



# Dysbiosis of Oral Microbiota During Oral Squamous Cell Carcinoma Development

Purandar Sarkar<sup>1</sup>, Samaresh Malik<sup>1</sup>, Sayantan Laha<sup>2</sup>, Shantanab Das<sup>2</sup>, Soumya Bunk<sup>3</sup>, Jay Gopal Ray<sup>4</sup>, Raghunath Chatterjee<sup>2</sup> and Abhik Saha<sup>1,3\*</sup>

<sup>1</sup> School of Biotechnology, Presidency University, Kolkata, India, <sup>2</sup> Human Genetics Unit, Indian Statistical Institute, Kolkata, India, <sup>3</sup> Department of Life Sciences, Presidency University, Kolkata, India, <sup>4</sup> Department of Oral Pathology, Dr. R Ahmed Dental College and Hospital, Kolkata, India

## OPEN ACCESS

### Edited by:

Alberto Paderno,  
University of Brescia, Italy

### Reviewed by:

Nora Staniene,  
Lithuanian University of Health  
Sciences, Lithuania  
Bhabatosh Das,  
Translational Health Science and  
Technology Institute (THSTI), India

### \*Correspondence:

Abhik Saha  
abhik.dbs@presiuniv.ac.in

### Specialty section:

This article was submitted to  
Head and Neck Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 06 October 2020

**Accepted:** 05 January 2021

**Published:** 23 February 2021

### Citation:

Sarkar P, Malik S, Laha S, Das S,  
Bunk S, Ray JG, Chatterjee R and  
Saha A (2021) Dysbiosis of Oral  
Microbiota During Oral Squamous Cell  
Carcinoma Development.  
Front. Oncol. 11:614448.  
doi: 10.3389/fonc.2021.614448

Infection with specific pathogens and alterations in tissue commensal microbial composition are intricately associated with the development of many human cancers. Likewise, dysbiosis of oral microbiome was also shown to play critical role in the initiation as well as progression of oral cancer. However, there are no reports portraying changes in oral microbial community in the patients of Indian subcontinent, which has the highest incidence of oral cancer per year, globally. To establish the association of bacterial dysbiosis and oral squamous cell carcinoma (OSCC) among the Indian population, malignant lesions and anatomically matched adjacent normal tissues were obtained from fifty well-differentiated OSCC patients and analyzed using 16S rRNA V3-V4 amplicon based sequencing on the MiSeq platform. Interestingly, in contrast to the previous studies, a significantly lower bacterial diversity was observed in the malignant samples as compared to the normal counterpart. Overall our study identified *Prevotella*, *Corynebacterium*, *Pseudomonas*, *Deinococcus* and *Noviherbaspirillum* as significantly enriched genera, whereas genera including *Actinomyces*, *Sutterella*, *Stenotrophomonas*, *Anoxybacillus*, and *Serratia* were notably decreased in the OSCC lesions. Moreover, we demonstrated HPV-16 but not HPV-18 was significantly associated with the OSCC development. In future, with additional validation, this panel could directly be applied into clinical diagnostic and prognostic workflows for OSCC in Indian scenario.

**Keywords:** oral squamous cell carcinoma, 16S rRNA sequence analysis, oral microbiology ecology, dysbiosis, human papillomavirus-16

## INTRODUCTION

Oral squamous cell carcinoma (OSCC), a subset of head and neck squamous cell carcinoma (HNSCC), is the most common oral malignancy, representing approximately 90% of all cancers in the oral cavity (1, 2). It is the sixth most common cancer worldwide and every year around 400,000 new cases are diagnosed (3, 4). The number of newly diagnosed cases is predicted to increase by 62% in 2035 (5). The prevalence of oral cancer is highest in India and it represents most prevailing cancer in male population and fifth most common cancer among women (6, 7). Despite easy accessibility of the oral cavity during physical examination as well as several technological advancements in

surgical procedures in addition to adjuvant radiotherapy and chemotherapy, due to the lack of early diagnosis based on appropriate molecular markers, OSCC patients are often diagnosed at more advanced stages, leading to poor survival outcomes. The overall 5-year survival rate of OSCC patients is roughly 50% across the globe (8–10). Thus, early detection, identification of biomarkers and understanding the role of various etiological agents can significantly improve the current situation of OSCC treatment. In developing countries like India, excessive tobacco usage including smoking, chewing betel quid and areca nut along with alcohol consumption are the major risk factors for OSCC development (7, 11). However, oral cancer often arises in patients without a history of tobacco usage or alcohol consumption, indicating contribution from other potential risk factors including genetic/epigenetic alterations or microbial infection (2, 6, 12). A number of oncogenic viruses including high risk human papillomavirus (HPV) genotypes and Epstein-Barr virus (EBV) have been identified as infectious etiological agents for OSCC (13–15). However, association of these oncoviruses with OSCC development is not strong, contributing approximately 20% of all oral cancers (13, 14). Thus, identification of other microbial factors influencing OSCC development is warranted.

Human body harbors a plethora of microbial species referred to as ‘commensal microbiota’ including bacteria, yeast, fungi, protozoa, archaea and viruses and develops a symbiotic ecosystem without eliciting a decimating immune response. However, alterations of the microbiome architecture (dysbiosis) often lead to a variety of human diseases including cancer (16, 17). The advent of next-generation sequencing technologies for example the 16S rRNA gene amplicon based sequencing has allowed an affordable and culture-free approach of identification of overall bacterial composition in cancerous lesions and its effect on the progression of the disease (18). As one of the prime territories of microbiome in human body, the oral cavity contains distinct niches with dynamic microbial communities (19). Oral microbial ecology is a critical factor in controlling both human physiology and pathophysiology. The oral microbiome and their produced metabolites translocate through gastrointestinal tract or due to periodontal pocket ulceration can affect various distant tissues and are associated with the development of a number of diseases like cardiovascular disorder, diabetes, rheumatoid arthritis and premature birth (20–23). The dysbiosis of oral microbiome is associated with a number of clinical symptoms that ranges from dental caries, periodontal disease to oral cancer (24–28). Importantly, chronic periodontitis has also been suggested as potential risk factor for the onset of oral pre-cancerous and cancerous lesions (27). There is, however, no consensus among reports regarding microbiome signature associated with the development of OSCC. For example, Schmidt et al. demonstrated depletion of Firmicutes and Actinobacteria in a study of 15 oral cancer patients (29), while Mager et al. using DNA hybridization technique reported elevation of *Capnocytophaga gingivalis*, *Prevotella melaninogenica*, and *Streptococcus mitis* in the saliva of OSCC patients (30). Recently, Zhao et al. demonstrated that a cluster of periodontitis associated

taxa such as *Fusobacterium*, *Dialister*, *Peptostreptococcus*, *Filifactor*, *Peptococcus*, *Catonella*, and *Parvimonas* was enriched in OSCC lesions (31). Al-hebshi et al. reported that several inflammatory bacterial species including *Pseudomonas aeruginosa* and *Fusobacterium nucleatum* are elevated in OSCC patients’ samples (32). Another report by Lee et al. demonstrated that salivary microbiome particularly five genera including *Bacillus*, *Enterococcus*, *Parvimonas*, *Peptostreptococcus*, and *Slackia* significantly varied between samples from the epithelial precursor and OSCC lesions (33), indicating a potential prognostic marker for OSCC development. Börnigen et al. identified a number of differentially abundant genera in oral cancer samples specifically *Dialister* as enriched and *Scardovia* as depleted (34). Overall, given the diversity of identified microbiome composition as well as limited number of samples, more in depth investigations with larger-scale epidemiologically designed cohorts are warranted to assess the role of microbiome dysbiosis in OSCC development.

Despite the highest oral cancer incidence in India, till date no efforts have been made in understanding the oral microbial imbalance during OSCC development among Indian patients. Here, in an aim to explore OSCC-associated bacterial composition fifty OSCC lesions and their anatomically matched normal tissue regions was profiled using 16S rRNA gene amplicon based sequencing by targeting the hypervariable V3-V4 region. Our analyses revealed while top five genera such as *Prevotella*, *Corynebacterium*, *Pseudomonas*, *Deinococcus* and *Noviherbaspirillum* were significantly enriched, while genera including *Actinomyces*, *Sutterella*, *Stenotrophomonas*, *Anoxybacillus*, and *Serratia* were notably depleted in the OSCC lesions as compared to matched control adjacent tissue samples. In sum, our results provided evidence of alterations of oral bacterial community during OSCC development and indicated the possibility of utilizing the identified microbiome signature as prognostic marker of oral malignancies in patients of Indian subcontinent.

## MATERIALS AND METHODS

### Ethics Statement

The study was approved by the Institutional Ethics Committee for Human Research, Indian Statistical Institute, Kolkata, India. Written informed consent was obtained from all participants and all methods in this study were performed in accordance with the relevant guidelines and regulations.

