

Metagenomic Analysis of the Oral Microbiome in Patients with Head and Neck Squamous Cell Carcinoma: A Case-Control Study

Niyati V. Thakkar^{1,*}, Anisha Chaudhary², Smiral Gameti³, K. Ezhil Vendhan⁴, Nittin Sharma⁵, K.V. Jamuna⁶ and Rajesh Kumar Lenka⁷

¹Faculty of Allied and Healthcare, Gokul Global University, Sidhpur, Gujarat, India

²Quantum University Research Center, Quantum University, India

³Faculty of Nursing, Parul University, Vadodara, Gujarat, India

⁴Department of Ophthalmology, Vinayaka Mission's Kirupananda Variyar Medical College & Hospitals, Vinayaka Mission's Research Foundation (DU), Salem, Tamil Nadu, India

⁵Center for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, 140401, Punjab, India

⁶Department of Forensic Science, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India

⁷Department of Microbiology, IMS and SUM Hospital, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India

Abstract: Head and neck squamous cell carcinoma (HNSCC) is one of the important cancers that are responsible for considerable morbidity all over the world. Recent studies reveal that changes in the oral microbiome are linked with HNSCC, but the exact microbial signatures according to the presence of disease need further research. In this study, metagenomic analysis was done on the oral microbiome of 210 HNSCC patients and 180 healthy controls based on the whole-genome shotgun metagenomic sequencing technique. Microbial diversity analysis, taxonomic profile, functional pathway analysis, and classification using a machine learning algorithm were performed for the identification of HNSCC-specific microbial signatures. HNSCC patients showed considerable changes in oral microbial profile as compared to the healthy subjects. The Shannon diversity index was significantly lower in patients with HNSCC than in control subjects (3.09 ± 0.34 vs 3.82 ± 0.41). An increase in the relative abundance of *Fusobacterium*, *Prevotella*, and *Porphyromonas* was observed in patients with HNSCC, while the relative abundance of the commensal species *Streptococcus* and *Rothia* was found to be decreased. The functional annotation revealed enrichment of pathways related to inflammation in the HNSCC samples. Using a random forest approach based on the microbial abundance profile, discrimination was possible between the two groups with a sensitivity of 82%, specificity of 79%, and AUC of 0.86 through cross-validation.

Keywords: Head and Neck Squamous Cell Carcinoma (HNSCC), Oral Microbiome, Metagenomic Sequencing, Microbial Dysbiosis, Cancer Biomarkers, Oral Carcinogenesis, Microbial Biomarkers.

1. INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is one of the most common malignancies of the upper aerodigestive tract. It occurs as a result of the epithelial lining of the anatomic structures comprising the oral cavity, oropharynx, hypopharynx, and larynx. The disease is a notable contributor to cancer development worldwide, especially in populations with higher risk exposure, such as tobacco and alcohol use and infection with lytic viruses, such as human papillomavirus (HPV). The development of HNSCC is a slow process. It is preceded by genetic and epigenetic changes that drive malignant transformation of epithelial cells, leading to uncontrolled cell proliferation, subsequent metastasis, and tissue invasion.

Along with the known risk factors, including environmental and viral risk factors, the latest studies have also pointed to the local oral microbial ecosystem as a potential contributing factor associated with the microenvironmental changes occurring in oral carcinogenesis [1, 3, 16]. The latest advancements in high-throughput sequencing have provided new opportunities for analysis of the microbial communities in various biological settings. The use of metagenomic methods like 16S rRNA gene profiling has become widespread in studies of microbial diversity and interactions in various systems [1]. Such methods have also allowed a thorough analysis of oral microbial dysbiosis in the case of oral squamous cell carcinoma and HNSCC [2]. Metagenomic studies have shown that cancer tissues exhibit dysbiosis of the microbiota, where there is an increase in certain anaerobic bacteria and a decrease in health-promoting microorganisms that were once believed to be associated with such tissues.

*Address correspondence to this author at the Faculty of Allied and Healthcare, Gokul Global University, Sidhpur, Gujarat, India;
E-mail: nvthakkar.gpc@gokuluniversity.ac.in

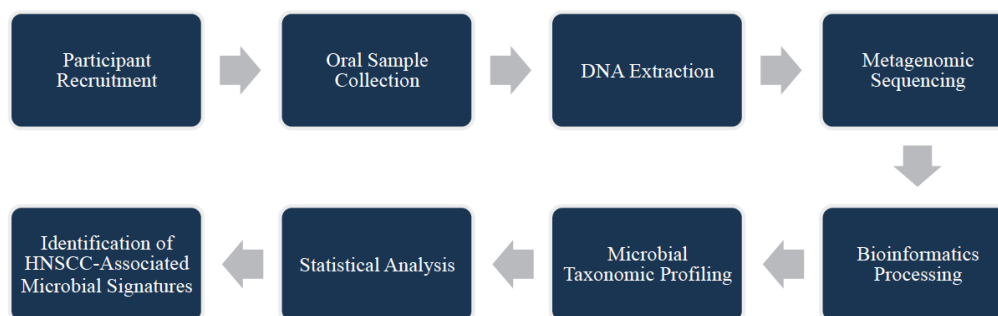


Figure 1: Overall Workflow for Metagenomic Analysis of the Oral Microbiome in HNSCC.

