Huang Z, Yang X, Ran S, Hu B, *et al*. Variations in oral microbiota associated with oral cancer. Sci Rep 2017;7:11773.
7. Jia G, Zhi A, Lai PFH, Wang G, Xia Y, Xiong Z, *et al*. The oral microbiota-A mechanistic role for systemic diseases. Br Dent J 2018;224:447-55.
8. Perera M, Al-Hebshi NN, Speicher DJ, Perera I, Johnson NW. Emerging role of bacteria in oral carcinogenesis: A review with special reference to perio-pathogenic bacteria. J Oral Microbiol 2016;8:32762.
9. Smędra A, Berent J. The influence of the oral microbiome on oral cancer: A literature review and a new approach. Biomolecules 2023;13:815.
10. La Rosa GRM, Gattuso G, Pedullà E, Rapisarda E, Nicolosi D, Salmeri M. Association of oral dysbiosis with oral cancer development. Oncol Lett 2020;19:3045-58.
11. Available from: https://www.qiagen.com/us/products/discovery-and-translational research/dna-rna-purification/dna-purification/microbial-dna/qiaamp-dna-Microbiome-kit/. [Last accessed on 2019 Oct 04].
12. Gaike AH, Paul D, Bhute S, Dhotre DP, Pande P, Upadhyaya S, *et al*. The gut microbial diversity of newly diagnosed diabetics but not of prediabetics is significantly different from that of healthy nondiabetics. mSystems 2020;5:e00578-19.
13. Available from: https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf. [Last accessed on 2021 Apr 28].
14. Available from: https://github.com/benjjneb/dada2. [Las accessed on 2021 Apr 28].
15. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High resolution sample inference from Illumina amplicon data. Nat. Methods 2016;13:581-3.
16. Available from: https://www.arb silva.de/documentation/release 138/. [Last accessed on 2021 Apr 28].
17. Matchado MS, Lauber M, Reitmeier S, Kacprowski T, Baumbach J, Haller D, *et al*. Network analysis methods for studying microbial communities: A mini review. Comput Struct Biotechnol J 2021;19:2687-98.
18. Khatoon J, Rai RP, Prasad KN. Role of *Helicobacter pylori* in gastric cancer: Updates. World J Gastrointest Oncol 2016;8:147-58.
19. Torralba MG, Aleti G, Li W, Moncera KJ, Lin YH, Yu Y, *et al*. Oral microbial species and virulence factors associated with oral squamous cell carcinoma [published correction appears in Microb Ecol. 2020 Nov 20]. Microb Ecol 2021;82:1030-46.
20. Pushalkar S, Mane S, Ji X, Li Y, Evans C, Crasta OR, *et al*. Microbial diversity in saliva of oral squamous cell carcinoma. FEMS Immunol Med Microbiol 2011;61:269-77.
21. Guerrero-Preston R, Godoy-Vitorino F, Jedlicka A, Rodríguez-Hilario A, González H, Bondy J, *et al*. 16S rRNA amplicon sequencing identifies microbiota associated with oral cancer, human papilloma virus infection and surgical treatment. Oncotarget 2016;7:51320-34.
22. Shin JM, Luo T, Kamarajan P, Fenno JC, Rickard AH, Kapila YL. Microbial communities associated with primary and metastatic head and neck squamous cell carcinoma—A high fusobacterial and low streptococcal signature. Sci Rep 2017;7:9934.
23. Panda M, Rai AK, Rahman T, Das A, Das R, Sarma A, *et al*. Alterations of salivary microbial community associated with oropharyngeal and hypopharyngeal squamous cell carcinoma patients. Arch Microbiol 2019;202:785-805.
24. Lim Y, Fukuma N, Totsika M, Kenny L, Morrison M, Punyadeera C. The performance of an oral microbiome biomarker panel in predicting oral cavity and oropharyngeal cancers. Front Cell Infect Microbiol 2018;8:267.
25. Perera M, Al-Hebshi N, Perera I, Ipe D, Ulett G, Speicher D, *et al*. Inflammatory bacteriome and oral squamous cell carcinoma. J Dent Res 2018;97:725-32.
26. Wang H, Funchain P, Bebek G, Altemus J, Zhang H, Niazi F, *et al*. Microbiomic differences in tumor and paired-normal tissue in head and neck squamous cell carcinomas. Genome Med 2017;9:14.
27. Li Z, Chen G, Wang P, Sun M, Zhao J, Li A, *et al*. Alterations of the oral microbiota profiles in Chinese patient with oral cancer. Front Cell Infect Microbiol 2021;11:780067.
28. Banerjee S, Tian T, Wei Z, Peck KN, Shih N, Chalian AA, *et al*. Microbial signatures associated with oropharyngeal and oral squamous cell carcinomas. Sci Rep 2017;7:4036.
29. Amer A, Whelan A, Al-Hebshi N, Healy CM, Moran GP. Acetaldehyde production by Rothiamucilaginosa isolates from patients with oral leukoplakia. J Oral Microbiol 2020;12:1743066.
30. Gopinath D, Kunnath Menon R, Chun Wie C, Banerjee M, Panda S, Mandal D, *et al*. Salivary bacterial shifts in oral leukoplakia resemble the dysbiotic oral cancer bacteriome. J Oral Microbiol 2020;13:1857998.
31. Schmidt BL, Kuczynski J, Bhattacharya A, Huey B, Corby PM, Queiroz EL, *et al*. Changes in abundance of oral microbiota associated with oral cancer. PLoS One 2014;9:e98741. Erratum in: PLoS One 2014;9:e106297.
32. Zhang Z, Yang J, Feng Q, Chen B, Li M, Liang C, *et al*. Compositional and functional analysis of the microbiome in tissue and saliva of oral squamous cell carcinoma. Front Microbiol 2019;10:1439.