### Sample Information

After the clinical diagnosis of oral squamous cell carcinoma (OSCC) the patients from Dr. R Ahmed Dental College and Hospital, were enlisted for the study. 50 patients were included in the study after confirmation of well-differentiated squamous cell carcinoma from histopathological reports. All participants were not on any local or systemic antibiotics prior to sample collection. Tissue samples were collected by incisional and 3 mm punch biopsy sample collection method from both the regions of cancerous lesions (N=50) and the adjoining clinically uninvolved normal area (matched control, N=50) for each of the 50 patients recruited in this study. A portion of the tissue

samples were collected in RNA Later (Invitrogen, Thermo Fisher Scientific Inc., Waltham, MA, USA) and stored at  $-80^{\circ}\text{C}$  for future use. Another portion was fixed in the formalin and used for histopathological evaluations.

## DNA Extraction

DNA was isolated from the cancerous lesion and adjacent unaffected normal tissue regions using DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) according to manufacturer's protocol. The quality and quantity of isolated DNA was determined by the A260/280 ratio using Synergy H1 Multimode Microplate Reader (BioTek Instruments, Inc., VT, USA). DNA samples were frozen at  $-20^{\circ}\text{C}$  for further analysis. Approximately 50 ng of genomic DNA from each sample was used for 16S rRNA V3-V4 amplicon sequencing.

## 16S Ribosomal RNA Sequencing and OTU Assignment

16S ribosomal RNA (rRNA) amplicon sequencing for metagenomics studies was performed on a MiSeq platform (Illumina, San Diego, CA, USA) using  $2\times 250$  bp chemistry. Clonal libraries for 16S rRNA V3-V4 hypervariable region were prepared using NEBNext Ultra DNA Library preparation kit (New England Biolabs, Ipswich, MA, USA) with the forward primer (5'-CTTCCCTACACGACGCT CTTCCGATCTACGGA GGCAGCAG-3') and the reverse primer (5'-GGAGTTCAGA CGTGTGCT CTTCCGATCTTACCAGGGTATCTAATCCT-3'). The amplicons were subjected to a number of enzymatic reactions for end-repairing, dA-tailing followed by adapter ligation and cleaning up using Solid Phase Reversible Immobilization (SPRI) technology (Beckman Coulter, Indianapolis, IN, USA). The adapter ligated fragments were indexed by limited PCR cycle to generate final libraries for paired-end sequencing. The concentration of the purified amplicons was measured using Qubit fluorometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) and the quality was checked using Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The multiplex amplified libraries were pooled in equimolar concentrations with unique indices, mixed with 15% PhiX control and sequenced using MiSeq reagent kit v2 (Illumina, San Diego, CA, USA) according to manufacturer's instruction.

The raw FASTQ sequencing files were further processed after checking base quality, base composition and GC content using FASTQC toolkit. The targeted amplicons were filtered out from other superfluous sequences by detecting the specific conserved region. Forward and reverse reads were stitched together with a minimum overlap of 30 bp and maximum overlap of 250 bp. De-replication was performed using USEARCH (35) (v11) for the identification of unique sequences and chimera sequences were filtered out using the UCHIME (36) algorithm in USEARCH package. Sequences that had a similarity of 97% were grouped together under a single operational taxonomic unit (OTU) against the GreenGenes database (release 2013-08: gg\_13\_8\_otus) using UPARSE (37) method. The taxonomy classification and relative

abundance assignments were performed using 'Quantitative Insights Into Microbial Ecology' (QIIME v. 1.9.0) (38) pipeline and singletons were discarded from the dataset to minimize the effect of low abundance sequences. To confirm the annotation, the resulting OTU representative sequences were then searched against the Ribosomal Database Project naïve Bayesian classifier (RDP 10 database, version 6) (39, 40) database, using the online program SEQMATCH (40). The taxonomy classifications at the phyla, order, family, genera and species level were performed using the GreenGenes and RDP databases.

## Diversity and Bacterial Enrichment Analyses

MicrobiomeAnalyst (41) was used for statistical analysis. The  $\alpha$ -diversity indexes including observed OTU numbers, Chao index, Simpson index, and Shannon index and the  $\beta$ -diversity – Bray-Curtis dissimilarity measurements were calculated. Evolutionary relation of the genera unique to the OSCC samples and the normal counterparts were analyzed and a cladogram was generated using Galaxy (42). The variation in genera as well as the unique bacterial composition in the normal and OSCC samples was identified using Random Forest (43) classification analysis within MicrobiomeAnalyst (41).

To estimate  $\beta$ -diversity, un-weighted and weighted UniFrac distances by Bray-Curtis method were calculated from the OTU abundance and utilized in Principal Component Analysis (PCoA) to analyze the unique clustering genera for the normal and OSCC affected tissue samples. PERMANOVA (44) algorithm on un-weighted and weighted UniFrac distance matrices was applied to generate PCoA plots. The differential abundances of OTUs and specific OTU enrichment between OSCC samples and matched controls were determined using LEfSe based on Kruskal-Wallis H test. Pairwise OTU enrichment analysis was performed to specifically identify the OTU abundance in each sample pair by comparing their true abundance values in the OSCC sample and its normal counterpart.

## Functional Prediction of Distinct Bacterial Communities

Functional compositions of the bacterial communities among two different groups were predicted using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) (45) according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (46).

## Real-Time PCR Primer Designing and Data Output

Primer-BLAST tool (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) in National Center for Biotechnology Information (NCBI) database was used to design primers for real-time quantitative PCR (qPCR) analyses. Primers for conserved sequence of bacterial 16S rRNA gene, human Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene and human oncogenic viruses

HPV16, HPV18 and EBV are listed in **Table S2**. qPCR primers were obtained from Integrated DNA Technologies, Inc. (Coralville, IA, USA). The optimum primer melting temperature ( $T_m$ ) was chosen at 60°C and the maximum GC content was kept at 55%. qPCR analysis was performed using iTaq Universal SYBR Green Supermix (BIO-RAD, Hercules, CA, USA) in CFX Connect Real-Time PCR detection System (BIO-RAD, Hercules, CA, USA) with the following thermal profile – one cycle: 95°C for 10 min; 40 cycles: 95°C for 10 s followed by 60°C for 10 s; and finally the dissociation curve at – 95°C for 1 min, 55°C 10 s, and 95°C for 10 s. Unless and otherwise stated, each sample was performed in duplicate and calculation was made using a  $-\Delta\Delta C_t$  method to quantify relative abundance compared with human genomic GAPDH control. The  $-\Delta C_t$  values of each OSCC samples and their match controls were plotted using GraphPad Prism 8.0.1.

## RESULTS

### Subject Characteristics and Oral Microbiota Profiling by 16S rRNA V3-V4 Amplicon Sequencing

To investigate changes in the oral microbiome associated with OSCC development, we prospectively collected cancerous lesions and anatomically matched adjacent normal tissue samples from four OSCC patients. Prior to proceeding into 16S rRNA amplicon based metagenomics studies, a preliminary verification was conducted for confirmatory presence of bacterial species in the isolated genomic DNA from the clinical samples. A real time quantitative PCR (qPCR) assay was performed using primer against the 16S rDNA conserved region. Excellent PCR amplification curves and  $-\Delta C_t$  values calculated against the housekeeping human GAPDH gene demonstrated the presence of bacterial species in both OSCC samples and their normal counterparts (**Figure S1**). To expand and substantiate our initial observations, we further collected another 46 pair of OSCC lesions and adjacent healthy tissue samples. The bacterial DNA was isolated from all 50 pair specimens (**Table S1**), followed by PCR amplification targeting the 16S rRNA V3–V4 hypervariable region. The 16S amplicons were purified, and a second round of index PCR was performed. The multiplex amplified libraries were pooled at equimolar concentrations and sequenced on an Illumina MiSeq platform. A total of 477,411 raw sequences were generated and after quality trimming and chimera checking, 322,656 high quality sequences were recovered for downstream analysis, with an average of 3,226 reads, ranging from 1,703 to 11,411 reads per sample.