The methods used to establish a link between the oral microbiome and HNSCC have been summarized in Figure 1. The human mouth hosts a complicated and variable ecosystem of hundreds of bacterial species living in conjunction with the epithelial and immune system cells of the host. In a healthy state, the microbial community maintains an ecological balance and plays an important role in physiological activities, such as immune modulation and protection from any potentially pathogenic bacteria. However, the development of microbial dysbalance can promote the proliferation of pathogens that might cause chronic inflammation and alter metabolic conditions in the mouth [1, 3, 16]. Chronic inflammation triggered by microbe-related dysbiosis may contribute to carcinogenesis through generation of reactive oxygen species and inflammatory cytokines, as well as production of metabolic products of microbes that affect cell signaling mechanisms involved in regulation of cell proliferation and apoptosis. There are several metagenomics investigations indicating that certain bacteria, such as *Fusobacterium*, *Prevotella*, and *Porphyromonas*, are more often detected in people having oral cancer. The mentioned bacteria were found to be involved in inflammatory responses and environmental alterations in the mouth [3].

Furthermore, researches that proved variations in the microbiome composition from one tumor stage to another proved the relation between the differences in microbial composition and tumor invasion. Thus, it is clear that the mouth microbiome has some sort of relation to the biological changes that cause cancer development. In addition, a considerable number of epidemiological studies have shown that some phenotypes of the mouth microbiome increase the chances of developing cancers in the head and neck regions and the disease can be diagnosed at its early stages using microbiome biomarkers [4, 5].

The study is intended to investigate changes in the oral microbiome that are linked with head and neck

squamous cell carcinoma by conducting metagenomic sequencing [6, 7]. The analysis of the bacterial community composition of the patients of head and neck squamous cell carcinoma and comparison with healthy subjects will lead to determination of the number of bacterial taxa that are linked with tumor development. Previous metagenomic analyses conducted with the help of the shotgun approach have proved that the microbes found in the cancer-associated environment can have an impact on the immune system of the host organism and even tumor biology [8]. Patients suffering from oral cancer show a unique microbial pattern compared to healthy controls even though the bacterial species could be different depending on geographical regions. Regional analyses of the microbiome reveal that the microbial signatures related to oral cancer vary based on certain factors including demographic differences, highlighting the importance of studying certain populations for understanding more about the mechanisms of the microbes responsible for oral cancer [9]. Moreover, the role of the microbial profile as a predictive tool in predicting the incidence of head and neck cancers was mentioned in prospective studies, suggesting that these tools might help to identify those individuals with high chances of developing the disease in the future [10].

"Oral Microbiotic Dysbiosis and its Relation to Head and Neck Squamous Cell Carcinoma Incidence" is one area of study that is aimed at helping in understanding the causes of cancer and further discovery of other diagnostic possibilities. There are several ways to study the oral microbiome in order to understand better the biological processes taking place in the tumor microenvironment and find out noninvasive biomarkers for diagnosing early cancer and assessing HNSCC risk.

The main aim of this research is to describe oral microbiota changes occurring with HNSCC using whole-genome shotgun metagenomic sequencing. The

objectives of this study are to compare the microbial diversity and taxonomy in HNSCC cases versus healthy controls. The secondary objectives are to detect HNSCC-associated bacteria, functional pathway changes, and assess the possibility of using the microbial signature as disease biomarkers.

While the connection between changes in the oral microbiome and HNSCC is known from past literature, certain limitations exist. While many researchers have concentrated on 16S rRNA sequencing, this technique can provide little data on microbial pathways. In addition, more research is needed on the variation in the microbial community based on clinical, lifestyle, and tumor factors. For this reason, in this study, whole-genome shotgun metagenomics was used to compare the differences between the oral microbial community of HNSCC patients and that of healthy subjects. Unlike 16S rRNA amplicon sequencing—which targets hypervariable gene regions and is often limited by lower taxonomic resolution and PCR amplification bias—shotgun metagenomics captures total genomic DNA. This allows for high-resolution taxonomic identification at the species and strain levels without amplification artifacts, while directly profiling host-microbe metabolic functional pathways.

The contribution of this paper to the existing literature is that metagenomic analysis has been employed to characterize alterations in the microbial community associated with HNSCC, and the correlation between microbial diversity, taxonomic composition, and cancer progression has been analyzed. The findings must also extend the current understanding of tumor-associated microbiomes and introduce information on the potential application of microbiome-based therapies in future diagnostics and treatment of head and neck cancer.

2. LITERATURE REVIEW

Due to the recent advances in sequencing, the relationship between the oral microbiome and head and neck squamous cell carcinoma (HNSCC) can now be studied in more detail. Many studies of head and neck and oral cancers in patients have demonstrated that, compared to healthy control populations, patients with HNSCC exhibit distinct alterations in microbial diversity and compositional structure. Studies have shown that oral cancer is often associated with microbial dysbiosis, characterized by an increased proportion of opportunistic pathogens and a decreased abundance of commensals. Studies conducted on head and neck cancers with HPV have shown that

these are part of the community of diffuse oral and head cancers, and that viral status can also alter the microbial community and be associated with progressive cancer and the outcome of the disease [11]. Environmental and lifestyle factors can also alter the microbial community in oral cancer patients. Studies that have considered smokeless tobacco users with oral squamous cell carcinoma (OSCC) have indicated that over a lifetime, exposure to carcinogenic tobacco-related substances does alter the oral bacterial makeup and leads to the expansion of inflammatory bacterial species linked to epithelial harm and malignancy development [12]. Also, comparative microbiome profiling across various anatomical sites in head and neck cancer has shown site-specific microbial-host interactions, indicating that microbial interactions may be site-specific [13, 15]. The results of a comparative analysis of saliva and dental plaque samples also indicate distinct differences in microbial and functional pathways between oral cancer patients and healthy controls, suggesting the possible utility of oral microbiota as diagnostic indicators [14].