33. Pushalkar S, Ji X, Li Y, Estilo C, Yegnanarayana R, Singh B, *et al*. Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. BMC Microbiol 2012;12:144.
34. Su S-C, Chang L-C, Huang H-D, Peng C-Y, Chuang C-Y, Chen Y-T, *et al*. Oral microbial dysbiosis and its performance in predicting oral cancer. Carcinogenesis 2021;42:127-35.
35. Zhang L, Liu Y, Zheng HJ, Zhang CP. The oral microbiota may have influence on oral cancer. Front Cell Infect Microbiol 2020;9:476.
36. Muto M, Hitomi Y, Ohtsu A, Shimada H, Kashiwase Y, Sasaki H, *et al*. Acetaldehyde production by non-pathogenic Neisseria in human oral microflora: Implications for carcinogenesis in upper aerodigestive tract. Int J Cancer 2000;88:342-50.
37. Gopinath D, Menon RK, Wie CC, Banerjee M, Panda S, Mandal D, *et al*. Differences in the bacteriome of swab, saliva, and tissue biopsies in oral cancer. Sci Rep 2021;11:1181.
38. Mager DL, Haffajee AD, Devlin PM, Norris CM, Posner MR, Goodson JM. The salivary microbiota as a diagnostic indicator of oral cancer: A descriptive, non-randomized study of cancer-free and oral squamous cell carcinoma subjects. J Transl Med 2005;3:27.
39. Yao Y, Shen X, Zhou M, Tang B. Periodontal pathogens promote oral squamous cell carcinoma by regulating ATR and NLRP3 inflammasome. Front Oncol 2021;11:722797.
40. Ganly I, Yang L, Giese RA, Hao Y, Nossa CW, Morris LGT, *et al*. Periodontal pathogens are a risk factor of oral cavity squamous cell carcinoma, independent of tobacco and alcohol and human papillomavirus. Int J Cancer 2019;145:775-84.
41. Chaudhari DS, Dhotre DP, Agarwal DM, Gaike AH, Bhalerao D, Jadhav P, *et al*. Gut, oral and skin microbiome of Indian patrilineal families reveal perceptible association with age. Sci Rep 2020;10:5685.