### Genera Abundance and Diversity of Oral Microecology Were Depleted During OSCC Development

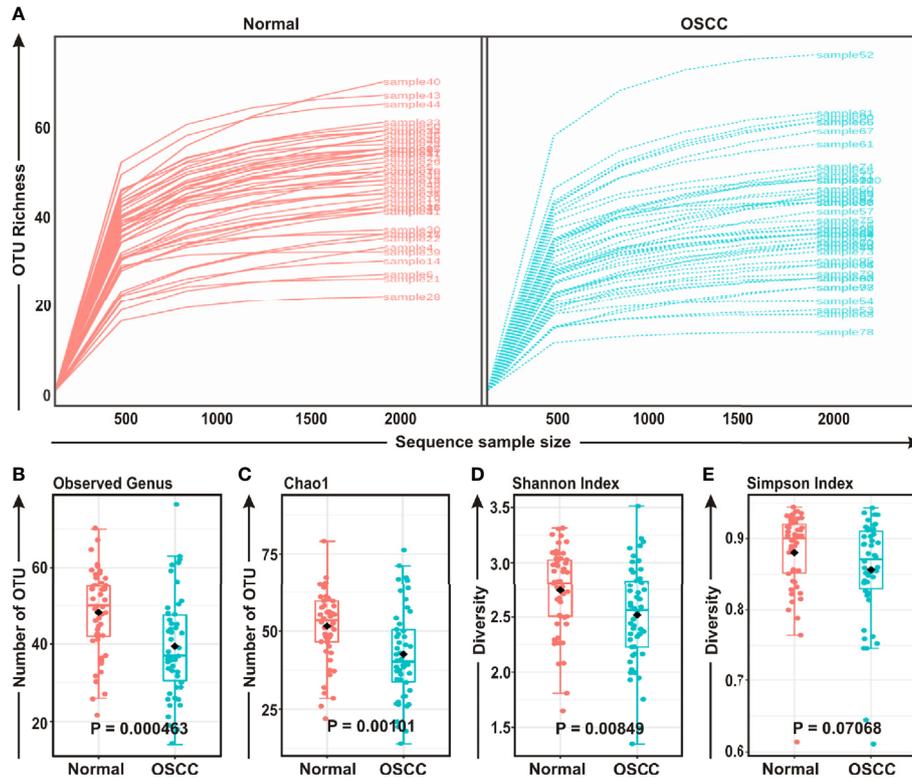
FASTQ files generated in 16S rRNA gene sequencing of all 50 samples containing high quality sequences were further analyzed using MicrobiomeAnalyst, a web-based online tool for comprehensive statistical analysis of microbiome data (41).

Rarefaction plot generated from the mapped reads indicated a clear difference of the OTU (Operational Taxonomic Unit) richness at the genus level between OSCC and normal samples (**Figure 1A**). Most of the samples, though not entirely, reached a saturated plateau phase, indicating further sequencing would possibly generate additional genera in those samples (**Figure 1A**). The average data of rarefaction plot demonstrated elevated genus richness in anatomically matched controls (samples 1–50) as compared to the paired contralateral OSCC lesions (samples 51–100), signifying a potential loss of certain genera during OSCC progression (**Figure 1A**). The OTU richness was further analyzed by calculating alpha diversity of oral microbiota, indicating the differences and similarities of the identified genera between the two sample groups (**Figures 1B–E**). The Observed genus ( $p = 0.000463$ ) and Chao1 index ( $p = 0.00101$ ) indicated that the OTU richness was significantly depleted in the OSCC samples as compared to the matched controls. The diversity estimator Shannon index ( $p = 0.00849$ ) indicated that relative diversity of bacterial genera was significantly decreased in cancerous lesions in contrast to the normal tissue sections (**Figures 1B–D respectively**). A similar trend of depletion in relative diversity of bacterial genera in OSCC samples as compared to the normal counterparts was also observed by employing Simpson index, although the data was not statistically significant ( $p = 0.070680$ ) (**Figure 1E**).

### Anatomically Matched OSCC and Normal Samples Comprised of Distinct Microbiome Composition

The beta diversity indicates the difference in the composition of bacterial community among different sample groups (47). To estimate  $\beta$ -diversity, weighted UniFrac distances as well as Bray-Curtis dissimilarity metric were calculated from the OTU abundance and utilized in Principal Component Analysis (PCoA) (47). Permutational Multivariate Analysis of Variance (PERMANOVA) (44) algorithm on Bray-Curtis dissimilarity and weighted UniFrac distance matrices was applied to generate PCoA plots (**Figures 2A, B, respectively**). The bacterial communities in the cancerous lesions and the anatomically matched controls clustered discretely, suggesting the overall structures of the bacterial communities in the groups were significantly different ( $p < 0.002$  and  $p < 0.009$ , respectively in two analyses) (**Figures 2A, B**).

A 'Random Forest' algorithm (43) was applied to further confirm the difference in bacterial community among the OSCC samples and anatomically matched healthy controls (**Figure 2C**). The decision trees extracted from the random forest classification identified distinct bacterial composition in diseased samples when compared with the normal counterparts. In the error plots identified from random forest analyses, while the red-line indicated the overall genera present in both OSCC and normal samples, green-line indicated the distinct genera present in the normal samples and the blue-line indicated the specific genera present in the OSCC lesions (**Figure 2C**). Moreover, in the total of 50 OSCC samples, 32 samples revealed unique and 18 samples demonstrated overlapping genera; whereas in case of contralateral



**FIGURE 1** | Comparison of oral microbiome compositions in the oral squamous cell carcinoma (OSCC) lesions and contralateral normal healthy groups. **(A)** Rarefaction analysis of bacterial 16S rRNA gene sequences of normal (red) and OSCC lesions (blue). Each line represents one sample. **(B–E)** Box-Whisker plots of **(B)** Observed operational taxonomic units (OTUs), **(C)** Chao 1 and **(D)** Shannon Index, **(E)** Simpson Index respectively.

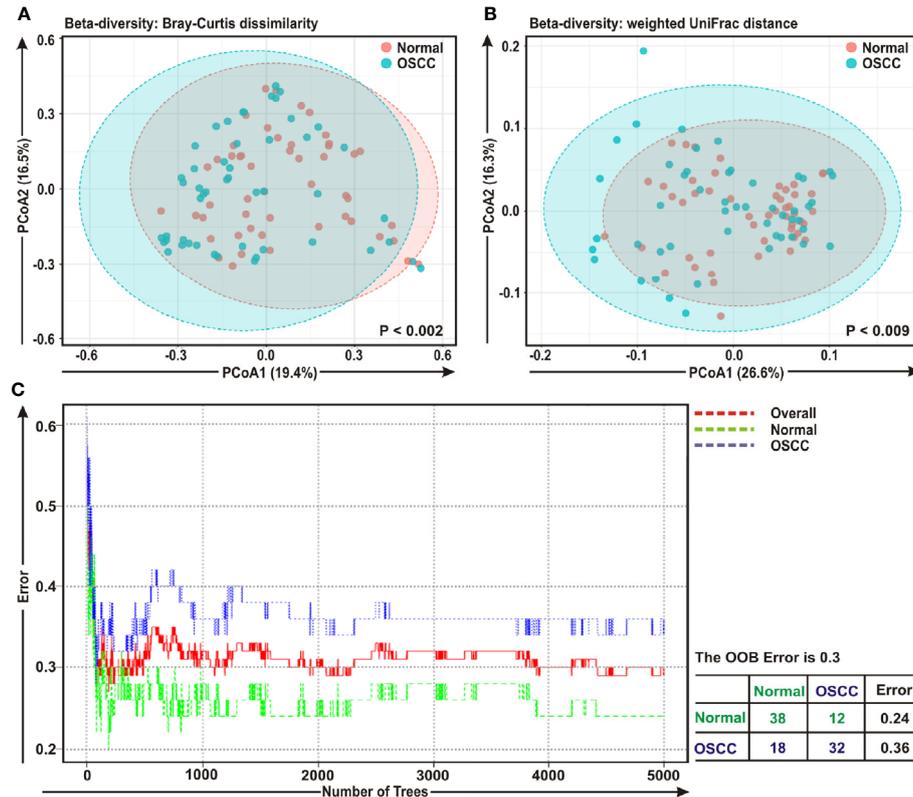
paired 50 normal tissue samples, 38 samples exhibited unique and 12 samples showed overlapping genera (Figure 2C).

### Phylogenetic Analysis Revealed Variations Among Common and Distinct Taxa in OSCC Lesions and Anatomically Matched Healthy Controls

The bacterial communities in the OSCC lesions and the anatomically matched healthy controls were first analyzed at phylum level (Figures 3A, B). The top five most abundant phyla including *Firmicutes*, *Proteobacteria*, *Fusobacteria*, *Bacteroidetes*, and *Actinobacteria* collectively comprised of 97.3% and 93% of all sequences in matched controls and OSCC lesions, respectively (Figure 3B). *Firmicutes* was the most abundant phylum in all samples, accounting for 36.1% of sequences in matched controls and 30.5% in OSCC lesions. In contrast, the abundances of the other detected phyla, including *Epsilonbacteraeota*, *Spirochaetes*, *Patescibacteria*, *Tenericutes*, *Synergistetes*, and *Deinococcus*, were less than 4.0%, ranges from 0.3% to 3.96%. While abundance of phyla including *Firmicutes*, *Proteobacteria* and *Actinobacteria* were reduced, *Fusobacteria*, *Bacteroidetes*, *Epsilonbacteraeota*, and *Spirochaetes* were elevated in OSCC lesions compared to normal healthy controls (Figures 3A, B). At the genus level, *Streptococcus*,

*Leptotrichia*, *Fusobacterium*, *Serratia*, *Neisseria*, *Haemophilus*, *Gemella*, *Campylobacter*, *Veillonella*, *Capsocytophaga*, *Prevotella*, *Porphyromonas*, *Rothia*, *Bifidobacterium*, and *Bacteroides* were the fifteen most abundant genera, comprising of 13.63%, 8.77%, 6.42%, 6.14%, 4.97%, 4.28%, 2.96%, 2.67%, 2.47%, 2.24%, 2.12%, 1.87%, 1.87%, 1.85%, and 1.84% sequence coverage, respectively (Figures S2A–C). Of all genera detected, 137 taxa were found common in all samples, while 26 and 29 taxa were distinctly present OSCC lesions and anatomically matched controls, respectively (Figure S2D). The shared genera among OSCC lesions and normal samples collectively represented more than 97.0% of all detected sequences in oral microbiota (Figure S2).