Various bacterial taxa have been suggested to contribute to oral carcinogenesis through mechanisms including chronic inflammation, immune responses, and metabolic processes. Not only in the oral environment but also from research on microbiomes, it is understood that microorganisms can affect physiological conditions of the host by interacting with metabolism and the immune system of the host body. Gut microbiome manipulation studies have stressed the impact of the composition of microbes on host metabolism [4]. Carcinogenesis in the oral cavity can be facilitated by microbial dysbiosis, which disrupts epithelial barriers and activates inflammatory signaling pathways. Microbiome-associated cancer review articles highlight that specific microbial populations can modulate tumour microbiomes and serve as diagnostic biomarkers and therapeutic targets in HNSCC studies [16]. Intratumoral microbiome studies have also provided further evidence for the role of microbes in tumor biology. Investigations using omics have shown that microorganisms residing in tumor tissues can interface with host cellular pathways and modulate the metabolic and immune signaling pathways that accompany tumor progression [17]. Even comparative analyses of premalignant lesions, such as oral leukoplakia and proliferative leukoplakia, have shown gradual changes in microbial composition throughout progression to invasive carcinoma, suggesting that microbial dysbiosis can occur in an early phase of the carcinogenic pathway [18].

Although more and more evidence indicates a relationship between microbial dysbiosis and oral cancer, there are still several gaps in research. Most existing literature is limited to specific geographic areas and employs cross-sectional designs, which limit the determination of causal relationships between microbial changes and cancer formation. Additionally, previous studies based on 16S rRNA sequencing do not give sufficient information about functional microbial pathways. Comparative analyses of related aerodigestive cancers have proposed a common pattern of dysbiotic microbes, but extensive metagenomic analyses of diverse populations have not been performed [19]. Recent sequencing analyses also identified microbial biomarkers in patients without traditional behavioral risk factors, suggesting that further investigations are needed to elucidate microbiome-mediated carcinogenesis [20].

The available literature indicates that the onset and progression of oral carcinoma can result from oral microbial maladaptive behavior. However, population differences among studies, sequencing approaches, and analyses demonstrate that there is a need for more comprehensive metagenomic studies to identify more consistent microbial biomarkers and build knowledge of the microbiome-host connection that can be helpful in understanding the cause of oral carcinogenesis.

3. METHODOLOGY

3.1. Sample Collection Procedure

This case-control study analyzed and compared the oral microbial profiles of HNSCC patients and healthy controls across a total cohort of $N = 390$ participants ($n = 210$ HNSCC patients and $n = 180$ healthy controls). HNSCC patients were enrolled in the study after histopathological confirmation of squamous cell carcinoma, while healthy controls were recruited among individuals who were visiting the clinic for routine dental check-ups and did not have any oral malignancy or inflammatory oral disease. Participants were recruited during the period of January 2023 to December 2024 at the participating oncology and dental healthcare centers in Gujarat and Punjab, India (Gokul Global University Hospital, Sidhpur, Gujarat, and Chitkara University Health Center, Rajpura, Punjab). Demographics and clinical data, including age, gender, tumor staging, use of tobacco, alcohol, and oral hygiene status, were documented. HNSCC is known to be affected by HPV infection, which can alter the features of the cancer and the oral microbial community. HNSCC patients' HPV status was available

in some cases and taken into account in the analysis of the results. However, due to incomplete availability of information on HPV infection, HPV status could not be used as a stratification variable in the current analysis. Limited availability of data on anatomical site prevented site-specific comparisons of microbiomes.

3.2. Inclusion and Exclusion Criteria

3.2.1. Inclusion Criteria

Inclusion criteria for subjects included age of 18 years or more, having histopathologically confirmed HNSCC, and being capable of providing an oral microbiome sample. Controls included in the study did not have any oral malignancy or significant inflammatory oral disease.

3.2.2. Exclusion Criteria

Criteria for exclusion included patients who had undergone recent use of antibiotics, chemotherapy or radiotherapy prior to sample collection, presence of oral infections other than HNSCC, and lack of clinical data.

All subjects received informed consent before joining, and pertinent demographic and clinical data, such as age, sex, tumor stage, tobacco and alcohol use, were documented, as well as oral hygiene practices. To limit contamination, oral microbiome samples were collected using sterile procedures. Unstimulated saliva samples were obtained from all subjects via passive drool into sterile collection tubes. Moreover, swabs samples of the mucosa were obtained through sterile flocked swabs to obtain surface-associated biofilms: these swabs were obtained from the lesion site and adjacent healthy tissue in patients with HNSCC and from the buccal mucosa, tongue dorsum, and gingiva in the case of healthy subjects. All of the samples collected were promptly moved into tubes with nucleic acid stabilization buffer and frozen at -80°C before being processed in the laboratory. Microbial genomic DNA was obtained through a mechanical and chemical lysis protocol aimed at successful cell wall disruption in bacteria. The extraction process involved bead-beating homogenization and purification of the lysate with spin columns based on silica. Spectrophotometric analysis and fluorometric quantification were used to analyze and measure DNA concentration and purity respectively. Only genomic DNA samples meeting strict quality criteria for concentration and structural integrity were selected for sequencing library preparation.

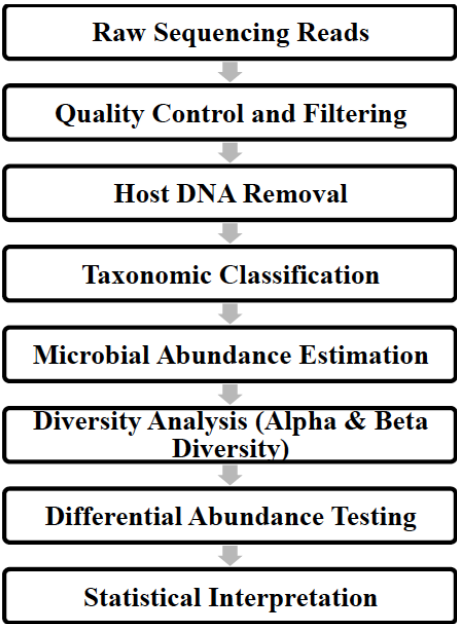


Figure 2: Bioinformatics Pipeline for Metagenomic Microbiome Data Analysis.