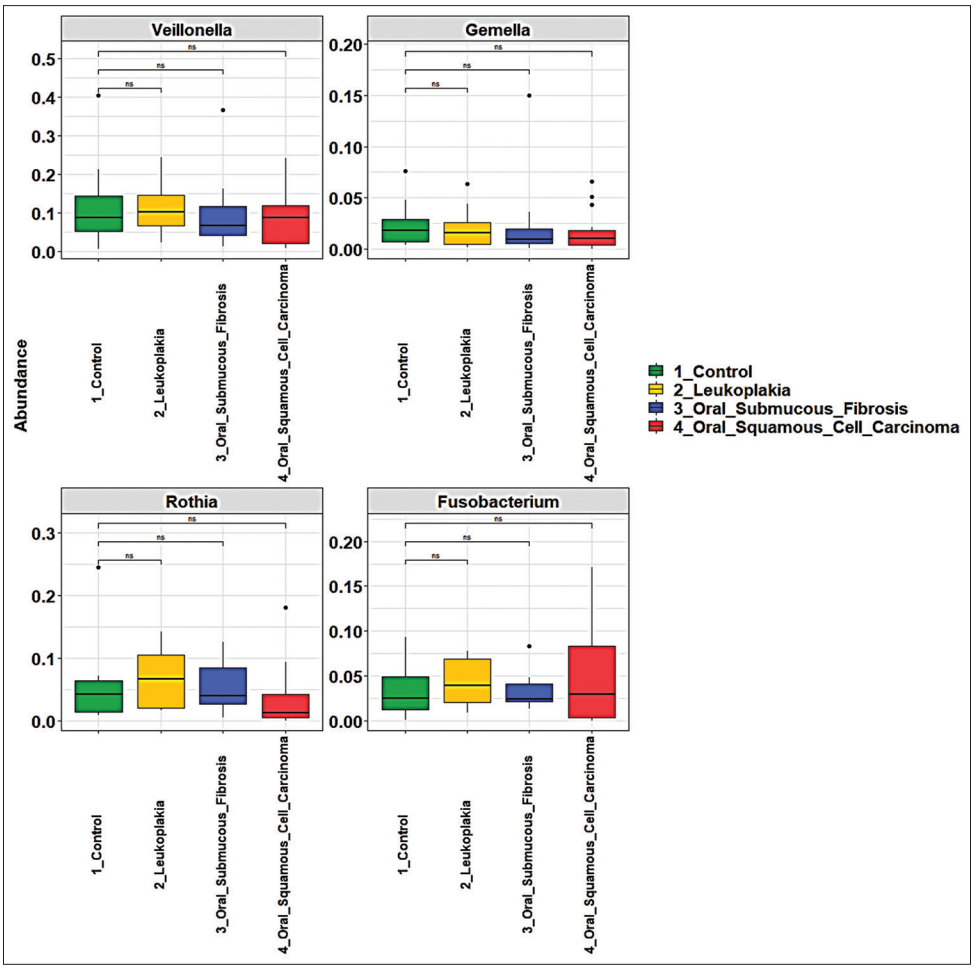

Figure S1: Comparison of Veillonella, Gemella, Rothia and Fusobacterium genera across all four study groups

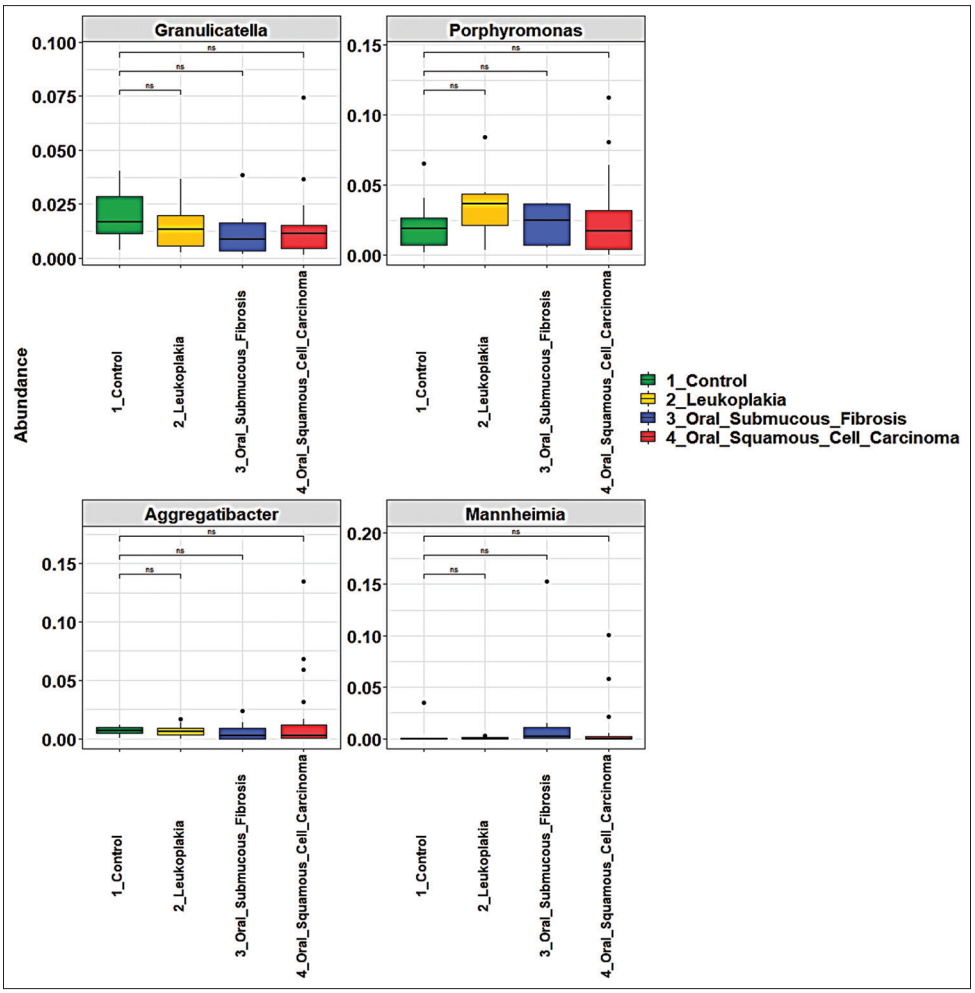

Figure S2: Comparison of Granulicatella, Porphyromonas, Aggregatibacter and Mannheimia genera across all four study groups