In order to further determine the differential presence and abundance of oral microbial community a phylogenetic tree was generated up to class level by MicrobiomeAnalyst (Figure 3C). The result demonstrated that *Rudrobacteria* class under phylum *Actinobacteria* was exclusively present in normal samples, whereas *Deinococcus* phylum was exclusively present in OSCC lesions (Figure 3C). A number of bacterial genera under the classes of *Gammaproteobacteria*, *Clostridia*, *Negativicutes*, *Bacilli*, *Actinobacteria*, *Bacteroidia* and *Fusobacteria* demonstrated an elevated abundance in both groups without significant deference in distribution between OSCC lesions and anatomically matched normal samples (Figure 3C).



**FIGURE 2** | Beta diversity analyses among cancerous and normal samples. Principal Component Analysis (PCoA) plots based on **(A)** Bray-Curtis dissimilarity and **(B)** weighted UniFrac distance matrices with respect to the bacterial abundance and composition among oral squamous cell carcinoma (OSCC) lesions and adjacent normal tissue samples. **(A)** Axis 1 (PCoA1): 19.4% of variation explained. Axis 2 (PCoA2): 16.5% of variation explained. **(B)** Axis 1 (PCoA1): 26.6% of variation explained. Axis 2 (PCoA2): 16.3% of variation explained. **(C)** The error plots identified from random forest classification analyses. Red-line indicates the overall genera present in both OSCC and normal samples, green-line indicates the distinct genera present in the normal samples and the blue-line indicates the specific genera present in the OSCC lesions.

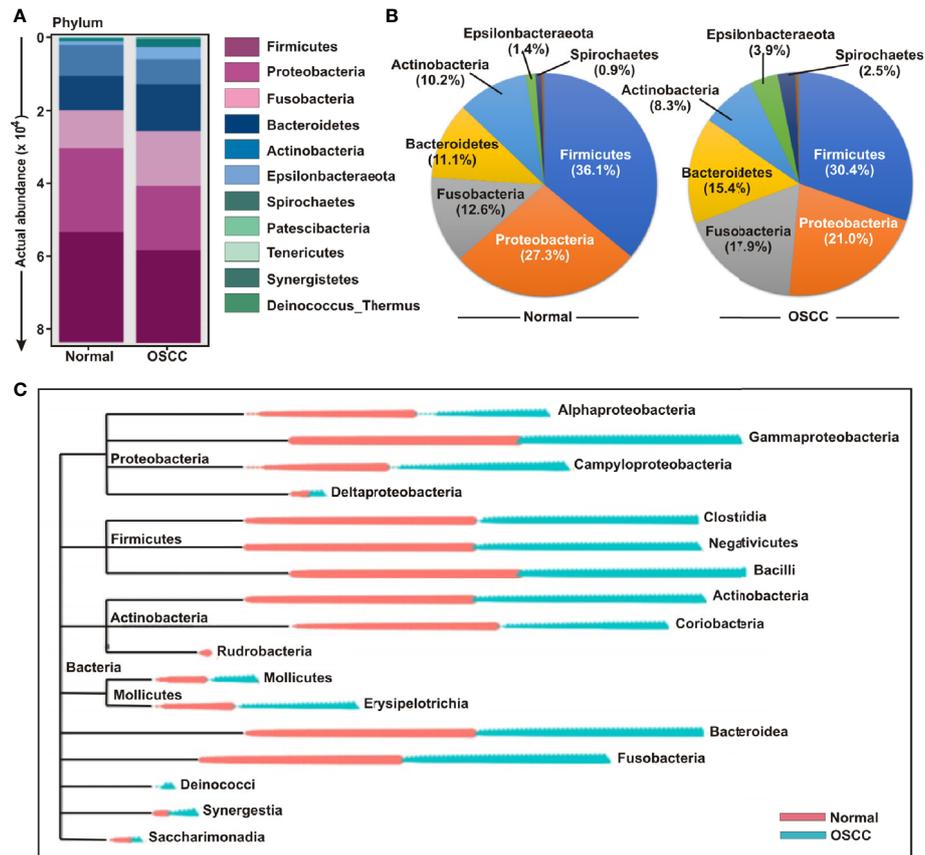
## Enrichment Analysis Identified Unique Genera for OSCC and Adjacent Normal Tissue Samples

Cladistic analysis allows for a precise definition of biological classification in which organisms are categorized in ‘clades’ (or groups) based on the most recent common ancestor and are best depicted by cladogram models indicating the relation between the different levels of clades in multiple sample groups. Identification of differential microbial ecology at phylum level would thus further facilitate to correlate their potential effect on OSCC progression. A cladogram was generated using Galaxy (42), a web-based platform for bioinformatic analysis, to visualize significantly enriched bacteria taxa identified in OSCC lesions and anatomically matched adjacent control tissue samples (**Figure 4A**). The result demonstrated that while *Bacteroidetes* phylum was phylogenetically predominant, a number of phyla including *Firmicutes* and *Actinobacteria* were depleted in the cancerous lesions as compared to the healthy controls (**Figure 4A**).

Linear Discrimination Analysis (LDA) Effect Size (LEfSe) algorithm allows identifying high dimensional biomarkers among multiple study groups (48). To identify the distinguishing taxa within OSCC lesions and matched controls, we applied LEfSe

method (**Figures 4B, C**). At the phylum level, *Bacteroidetes* and *Deinococcus* were significantly enriched in OSCC lesions, while *Proteobacteria*, *Firmicutes* and *Actinobacteria* were considerably diminished (**Figure 4B**). At the genus level, 22 taxa including *Serratia*, *Anoxybacillus*, *Stenotrophomonas*, *Sutterella*, *Actinomyces*, *Bacillus*, *Lysobacter*, *Paenibacillus*, *Ammoniphilus*, *Bifidobacterium*, *Megamonas*, *Collinsella*, *Brevibacillus*, *Megasphaera*, *Blautia*, *Methylobacterium*, *Prevotella\_9*, *Roseburia*, *Phenylobacterium*, *Pseudopropionibacterium*, *Parabacteroides*, and *Anaerobacillus* were significantly declined in the OSCC lesions as compared to the healthy controls (**Figure 4C**). In contrast, only five taxa including *Prevotella\_7*, *Corynebacterium1*, *Pseudomonas*, *Deinococcus*, and *Noviherbaspirillum* were significantly enriched in the OSCC lesions as compared to the anatomically matched control samples (**Figure 4C**). Box-Whisker dot plots along with the pair-wise genus enrichment analysis were also performed to clearly visualize the differential enrichment pattern of top five bacterial genera identified by the LEfSe analyses between OSCC lesions and contralateral anatomically matched healthy controls (**Figure S3**).