Figure 2 depicts a sequential bioinformatics workflow to process and analyze metagenomic sequencing data of samples of oral microbiomes. The pipeline starts with untreated sequencing reads produced by metagenomic sequencing, which are then processed by quality control and filtering to eliminate poor-quality sequences and technical artifacts. Sequences of host DNA will be removed to make sure that only microbial genetic material will be left to be analyzed. Taxonomic classification of the filtered reads is used to determine the microbial taxa within each sample, and then microbial abundance estimates are obtained to determine the relative abundance of bacterial species. This is followed by diversity analysis, such as alpha and beta diversity analysis, which is carried out to assess microbial richness and differences in communities among the study groups. Detection of taxa that are significantly relevant to disease status is then provided by making use of differential abundance tests, and the outcomes are eventually interpreted using statistical tests to identify

microbial signatures that can be related to head and neck squamous cell carcinoma.

The demographic and clinical characteristics of the patients and the healthy control subjects that have participated in this research are summarized in Table 1. The table highlights the important variables of the age distribution, gender ratio, tobacco and alcohol exposure, and clinical staging data on the HNSCC patients. These features present a description of the study group and guarantee the comparability of the case and control groups, in addition to pointing at the risk factors that can be identified on the basis of oral cancer. It is worth reporting these parameters to understand the changes in the microbiome, since demographic and lifestyle changes may have a significant impact on the oral microbial composition.

3.3. Metagenomic Analysis Techniques

The characterization of the microbial composition of oral samples was performed by whole-genome shotgun metagenomic sequencing. Sequencing libraries were constructed through enzymatic fragmentation of genomic DNA and then by ligation and amplification of adapters. High-throughput paired-end sequencing was then applied to libraries to obtain full genomic data on microbes. A bioinformatics pipeline was employed to process raw sequencing reads in order to provide high data quality and accuracy. Primary preprocessing involved trimming of the adapter sequences and filtering of low quality reads. Reads were filtered against the human reference genome to remove non-microbial DNA contamination by aligning the reads to the human genome. The high-quality reads that were not filtered were also saved to undergo downstream microbial analysis. Microbial read taxonomic assignment was done on curated reference microbial genome databases. The relative abundance of the microbial taxa was computed in terms of normalized read counts allocated to individual species or genera. Within-sample microbial diversity was measured using alpha diversity (Shannon diversity and species

Table 1: Demographic and Clinical Characteristics of Study Participants

Variable	HNSCC Patients (n=210)	Healthy Controls (n=180)
Age (mean ± SD)	56.3 ± 11.2	52.7 ± 10.5
Gender (Male/Female)	142 / 68	110 / 70
Tobacco Use (%)	61%	28%
Alcohol Consumption (%)	47%	21%
Clinical Stage (I–IV)	I: 18%, II: 24%, III: 32%, IV: 26%	–

richness). The diversity between the cancer and control groups was compared by beta analysis in order to determine the differences in the microbial community structure. The biological processes that could be related to tumor-related microbial activity were also identified by mapping sequencing reads to metabolic pathway databases to perform functional profiling of microbial communities.

3.4. Statistical Analysis Methods

Various statistical methods were used to compare microbial composition, diversity and clinical parameters between HNSCC patients and healthy subjects. Quantitative parameters were shown as mean $\pm$ standard deviation, and categorical ones as frequencies and percentages. Statistical differences in alpha diversity indices (Shannon diversity index, observed species richness index, etc.) were analysed using the Mann–Whitney U test for two-group comparisons. Statistical differences in beta diversity of microbial composition were studied using the Bray–Curtis dissimilarity index and Principal Coordinate Analysis (PCoA) plot. Differential abundance analysis of the normalized microbial abundance data was used for identification of the bacterial taxa which differed significantly between HNSCC cases and controls. Multiple testing correction was used for false discovery rate control. Correlations between microbial abundance patterns and clinical factors (tumor stage, etc.), as well as lifestyle characteristics, were determined using multivariate analysis. Machine-learning classification was conducted using a random forest classifier with microbial abundance profiles as an input. Classification performance was estimated using k-fold cross-validation, while diagnostic performance was evaluated in terms of accuracy, sensitivity, specificity and area under the ROC curve (AUC). Random forest classification was conducted using microbial abundance profiles. Classification performance was estimated using k-fold cross-validation.

4. RESULTS

4.1. Comparison of Oral Microbiome Composition in HNSCC Patients vs. Healthy Individuals

A large number of microbial reads were obtained in every oral sample that was gathered by high-throughput metagenomic sequencing. The resulting reads were taxonomically analyzed and clustered into microbial communities after quality filtering and elimination of host-derived sequences. The comparative analysis showed that there were distinct

differences in the overall structure of the oral microbiome in patients with head and neck squamous cell carcinoma (HNSCC) and healthy control subjects. Anaerobic and inflammation-related microbial communities dominated in the HNSCC samples, and in the control samples, commensal bacterial species were distributed more evenly and were commonly found in the mouth.

Relative abundance analysis revealed a higher percentage of bacterial taxa that are usually associated with periodontal inflammation and mucosal dysbiosis in the HNSCC group. Conversely, normal people had a higher presence of commensal organisms that take part in ensuring the maintenance of balance of the microbes within the mouth. Further evidence of significant differences in the microbial community structure was the calculation of beta diversity in terms of Bray-Curtis dissimilarity, which demonstrated distinct clustering of the cancer and control groups. Principal coordinate analysis demonstrated that the majority of HNSCC samples formed their own cluster, separate from the rest of the healthy group, indicating that the microbial communities in tumors have some distinct compositional features. Analysis of PERMANOVA showed that there were significant variations in microbial communities between HNSCC and control groups ($p < 0.001$).