Table S1: Sequence statistics for Bio-informatics data analysis

Sample Id	Input reads	Filtered reads	Merged reads	Non-chimera reads
Normal-1	95618	82872	79380	75355
Normal-2	95050	81853	76607	70638
Normal-3	99342	85432	80344	70833
Normal-4	96460	85290	80279	77805
Normal-5	96955	86088	82762	68602
Normal-6	111252	98238	90743	79015
Normal-7	114621	100821	93061	83801
Normal-8	108768	96766	90083	81998
Normal-9	102579	92064	85350	66911
Normal-10	92720	83109	76996	63021
Leu-1	96162	86836	80416	67305
Leu-2	98409	89927	82702	64970
Leu-3	105786	91550	82588	71910
Leu-4	99721	86271	82639	71799
Leu-5	97421	88451	82802	79000
Leu-6	94237	85986	79889	68882
Leu-7	101737	91234	87395	80996
Leu-8	111384	95852	88233	80620
Leu-9	238972	211517	195022	181025
Leu-10	116427	106254	102450	96011
OSMF-2	107212	93593	87223	74705
OSMF-3	95489	84354	77181	68657
OSMF-4	95268	85085	78271	70663
OSMF-5	103061	92510	90518	88101
OSMF-7	89282	82418	79701	72971
OSMF-8	88324	79919	73753	69781
OSMF-9	98971	89957	86522	84052
OSMF-10	95849	84412	80211	76279
OSCC-1	91995	85274	82698	64385
OSCC-2	95961	88079	81123	69634
OSCC-3	94294	86865	79009	66103
OSCC-4	89440	80784	72549	62195
OSCC-5	172791	152478	144937	124676
OSCC-6	96309	85830	81941	76611
OSCC-7	204196	181075	167458	153040
OSCC-8	83356	75342	68631	62537
OSCC-9	90496	81064	75059	72792
OSCC-10	93033	85512	77743	61776
OSCC-11	92064	80953	73990	62852
OSCC-13	101551	86096	83332	82248
OSCC-14	92924	84336	80649	77230
OSCC-15	93842	85382	81447	79642
OSCC-16	88993	81527	78573	77792
OSCC-17	93306	85173	81184	78753
OSCC-18	183654	165815	158352	143760
OSCC-19	132970	119675	114896	105167
OSCC-20	92611	84045	79739	75537
	5030863	4493964	4216431	3802436

Total input reads – 5030863. Filtered reads – 4493964. Merged paired reads – 4216431. Non-chimeric reads – 3802436

Table S2: Bacterial genera detected only in Oral Squamous Cell Carcinoma samples

Name of the taxa	Relative abundance
Ezakiella	0.049425
Anaerococcus	0.055324
Escherichia/Shigella	0.036831
Finegoldia	0.013553
Pantoea	0.013943
Family_XIII_AD3011_group	0.011344
Bilophila	0.008274
Morganella	0.004425
Odoribacter	0.003183
Succiniclasticum	0.007171
Lentibacillus	0.002796
Salinicoccus	0.002471
Sutterella	0.002341
Brucella	0.002496
Providencia	0.001581
Enterococcus	0.000635
Incertae_Sedis	0.001154
UCG-005	0.00078
Trueperella	0.000793
Faecalibacterium	0.00065
Sphingobacterium	0.000668
Acholeplasma	0.000721
DNF00809	0.000353
Erysipelatoclostridium	0.000577
Oscillibacter	0.000666
CAG-352	0.000538
Clostridium_sensu_stricto_1	0.000445
Chryseobacterium	0.000334
Phascolarctobacterium	0.000193
Dermabacter	0.000361
Lachnospiraceae_NK4A136_group	0.00031
Candidatus_Methanomethylophilus	0.000344
Schlesneria	0.000201
Paenibacillus	0.00021
Agathobacter	0.000216
Fournierella	0.00025
Halomonas	0.00013
Acetobacter	0.00013
Paracoccus	0.000134
Trichomonas	0.00014