In addition, a ‘Random Forest’ algorithm was employed to further assess the diversity in bacterial community at the species level among the OSCC samples and anatomically matched



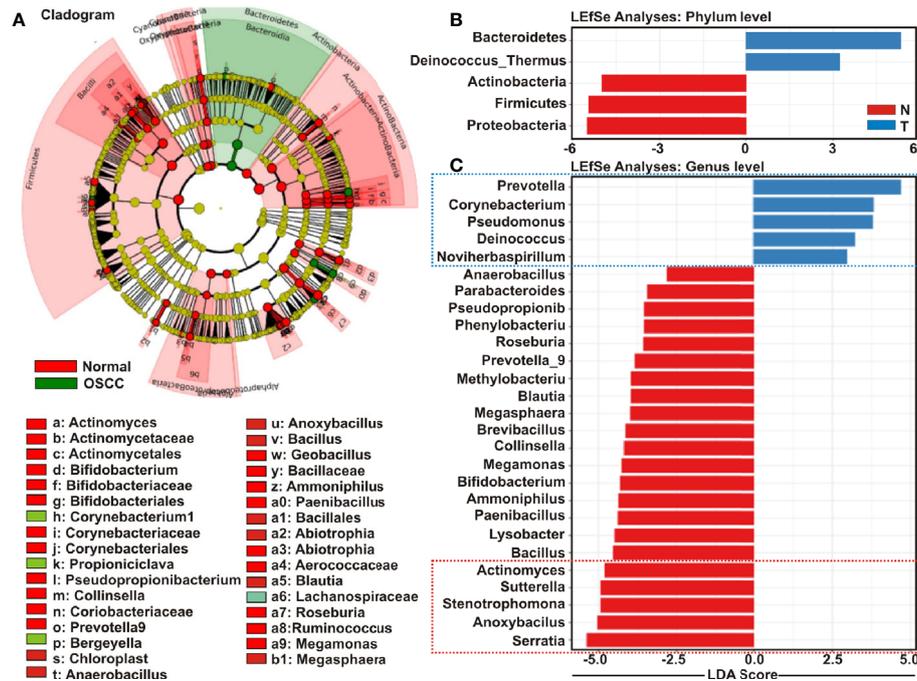
**FIGURE 3** | Composition of bacterial communities across samples at the phylum and genus levels. **(A)** Actual and **(B)** relative abundance of bacterial communities at phylum level of OSCC lesions and anatomically matched controls. **(C)** Phylogenetic tree with operational taxonomic unit (OTU) abundances between OSCC and normal samples.

healthy controls (**Figure S4A**). The decision trees extracted from the random forest classification identified distinct bacterial species in OSCC lesions when compared with the normal samples (**Figure S4A**). The results demonstrated that in the total of 50 OSCC samples, 34 samples exhibited unique and 16 samples showed overlapping species, while in normal counterparts, 30 samples exhibited unique and 20 samples showed overlapping species (**Figure S4A**). LEfSe analyses further revealed *Capnocytophaga*, unidentified *Micrococcaceae* and uncultured *Cornebacterium* 1 species were considerably enriched in OSCC lesions, 29 different species, however, mostly unidentified and uncultured species, were significantly declined as compared to the paired contralateral anatomically matched controls (**Figure S4B**). Recent studies suggested that 16S rRNA based sequencing technologies targeting one or more hypervariable regions allow reliable identification of bacterial genera, but can potentially misguide identification of bacterial species (49, 50). In agreement to this, our study also demonstrated that the sequencing depth was not sufficient to accurately identify the oral microbial composition at the species level responsible for OSCC development. Therefore, to nullify the false positives at the species level, we have limited our

analyses up to genus level for further investigation and subsequent conclusion.

## Functional Prediction of Oral Microbiome Associated With OSCC Development

To envisage oral microbial functions connected to the development of OSCC, we employed the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) (45) and accordingly Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (46) were generated specific for OSCC lesions and anatomically matched healthy controls (**Figure 5**). The LEfSe outputs demonstrated that function related to nucleotide synthesis and maintaining the fundamental functions of a cell such as pyrimidine and purine metabolism, DNA repair and recombination proteins, DNA replication, transcription machinery, amino and nucleotide sugar metabolism, protein translation related function such as ribosome, amino acid related enzymes, aminoacyl tRNA biosynthesis, peptidases, as well as peptidoglycan biosynthesis were associated with the progression of OSCC (**Figure 5A**). In contrast, parameters related to flagellar assembly, butanoate metabolism, secretion system, bacterial motility proteins, two-



**FIGURE 4** | Cladogram and enrichment analysis among oral squamous cell carcinoma (OSCC) lesions and clinically normal tissue samples. **(A)** Cladogram for phylogenetic relation of Normal and OSCC genus. Cladogram was constructed using the Linear Discrimination Analysis (LDA) Effect Size (LEfSe) method to indicate the phylogenetic distribution of bacteria that were significantly enriched in the tumor and normal groups. LDA scores showed significant bacterial differences within groups OSCC and clinically normal counterparts at the **(B)** phylum level and **(C)** genus level.

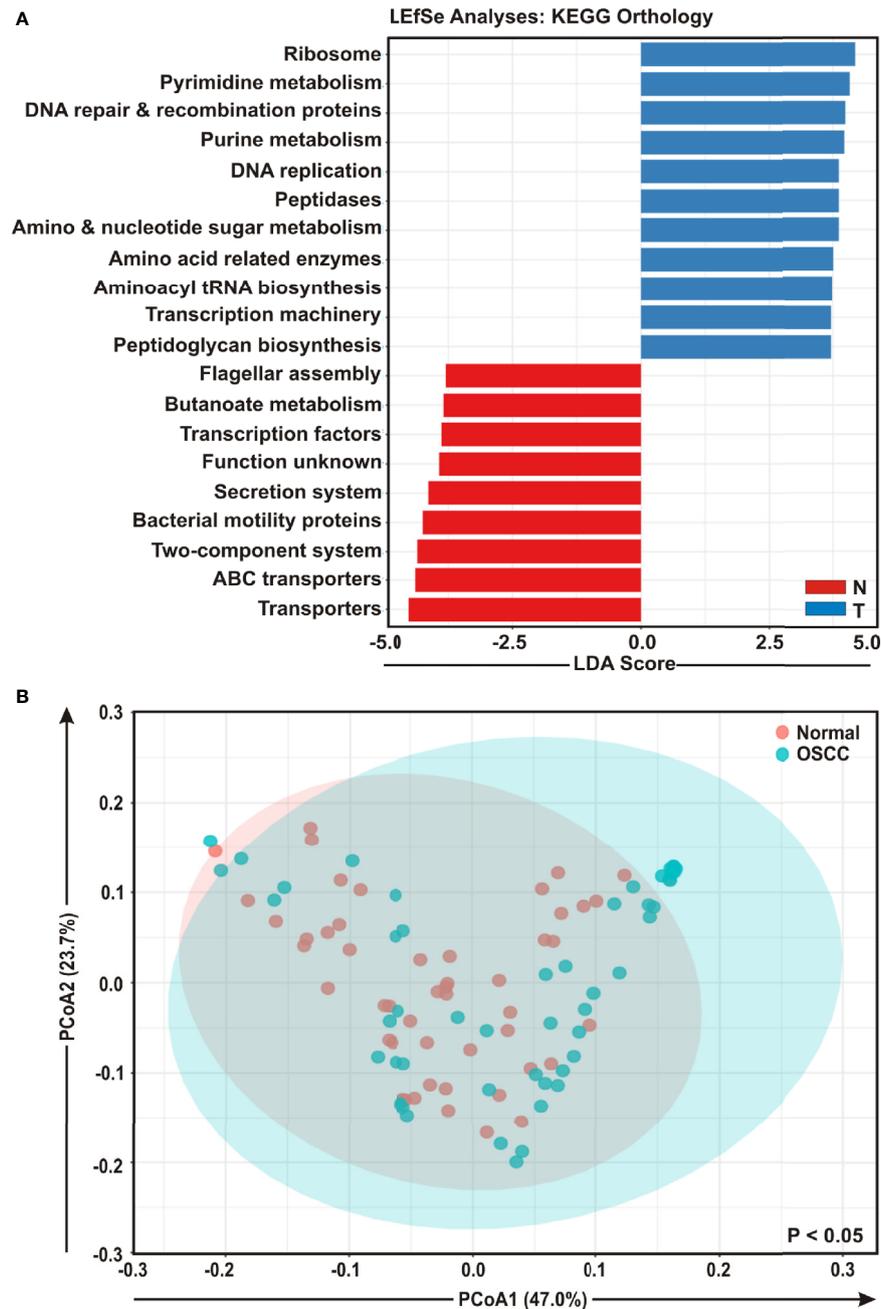
component system and ABC transporters, were inversely associated with OSCC development (**Figure 5A**). PCoA analyses also demonstrated that the predicted functions of bacterial compositions in two groups – OSCC lesions and the anatomically matched controls were significantly clustered ( $p < 0.05$ ) (**Figure 5B**).

## Quantitative PCR Profile of Oncogenic Viruses Revealed Significant Association of HPV16 With OSCC Development

Studies suggest that a number of human oncogenic viruses including human papilloma viruses (HPVs) and Epstein-Barr virus (EBV) are associated with OSCC development (13–15). To assess the potential involvement of viral etiology in our samples we designed real time PCR primers against EBV encoded EBNA3A oncogene (GeneID: 3783762) along with two high risk HPV isotypes HPV-16 encoded E2 oncogene (GeneID: 1489080) and HPV-18 encoded E6 oncogene (GeneID: 1489088) and subsequently employed in quantitative PCR (qPCR) analyses (**Figures 6A–C**). The housekeeping gene human GAPDH gene was utilized as control assuming the genomic segment bearing GAPDH gene remained unaffected in both normal and OSCC affected tissue sections. A higher negative  $-\Delta C_t$  (average GAPDH  $C_t$  value – average target primer  $C_t$  value) indicated elevated presence of the virus in the sample as

detected by specific primer set targeting specific viral gene. Our results clearly demonstrated that only HPV-16 ( $p = 0.004$ ) was significantly associated with OSCC lesions as compared to the control tissue sections (**Figure 6A**). In contrast, no significant association for both HPV-18 ( $p = 0.221$ ) and EBV ( $p = 0.326$ ) between the two sample groups was observed (**Figures 6B, C**, respectively). However,  $-\Delta C_t$  values for both HPV-18 and EBV were higher than that of HPV-16, indicating a higher prevalence in oral tissue samples (compare **Figures 5B, C** with **5A**). Altogether these oncogenic viruses might regulate the onset as well as progression of oral cavity oncogenesis and thereby demands their detection along with bacterial dysbiosis.