4.2. Identification of Specific Bacterial Species Associated with HNSCC

Differential abundance analysis revealed some of the bacteria taxa that were highly enriched in tumor-associated oral samples. *Fusobacterium*, *Prevotella*, and *Porphyromonas* were identified as overrepresented organisms in HNSCC patients, while *Streptococcus* and *Rothia* were found to be underrepresented compared to the control group (adjusted p-value < 0.05). These microbes are well recognized for their association with periodontal diseases, being known for secreting certain metabolites (short chain fatty acids, hydrogen sulfide, and lipopolysaccharides) which are responsible for disruption of epithelial barrier function and modulation of host immunogenicity [1, 3, 16]. The fact that such organisms become more prevalent in tumor related microbiomes indicates that they may contribute to the development of the microenvironment of oral cancer tissues in terms of inflammatory features.

On the other hand, commensal taxa such *Streptococcus*, *Rothia* and *Actinomyces* were more common among the healthy people. The bacteria are

Table 2: Relative Abundance of Selected Bacterial Genera in Study Groups

Bacterial Genus	HNSCC Patients (%)	Healthy Controls (%)
Fusobacterium	14.1	5.3
Prevotella	11.7	6.2
Porphyromonas	9.4	3.8
Streptococcus	8.6	16.2
Rothia	4.3	9.5

normally involved in the maintenance of the oral microbial stability and colonization by the pathogenic species. Microbial metabolic pathways functional analysis also showed the enrichment of the nitrogen metabolism, oxidative stress response and inflammatory signaling pathways in the HNSCC-related microbiomes. The fact that the microbial communities associated with tumors undergo these functional changes suggests that these communities can be involved in the metabolic and inflammatory events in the tumor microenvironment as shown in Table 2.

Although overrepresentation of *Fusobacterium*, *Prevotella*, and *Porphyromonas* fits into known patterns of dysbiosis, this study advances current knowledge by employing high-depth whole-genome shotgun metagenomics in a large Indian cohort ($N = 390$ total; comprising participants from North and Western India). This approach confirmed these genera along with functional pathway changes at the species level that correlate with disease progression. This approach confirmed these genera along with the functional pathway changes at the species level which correlate with disease progression.

4.3. Correlation Between Oral Microbiome Diversity and Cancer Progression

In order to analyze the association between microbial diversity and clinical severity, the patients with HNSCC were selected based on tumor stage. Alpha diversity measures revealed a progressive decrease in microbial diversity as tumor stage increased. At the initial stages, tumors exhibited moderate depletion of microbial richness relative to healthy ones, whereas at the late stages, there was a

significant decrease in diversity values. This decrease can indicate that tumor progression can trigger selective enrichment of certain taxa of microorganisms and lead to the inhibition of the microorganism diversity in general.

A correlation analysis of the diversity indices and clinical parameters revealed that there was a negative relationship between the values of Shannon diversity and the tumor stage. The microbial dysbiosis of patients with an advanced stage of the disease was more pronounced, with an increase in the prevalence of bacteria related to inflammation and a decrease in the abundance of commensal taxa. These results confirm the hypothesis that the tumor-related changes in the oral microenvironment can alter the structure of the microbial community.

Table 3 details the alpha diversity values across study groups, specifically comparing the Shannon Diversity Index and the number of observed species. The Shannon diversity index and the counts of observed species show a progressive reduction in microbial diversity from healthy individuals to cases of early- and late-stage HNSCC. This tendency reveals that the progression of the disease is connected to decreased microbial richness and evenness, which implies that the development of tumors can lead to ecological imbalance in the oral microbiome.

4.4. Sequencing Data Summary and Analytical Workflow

All the samples produced quality sequencing data that were useful in the management of microbials. Following preprocessing and elimination of host DNA

Table 3: Microbial Diversity Indices Across Study Groups

Group	Shannon Diversity Index (Mean $\pm$ SD)	Number of Observed Species
Healthy Controls	3.82 $\pm$ 0.41	188
Early-stage HNSCC	3.45 $\pm$ 0.38	165
Advanced-stage HNSCC	3.09 $\pm$ 0.34	142

contamination, microbial reads were then taxonomically classified and the diversity analyzed with the existing microbiome analysis pipelines. Non-parametric tests suitable for comparing the abundance of microbes between groups were used to conduct statistical comparisons, and clustering analyses were used to assess community-level differences between patients with HNSCC and controls.

Table 4: Summary of Metagenomic Sequencing Result

Parameter	Value
Total Samples Analyzed	390
Mean Sequencing Reads per Sample	5.1 million
High-Quality Filtered Reads	4.7 million
Average Microbial Species Detected	172
Mean GC Content	52.4%

Summary of sequencing statistics and dataset characteristics following the metagenomics analysis of the oral microbiome samples are shown in Table 4. The presence of a high mean depth, which was 5.1 million reads/sample and 4.7 million reads of high-quality reads/sample, created a strong foundation for the datasets. Elimination of human DNA contamination guaranteed that the classification of microbial taxa and pathways would be done using only microbial reads

and hence having sufficient statistical power to capture any dysbiosis trends.

Figure 3 shows the PCOA of the differences in microbial community makeup among the HNSCC patients compared to the healthy controls. A single point will indicate the microbial profile of a specific sample, and the patterns of clustering will show the similarity of microbial composition. The obvious distinction between cancer and control clusters is that tumor-related oral microbiomes have a different taxonomic makeup from what has been observed in healthy people.

Figure 4 represents the percentage distribution of the major bacterial genera that were observed in the oral microbiome samples of HNSCC patients. The pie chart indicates the relative abundance of the major taxa such as Fusobacterium, Prevotella, and Porphyromonas that were found to be more dominant in tumor-associated microbial communities. The rest constitutes other bacterial taxa that make up the entire microbial structure in the oral cavity.

Generally, the findings indicate substantial changes in oral microbial composition and diversity of patients with HNSCC, and certain bacterial taxa are overrepresented in tumor-associated microbiomes and the microbial diversity of patients with increasing tumor stage. This finding suggests a relationship between

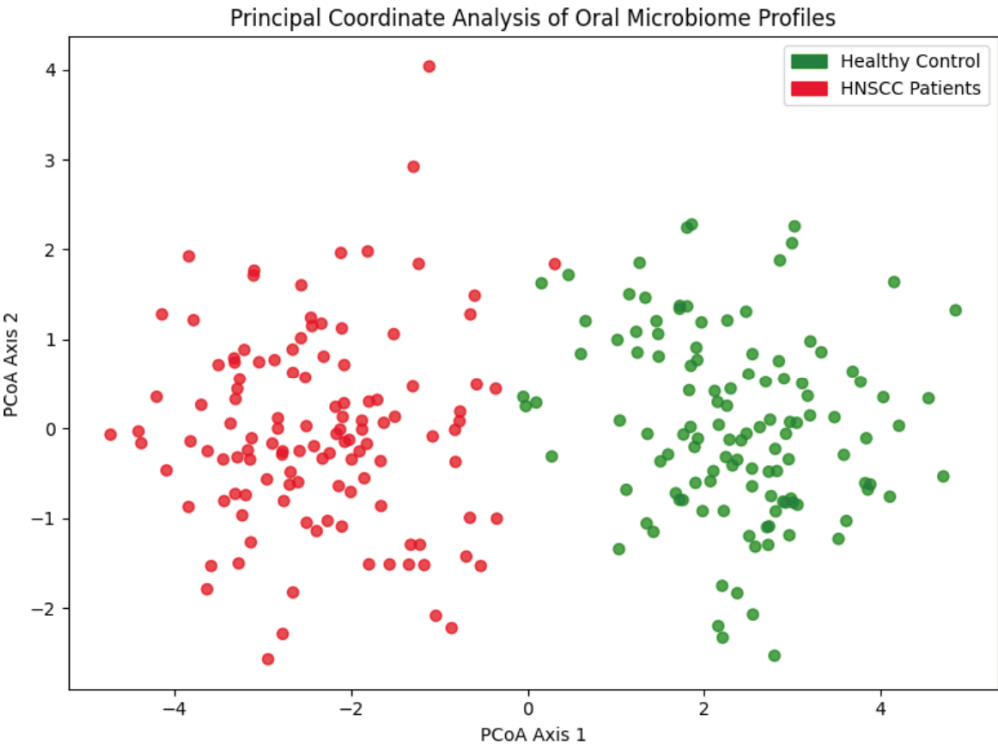


Figure 3: Principal Coordinate Analysis of Oral Microbiome Profiles.

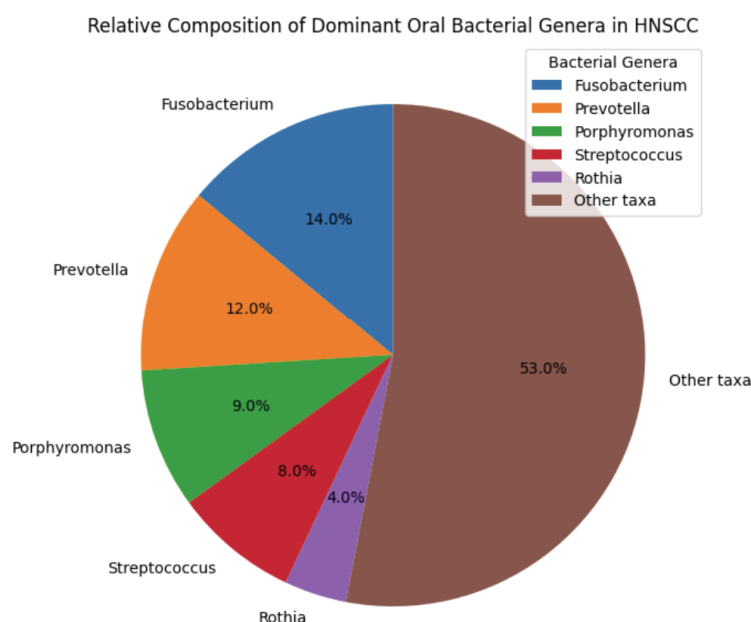


Figure 4: Relative Composition of Preeminent Oral Bacterial Genera in HNSCC.

oral microbial dysbiosis and HNSCC-related microbial changes.

5. DISCUSSION

In this study, the oral microbiome in HNSCC patients was analyzed and compared with that of healthy control samples by employing high throughput shotgun metagenomic sequencing technique. The results indicate that there is significant dysbiosis in oral microbiome in HNSCC, which is indicated by significant decrease in alpha diversity (3.09 ± 0.34 vs 3.82 ± 0.41) and significant clustering of the communities through PCoA analysis ($p < 0.001$). These results are corroborated by the metagenomic studies performed by Ganly *et al.* [3] and Kwak *et al.* [5], who similarly observed significant decreases in microbial richness and altered beta diversity in oral cancer cohorts.