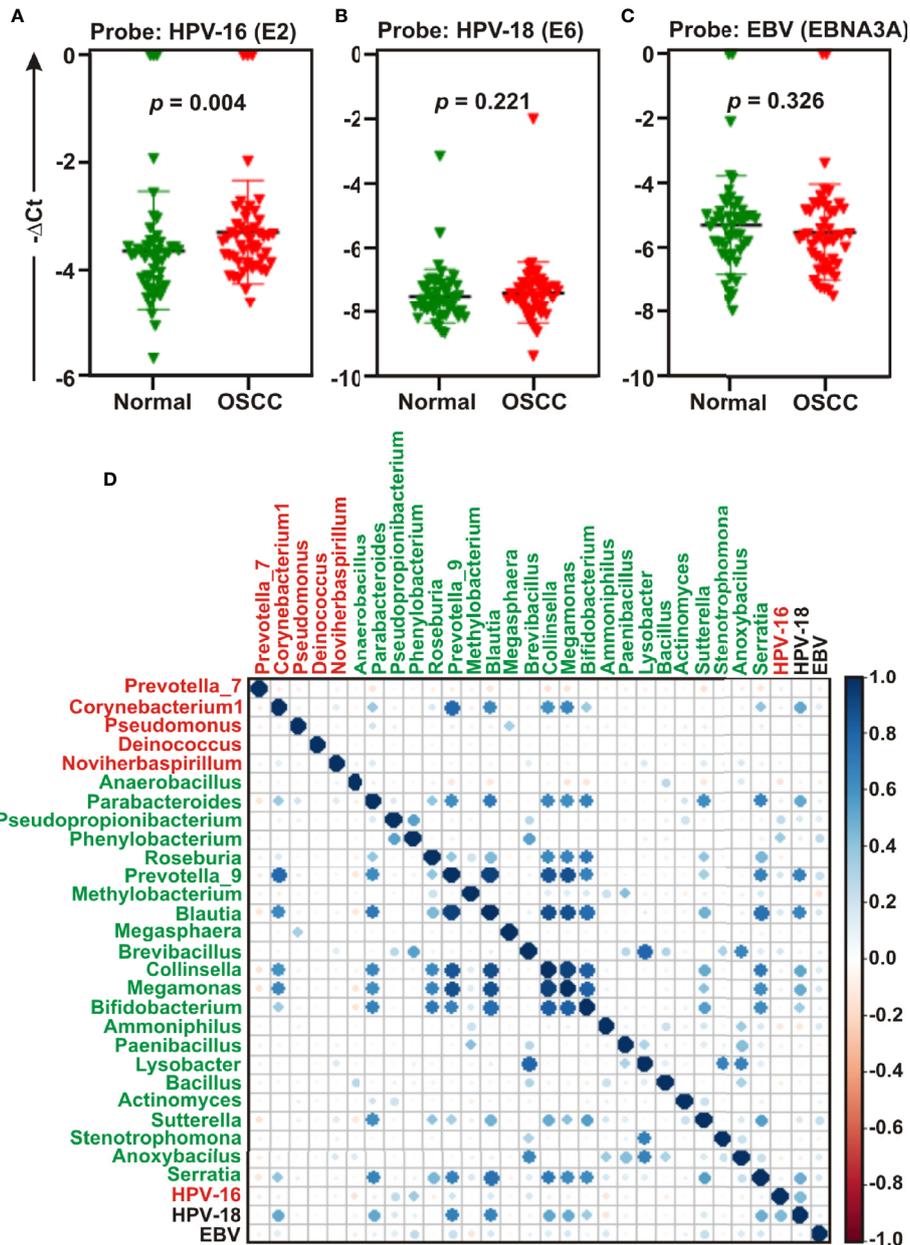
Next, the co-occurrence and co-exclusion patterns of these oncogenic viruses with the 27 most abundant bacterial genera identified in LEfSe analyses in each group of OSCC lesions and contralateral matched controls were further investigated (**Figure 6D**). Overall, there was no negative correlation found in our analyses. In matched normal controls, *Prevotella\_9* was found to be positively correlated with a number of bacterial genera. For example, *Prevotella\_9* and *Blautia* were the most positively correlated ( $\rho = 0.926$ ), followed by *Megamonas* ( $\rho = 0.872$ ), *Collinsella* ( $\rho = 0.852$ ), *Serratia* ( $\rho = 0.679$ ), *Bifidobacterium* ( $\rho = 0.673$ ), and *Parabacteroides* ( $\rho = 0.626$ ). HPV-18 demonstrated moderate positive correlation with most of these genera – including *Prevotella\_9* ( $\rho = 0.683$ ), *Blautia* ( $\rho = 0.660$ ), *Megamonas* ( $\rho = 0.509$ ), *Collinsella* ( $\rho = 0.527$ ), and *Parabacteroides*



**FIGURE 5** | Functional analyses of identified microbial compositions in cancerous lesions and clinically normal samples. **(A)** Linear Discrimination Analysis (LDA) scores predicting gene function enriched among two different groups of oral squamous cell carcinoma (OSCC) lesions and normal samples using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUST) according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. **(B)** Principal Component Analysis (PCoA) of bacterial functions associated OSCC lesions and contralateral matched controls. Axis 1 (PCoA1): 47.0% of variation explained. Axis 2 (PCoA2): 23.7% of variation explained.

( $\rho = 0.514$ ) (**Figure 6D**, **Table S3**). In contrast, HPV-16 demonstrated no positive correlation with any bacterial genera identified in cancerous lesions (**Figure 6D**). Although *Corynebacterium1* was significantly associated with OSCC lesions, it demonstrated positive correlation with several bacterial

taxa abundantly enriched in normal samples (**Figure 6D**). Among these, the most positively correlated genera were *Corynebacterium1* and *Prevotella\_9* ( $\rho = 0.785$ ) followed by *Megamonas* ( $\rho = 0.657$ ), *Blautia* ( $\rho = 0.643$ ), and *Collinsella* ( $\rho = 0.601$ ) (**Figure 6D**, **Table S3**).



**FIGURE 6** | Quantitative PCR (qPCR) and co-occurrence analysis of human oncogenic viruses with identified bacterial genera in normal and oral squamous cell carcinoma (OSCC) samples. Comparative qPCR data of (A, B) high risk human papilloma viruses (HPVs) - (A) HPV-16 (B) HPV-18 and (C) Epstein-Barr virus (EBV). PCR calculation was performed by  $-\Delta\text{Ct}$  method to quantify relative abundance of each tumor virus using human genomic GAPDH as control. The  $-\Delta\text{Ct}$  values of each sample were plotted using GraphPad Prism 8.0.1. (D) Pearson correlations among human oncogenic viruses and the top 27 most abundant bacterial genera identified by Linear Discrimination Analysis (LDA) Effect Size (LEfSe) analyses were calculated and analyzed. Correlation values range from  $-1.00$  (red) to  $1.00$  (blue).

## DISCUSSION

In spite of the highest oral cancer incidence, accounting to 30% of all cancers in India (6, 7), so far there are no reports describing changes of oral microbiome in OSCC among Indian patients. The purpose of the current investigation was to profile the dysbiosis of oral microbiota between OSCC lesions and contralateral anatomically matched control tissue samples

prospectively collected from fifty patients of eastern region of India. In agreement with Guerrero-Preston et al. study (51), we also observed a significant loss in richness and diversity of oral bacterial communities in OSCC lesions compared to matched controls. However, several reports revealed enhanced diversity of bacterial communities in OSCC samples (31, 52, 53). Nevertheless, dysbiosis of oral microbiome appears to be strongly associated with OSCC development. Overall, the

results demonstrated that *Prevotella*, *Corynebacterium*, *Pseudomonas*, *Deinococcus*, and *Noviherbaspirillum* genera were significantly enriched, while genera including *Actinomyces*, *Sutterella*, *Stenotrophomonas*, *Anoxybacillus*, and *Serratia* were depleted in the OSCC lesions as compared to the matched healthy controls.