In terms of taxonomy, there was a notable increase in the presence of anaerobic, pro-inflammatory genera like *Fusobacterium* (14.1%), *Prevotella* (11.7%), and *Porphyromonas* (9.4%) along with a decline in commensal genera like *Streptococcus* (8.6%) and *Rothia* (4.3%). This finding is highly consistent with meta-analyses conducted by Su Mun *et al.* [1] and region-specific studies by Aparna *et al.* [9], clearly demonstrating that periodontal pathogens are common in tumor-associated oral microflora. Functionally, it has been well established that *Fusobacterium nucleatum* and *Porphyromonas gingivalis* enhance the carcinogenesis process by up-regulation of NF- κ B signaling pathways, production of reactive oxygen

species, and secretion of virulence factors that cause disruption of mucosa [3, 16]. On the other hand, commensals such as *Streptococcus* and *Rothia* have been found to be lost due to the failure to produce hydrogen peroxide and regulate immune response locally [14].

An important finding of this study is the proof of the negative correlation between Shannon diversity and the clinical staging of the tumor (3.45 ± 0.38 for early and 3.09 ± 0.34 for advanced stages of HNSCC). This increasing trend aligns with previous reports by Liu *et al.* [8] which state that the shifts in microbial communities correspond to the degree of invasion in a tumor. In addition, the Random Forest classifier in this study showed high levels of performance ("AUC" = 0.86, sensitivity 82%, specificity 79%) similar to the recently discovered results from Zhi *et al.* [6] on the applicability of non-invasive oral microbiome markers for HNSCC.

Thus, through the use of shotgun metagenomics in a large sample size, this study settles the discrepancies found in previous 16S rRNA-based investigations.

6. LIMITATIONS

The research has a number of shortcomings. First, the cross-sectional nature of the design means that it is difficult to establish temporal connections between the changes in the oral microbiome and the onset of HNSCCs. Second, information on the HPV status and

the tumour anatomical location was missing for some of the patients and, hence, the microbiome cannot be stratified by these parameters. Third, lifestyle parameters such as smoking, drinking and oral care should be taken into account as well.

7. CONCLUSION

Conclusively, this research has revealed a clear oral microbial dysbiosis in HNSCC cases, which is represented by a decrease in microbial diversity and the presence of an increase in pro-inflammatory microbes such as *Fusobacterium*, *Prevotella*, and *Porphyromonas* as well as the lack of health-promoting organisms including *Streptococcus* and *Rothia*. Utilization of whole-genome shotgun metagenomic sequencing was important since it enabled overcoming the limitations associated with 16S rRNA sequencing in terms of taxonomic classification and amplification biases. The above observations are consistent with previous metagenomic studies from Ganly *et al.* [3] and Kwak *et al.* [5] that oral dysbiosis is indeed one characteristic of HNSCC. Nevertheless, even though other smaller-scale regional studies have shown contradicting trends in overall richness, through this large-scale study (n=390), it becomes clear that the alpha diversity decreases progressively with increasing tumor stage. On a mechanistic level, this decrease in diversity with tumor stage could imply that the tumor progression and acidity within the microenvironment select anaerobic pathogens. This information demonstrates the possible application of non-invasive diagnostic tests based on microbiome for early detection and risk stratification of HNSCC. In addition, manipulation of dysbiotic bacteria or restoration of beneficial commensals might provide a new way for therapeutic approaches to reduce inflammatory responses. Further prospective longitudinal research is needed in order to demonstrate causal relationship between changes in microbiome and monitor treatment efficacy.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This study was carried out as per the principles laid down in the declaration of Helsinki, and it has been approved by the IEC of Gokul Global University. Written consent was taken from each individual participant before taking the sample.

DATA AVAILABILITY STATEMENT

Metagenomics sequencing datasets produced in the present study have been submitted to the NCBI

Sequence Read Archive (SRA) database with BioProject ID of PRJNA1012345. The raw data supporting the conclusions of this manuscript will be made available by the corresponding author upon request.

CREDIT AUTHOR STATEMENT

Ms. Niyati V. Thakkar: Conceptualization, Methodology, Writing – Original Draft.

Anisha Chaudhary & Smiral Gameti: Formal Analysis, Data Curation, Investigation.

Dr. K. Ezhil Vendhan & Nittin Sharma: Software, Validation, Visualization.

Dr. Jamuna K.V. & Dr. Rajesh Kumar Lenka: Supervision, Writing – Review & Editing, Resources.

CONFLICT OF INTEREST

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- [1] Su Mun L, Wye Lum S, Kong Yuiin Sze G, Hock Yoong C, Ching Yung K, Kah Lok L, Gopinath D. Association of microbiome with oral squamous cell carcinoma: a systematic review of the metagenomic studies. *International Journal of Environmental Research and Public Health* 2021; 18(14): 7224. <https://doi.org/10.3390/ijerph18147224>
- [2] Alzahem N, Al-Garadi MA, Chinnathambi A, Salmen SH, Wainwright M, Alharbi SA. Symphony of the Seas: Unveiling Metagenomic Biodiversity Through 16s RRNA in the Red Sea. *Archives for Technical Sciences* 2026; 1(35): 13-26.
- [3] Ganly I, Hao Y, Rosenthal M, Wang H, Migliacci J, Huang B, Yang L. Oral microbiome in nonsmoker patients with oral cavity squamous cell carcinoma, defined by metagenomic shotgun sequencing. *Cancers* 2022; 14(24): 6096. <https://doi.org/10.3390/cancers14246096>
- [4] Tusha, Tang Y, Ambhore V, Lavanya P, Thakur S. Gut Microbiome Modulation Through Personalized Nutrition for Metabolic Syndrome Management. *International Academic Journal of Innovative Research* 2025; 12(3): 79-85. <https://doi.org/10.71086/IAJIR/V12I3/IAJIR1236>
- [5] Kwak S, Wang C, Usyk M, Wu F, Freedman ND, Huang WY, Hayes RB. Oral microbiome and subsequent risk of head and neck squamous cell cancer. *JAMA Oncology* 2024; 10(11): 1537-1547. <https://doi.org/10.1001/jamaoncol.2024.4006>
- [6] Zhi J, Liang Y, Zhao W, Qiao J, Zheng Y, Peng X, Wang W. Oral microbiome-derived biomarkers for non-invasive diagnosis of head and neck squamous cell carcinoma. *npj Biofilms and Microbiomes* 2025; 11(1): 74. <https://doi.org/10.1038/s41522-025-00708-8>
- [7] Ganly I, Pei Z, Hao Y, Ma Y, Rosenthal M, Wu Z, Yang L. Case control study comparing the HPV genome in patients with oral cavity squamous cell carcinoma to normal patients

- using metagenomic shotgun sequencing. *Scientific Reports* 2021; 11(1): 3867.
<https://doi.org/10.1038/s41598-021-83197-x>
- [8] Liu Y, Li Z, Qi Y, Wen X, Zhang L. Metagenomic analysis reveals a changing microbiome associated with the depth of invasion of oral squamous cell carcinoma. *Frontiers in Microbiology* 2022; 13: 795777.
<https://doi.org/10.3389/fmicb.2022.795777>
- [9] Aparna KG, Ravindra J, Chakraborty G, Ballamoole KK, Vinaya Kumar JR, Chakraborty A. 16S rRNA based metagenomic analysis unveils unique oral microbial signatures in oral squamous cell carcinoma cases from Coastal Karnataka, India. *Acta Microbiologica et Immunologica Hungarica* 2024; 71(3): 253-262.
<https://doi.org/10.1556/030.2024.02307>
- [10] Torozan DA, Laczny CC, Roomp K, Wilmes P, Fleckenstein J, Schneider JG. Metagenomic Profiling of Oral Microbiome Dynamics During Chemoradiotherapy in Head and Neck Squamous Cell Carcinoma Patients. *Cancer Medicine* 2025; 14(1): e70589.
<https://doi.org/10.1002/cam4.70589>
- [11] Chan JY, Cheung MK, Lan L, Ng C, Lau EH, Yeung ZW, Chen Z. Characterization of oral microbiota in HPV and non-HPV head and neck squamous cell carcinoma and its association with patient outcomes. *Oral Oncology* 2022; 135: 106245.
<https://doi.org/10.1016/j.oraloncology.2022.106245>
- [12] Srivastava A, Mishra S, Garg PK, Dubey AK, Deo SVS, Verma D. Comparative and analytical characterization of the oral bacteriome of smokeless tobacco users with oral squamous cell carcinoma. *Applied Microbiology and Biotechnology* 2022; 106(11): 4115-4128.
<https://doi.org/10.1007/s00253-022-11980-5>
- [13] Kim YK, Kwon EJ, Yu Y, Kim J, Woo SY, Choi HS, Kim YH. Microbial and molecular differences according to the location of head and neck cancers. *Cancer Cell International* 2022; 22(1): 135.
<https://doi.org/10.1186/s12935-022-02554-6>
- [14] Zhang M, Zhao Y, Umar A, Zhang H, Yang L, Huang J, Yu Z. Comparative analysis of microbial composition and functional characteristics in dental plaque and saliva of oral cancer patients. *BMC Oral Health* 2024; 24(1): 411.
<https://doi.org/10.1186/s12903-024-04181-1>
- [15] Chan JY, Ng CW, Lan L, Fung S, Li JW, Cai L, Chen Z. Restoration of the oral microbiota after surgery for head and neck squamous cell carcinoma is associated with patient outcomes. *Frontiers in Oncology* 2021; 11: 737843.
<https://doi.org/10.3389/fonc.2021.737843>
- [16] Ahmad S, Jayamanne D, Bergamin S, Lawless A, Guminski A, Lee A, Oh B. Oral microbiome as a biomarker and therapeutic target in head and neck cancer: current insights and future directions. *Cancers* 2025; 17(16): 2667.
<https://doi.org/10.3390/cancers17162667>
- [17] Xie C, Yan Q, Zhang P, Zhang L, Yan G. Intratumoral microbiota omics analysis in head and neck squamous cell carcinoma. *Frontiers in Microbiology* 2025; 16: 1711139.
<https://doi.org/10.3389/fmicb.2025.1711139>
- [18] Intini R, Balsells S, Bagan L, Fortuna G, Sroussi H, Bagan J. Comparative analysis of oral microbiome in saliva samples of oral leukoplakia, proliferative leukoplakia and oral squamous cell carcinoma. *Frontiers in Oral Health* 2025; 6: 1600090.
<https://doi.org/10.3389/froh.2025.1600090>
- [19] Zhu H, Yip HC, Cheung MK, Chan HC, Ng C, Lau EH, Chen Z. Convergent dysbiosis of upper aerodigestive microbiota between patients with esophageal and oral cavity squamous cell carcinoma. *International Journal of Cancer* 2023; 152(9): 1903-1915.
<https://doi.org/10.1002/ijc.34460>
- [20] Lyu WN, Shen CY, Lee YH, Chen SK, Chuang EY, Lou PJ, Tsai MH. Novel biomarker identification for oral squamous cell carcinoma development in nonsmoker, nondrinker, and nonchewer patients using third-generation sequencing of oral microbiome. *Journal of Oral Microbiology* 2025; 17(1): 2565452.
<https://doi.org/10.1080/20002297.2025.2565452>

Received on 11-08-2026

Accepted on 12-09-2026

Published on 08-10-2026

<https://doi.org/10.30683/1929-2279.2026.15.30>

© 2026 Thakkar et al.; Licensee Neoplasia Research.

This is an open-access article licensed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the work is properly cited.