Previously, several models of microbial infection and potential oral microbiome signature link to the pathology of a number of oral diseases including cancer have been established. For example, certain oral bacterial pathogens, including *Porphyromonas gingivalis* and *Fusobacterium nucleatum* have been reported to disrupt the equilibrium of oral microbiome and along with deregulated immune response eventually initiate periodontal diseases (periodontitis) (54–56). These well studied periodontal organisms subsequently prompted researchers to further investigate the precise role of dysbiotic oral microbiota in developing oral cancer (57, 58). In general, our results agreed with the previously published data of enriched and depleted microbes associated with the OSCC development. Overall, five of the most abundant phyla including *Proteobacteria*, *Firmicutes*, *Actinobacteria*, *Bacteroidetes*, and *Fusobacteria* identified in our study were consistent with those found in previous studies. However, the less abundant phyla including *Tenericutes*, *Deinococcus*, and *Patescibacteria* detected were significantly varied among multiple studies. In addition, in line with the previous studies (29, 52, 53) *Firmicutes* was also found as the most abundant phylum in overall oral microbiome in our study. Of the significantly elevated genera in cancerous lesions, *Prevotella* and *Pseudomonas* were previously shown to be highly abundant in both periodontitis and OSCC samples when compared to healthy controls (32, 53, 59). Importantly, periodontitis has been suggested as a self-governing risk factor for OSCC development (27). Interestingly, in contrast to our finding, *Corynebacterium* was previously found to be decreased in oral cavity cancer (OCC) and oropharyngeal cancers (OPC) (60). Our results indicated presence of a unique genus *Deinococcus* although relatively less abundantly only in cancerous lesions and could not be detected in control tissue sections. Since in our study design, paired OSCC lesion and control tissue samples were obtained from single patient, it nullified the possibilities of inter-individual variation. Thus even small differences of bacterial communities among two these groups would represent significance in OSCC development. Species of the *Deinococcus* genus are recognized for their extreme resistance to ionizing radiation and oxidative stress and other damaging conditions (61). Although a number of earlier studies indicated the presence of *Deinococcus* genus (62, 63), the precise role of the members of this genus in OSCC is yet to be defined.

In our study, although *Fusobacteria* was identified as one of the most abundant phyla in overall oral bacteriome, its abundance showed no significant difference between OSCC lesions and normal tissue samples. This is in contrast to a number of recent reports which demonstrated significant abundance of several members of *Fusobacterium* in OSCC lesions when compared to normal samples (31, 32). Mager et al. detected *F. periodonticum* in the saliva sample from OSCC patients using specific bacteria probes,

but its abundance showed no significant difference between OSCC-positive and OSCC-free patients (30). Yang et al. determined significant elevation of *F. periodonticum* species in OSCC lesions, whereas no significant difference was observed in case of *F. nucleatum* between tumor and normal samples (52). In contrast, Al-Hebshi et al. indicated that *F. nucleatum* was the most significantly enriched species in OSCC lesions as compared to the control normal tissues (32). The diverse presence of different members of *Fusobacterium* species identified in OSCC samples in multiple studies possibly arose due to varied sample types as well as subjects recruited of different ethnicity across the world. In addition, *Fusobacterium nucleatum* was also identified as one of the highly enriched bacterial species in colorectal cancer (64). Moreover, Komiya et al. showed that patients with colorectal cancer (CRC) have identical strains of *Fusobacterium nucleatum* in their CRC tissue section and oral cavity (65). Given the importance of *Fusobacterium* in various human cancers, further in depth investigation is required to verify *Fusobacterium* association with OSCC in Indian scenario with larger patients sample size.

Nucleotide metabolism is an important pathway that provides purine and pyrimidine molecules for DNA replication, RNA biogenesis, as well as cell bioenergetics. Increased nucleotide metabolism supports uncontrolled proliferation of cancer cells and thus represents a hallmark of cancer (66). Apart from nucleotide metabolisms, several critical pathways like DNA repair, recombination, protein synthesis and transcription machineries are frequently altered in tumor cells (67–69). Moreover, inhibitors that specifically blocks DNA replication and induce DNA damages have been widely used as chemotherapeutic agents against numerous cancers (70). In agreement to this, our PICRUSt analyses showed that function related to nucleotide metabolisms including both purine and pyrimidine synthesis as well as basic cell functions like DNA repair and replication and functions related to mRNA translation including ribosome, amino acid related enzymes, aminoacyl tRNA biosynthesis and peptidases were significantly linked to OSCC development. Although Yost et al. using metatranscriptomic analyses suggested importance of these pathways for OSCC development (71), so far there are no robust studies that directly linked microbes with these pathways in a tumor microenvironment. Moreover, in contrast to our study, Yang et al. demonstrated that parameters related to protein and amino acids metabolisms were inversely associated with OSCC progression from stage 1 to stage 4 patients (52). Previously a number of reports demonstrated that pathways related to bacterial chemotaxis and flagellar assembly were remarkably enriched in the OSCC group (32, 59). However, in contrast, our study showed that pathways related to flagellar assembly and bacterial motility proteins were inversely associated with the OSCC development.

A growing body of evidence suggested a potential association of several human tumor viruses with oral cancers (13–15, 72). For example, while low risk HPV subtypes including HPV-6 and HPV-11 are associated with a variety of oral benign papillomatous lesions such as oral squamous papilloma, oral verruca vulgaris, oral condyloma accuminatum and focal epithelial hyperplasia, high risk HPV subtypes including HPV-16 and HPV-18 are associated with malignant lesions (72–75). The transformation

of normal oral mucosa in OSCC is potentially linked to precancerous lesions, such as OLP (76, 77). Although the precise role of viral mediated malignant transformation of precancerous lesions is not clear, HPV infection is significantly associated with OLP (72, 78). Overall, previous studies suggested that HPV-16 is the most frequently detected HPV subtype in oral cancers (79) and accordingly in 2012 the International Agency of Research of Cancer (IARC) acknowledged the significant association of HPV-16 high risk group with oral cancer development (14). In agreement to this, our results also demonstrated that HPV-16 but not HPV-18 was significantly associated with OSCC lesions as compared to anatomically matched control tissue sections.

In sum, using a carefully controlled patients' cohort, herein we identified specific microbial signature associated with OSSCC development. However, the current study has several limitations including constraints associated with the 16S rRNA gene amplification based sequencing technologies (49, 50). Recent studies suggested although more than 99% of sequencing reads could be correctly classified at the genus level, a significant proportion at the species level might be misclassified during identification of bacterial populations by targeting various variable regions of the 16S rRNA gene. In agreement to this, our study also showed that the sequencing depth was not adequate to precisely classify the oral microbiota at the species level and thus, in order to nullify false positives, we restricted our analyses up to genus level. Another limitation of this study was relatively smaller sample size. In future, to validate the results a larger sample size with distinct cancer stages among population in different regions and different socioeconomic background would be highly preferable. We additionally lacked information on the involvement of other organisms particularly fungus and the association of different viruses with OSCC development. Strategies such as whole genome sequencing and metabolomics would allow in identification of changes of overall oral microbiota and their metabolites during OSCC development. Moreover, whole genome shotgun sequencing (metagenomics) would further validate the functional inferences from 16S rRNA amplicon sequences obtained using PICRUSt. Since the current study was conducted using tissue biopsy samples, it would be interesting to investigate whether the results could be extended in a non-invasive method by utilizing saliva samples. Altogether, longitudinal research activities are greatly demanded to explore the functional implications of the oral microbiota in terms of diagnosis and risk assessment of OSCC development, as well as potential expansion of current therapeutic strategies to restore the health of the oral ecosystem.

## DATA AVAILABILITY STATEMENT

The 16S rRNA amplicon sequencing data from this study have been deposited in the NCBI BioProject under accession number PRJNA666746.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Ethics Committee for Human

Research, Indian Statistical Institute, Kolkata, India. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

PS and AS wrote the main manuscript text. PS, SM, SL, RC, and AS performed the bioinformatic analysis. PS and SB performed the experiments. JR recruited the patients and conducted the histopathological studies. SD and RC collected the samples and performed sampling. AS conceived, designed, and successfully sought funding for the study. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by grants from DBT/Wellcome Trust India Alliance Intermediate Fellowship research grant [IA/I/14/2/501537] and Science & Technology and Biotechnology, Govt. of West Bengal [1798 (Sanc.)/ST/P/S&T/9G-5/2019] to A.S. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## ACKNOWLEDGMENTS

The authors would like to thank all patients for making this study possible. The authors thank all members of both AS and RC laboratories for valuable suggestions and discussions of the experiment. AS is Wellcome Trust/DBT India Alliance Intermediate Fellow. PS and SL are recipients of CSIR-NET Research Fellowships and SM is a recipient of UGC-NET Research Fellowships.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2021.614448/full#supplementary-material>

**Supplementary Figure 1 |** Preliminary quantitative PCR (qPCR) analyses of four OSCC and anatomically matched normal samples. PCR calculation was performed by  $-\Delta\Delta C_T$  method to quantify relative abundance of overall bacteria using primers against the conserved region of 16S rRNA gene and human genomic GAPDH gene segment as control. The  $-\Delta C_T$  values of each sample were plotted using GraphPad Prism 8.0.1.

**Supplementary Figure 2 |** Differentially abundant genera in OSCC lesions and anatomically matched normal samples. Relative abundance (%) of the taxa at the genus level in (A) all samples, (B) anatomically matched control tissues and (C) OSCC lesions. (D) Venn diagram depicts distinct and overlapped genera among OSCC and normal samples.

**Supplementary Figure 3 |** Top five taxa at genus level identified in the LEfSe analysis among OSCC lesions and healthy matched controls. Box Whisker Plot (top) and Pair-wise (bottom) genus enrichment analysis of top five bacterial genera

identified in LEFSe analysis as described in **Figure 4** among **(A)** normal samples and **(B)** OSCC lesions.

**Supplementary Figure 4 |** Bacterial composition at the species level between paired OSCC lesions and contralateral normal tissue samples. **(A)** The error plots identified from random forest classification analyses. Red-line indicates the overall species present in both OSCC and normal samples, green-line indicates the distinct species present in the normal samples and the blue-line indicates the specific species present in the OSCC lesions. **(B)** Linear discriminant analysis effect size (LEFSe) analysis

demonstrating differential bacterial composition at the species level between the OSCC lesions and anatomically matched healthy controls.

**Supplementary Table 1 |** Clinical characteristics of patients' samples.

**Supplementary Table 2 |** Real time PCR primers.

**Supplementary Table 3 |** Pearson correlations among HPV-16, HPV-18, EBV and the top 27 most abundant bacterial genera identified by LEFSe analyses.

## REFERENCES

- Vigneswaran N, Williams MD. Epidemiologic trends in head and neck cancer and aids in diagnosis. *Oral Maxillofac Surg Clin North Am* (2014) 26(2):123–41. doi: 10.1016/j.coms.2014.01.001
- Rischin D, Ferris RL, Le QT. Overview of Advances in Head and Neck Cancer. *J Clin Oncol* (2015) 33(29):3225–6. doi: 10.1200/JCO.2015.63.6761
- Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* (2005) 55(2):74–108. doi: 10.3322/canjclin.55.2.74
- Warnakulasuriya S. Causes of oral cancer—an appraisal of controversies. *Br Dent J* (2009) 207(10):471–5. doi: 10.1038/sj.bdj.2009.1009
- Nunez-Gonzalez S, Delgado-Ron JA, Gault C, Simancas-Racines D. Trends and Spatial Patterns of Oral Cancer Mortality in Ecuador, 2001–2016. *Int J Dent* (2018) 2018:6086595. doi: 10.1155/2018/6086595
- Basu B, Chakraborty J, Chandra A, Katarkar A, Baldevbhai JRK, Dhar Chowdhury D, et al. Genome-wide DNA methylation profile identified a unique set of differentially methylated immune genes in oral squamous cell carcinoma patients in India. *Clin Epigenet* (2017) 9:13. doi: 10.1186/s13148-017-0314-x314
- Byakodi R, Byakodi S, Hiremath S, Byakodi J, Adaki S, Marathe K, et al. Oral cancer in India: an epidemiologic and clinical review. *J Community Health* (2012) 37(2):316–9. doi: 10.1007/s10900-011-9447-6
- Tajmirriahi N, Razavi SM, Shirani S, Homayooni S, Gasemzadeh G. Evaluation of metastasis and 5-year survival in oral squamous cell carcinoma patients in Isfahan (2001–2015). *Dent Res J (Isfahan)* (2019) 16(2):117–21. doi: 10.4103/1735-3327.250974
- Geum DH, Roh YC, Yoon SY, Kim HG, Lee JH, Song JM, et al. The impact factors on 5-year survival rate in patients operated with oral cancer. *J Korean Assoc Oral Maxillofac Surg* (2013) 39(5):207–16. doi: 10.5125/jkaoms.2013.39.5.207
- Rogers SN, Brown JS, Woolgar JA, Lowe D, Magennis P, Shaw RJ, et al. Survival following primary surgery for oral cancer. *Oral Oncol* (2009) 45(3):201–11. doi: 10.1016/j.oraloncology.2008.05.008
- Parra C, Jodar-Sanchez F, Jimenez-Hernandez MD, Vigil E, Palomino-Garcia A, Moniche-Alvarez F, et al. Development, Implementation, and Evaluation of a Telemedicine Service for the Treatment of Acute Stroke Patients: TeleStroke. *Interact J Med Res* (2012) 1(2):e15. doi: 10.2196/ijmr.2163
- Bagan J, Sarrion G, Jimenez Y. Oral cancer: clinical features. *Oral Oncol* (2010) 46(6):414–7. doi: 10.1016/j.oraloncology.2010.03.009
- Gupta K, Metgud R. Evidences suggesting involvement of viruses in oral squamous cell carcinoma. *Patholog Res Int* (2014) 2013:642496. doi: 10.1155/2013/642496
- Kim SM. Human papilloma virus in oral cancer. *J Korean Assoc Oral Maxillofac Surg* (2016) 42(6):327–36. doi: 10.5125/jkaoms.2016.42.6.327
- She Y, Nong X, Zhang M, Wang M. Epstein-Barr virus infection and oral squamous cell carcinoma risk: A meta-analysis. *PloS One* (2017) 12(10):e0186860. doi: 10.1371/journal.pone.0186860PONE-D-17-18573
- Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. *Nat Rev Genet* (2012) 13(4):260–70. doi: 10.1038/nrg3182
- Elinav E, Garrett WS, Trinchieri G, Wargo J. The cancer microbiome. *Nat Rev Cancer* (2019) 19(7):371–6. doi: 10.1038/s41568-019-0155-310.1038/s41568-019-0155-3
- Malla MA, Dubey A, Kumar A, Yadav S, Hashem A, Abd Allah EF. Exploring the Human Microbiome: The Potential Future Role of Next-Generation Sequencing in Disease Diagnosis and Treatment. *Front Immunol* (2019) 9:2868:2868. doi: 10.3389/fimmu.2018.02868
- Deo PN, Deshmukh R. Oral microbiome: Unveiling the fundamentals. *J Oral Maxillofac Pathol* (2019) 23(1):122–8. doi: 10.4103/jomfp.JOMFP\_304\_18JOMFP-23-122
- Beck JD, Offenbacher S. Systemic effects of periodontitis: epidemiology of periodontal disease and cardiovascular disease. *J Periodontol* (2005) 76(11 Suppl):2089–100. doi: 10.1902/jop.2005.76.11-S.2089
- Bingham CO, Moni M. Periodontal disease and rheumatoid arthritis: the evidence accumulates for complex pathobiologic interactions. *Curr Opin Rheumatol* (2013) 25(3):345–53. doi: 10.1097/BOR.0b013e32835fb8ec
- Daniel R, Gokulanathan S, Shanmugasundaram N, Lakshmigandhan M, Kavin T. Diabetes and periodontal disease. *J Pharm Bioallied Sci* (2012) 4(Suppl 2):S280–2. doi: 10.4103/0975-7406.100251JPBS-4-280
- Walia M, Saini N. Relationship between periodontal diseases and preterm birth: Recent epidemiological and biological data. *Int J Appl Basic Med Res* (2015) 5(1):2–6. doi: 10.4103/2229-516X.149217IJABMR-5-2
- Curtis MA, Zenobia C, Darveau RP. The relationship of the oral microbiota to periodontal health and disease. *Cell Host Microbe* (2011) 10(4):302–6. doi: 10.1016/j.chom.2011.09.008
- Preza D, Olsen I, Aas JA, Willumsen T, Grinde B, Paster BJ. Bacterial profiles of root caries in elderly patients. *J Clin Microbiol* (2008) 46(6):2015–21. doi: 10.1128/JCM.02411-07
- Chhour KL, Nadkarni MA, Byun R, Martin FE, Jacques NA, Hunter N. Molecular analysis of microbial diversity in advanced caries. *J Clin Microbiol* (2005) 43(2):843–9. doi: 10.1128/JCM.43.2.843-849.2005
- Tezal M, Sullivan MA, Hyland A, Marshall JR, Stoler D, Reid ME, et al. Chronic periodontitis and the incidence of head and neck squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* (2009) 18(9):2406–12. doi: 10.1158/1055-9965.EPI-09-0334
- Karpinski TM. Role of Oral Microbiota in Cancer Development. *Microorganisms* (2019) 7(1):20. doi: 10.3390/microorganisms7010020
- 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(6):e98741. doi: 10.1371/journal.pone.0098741PONE-D-13-34607
- 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. doi: 10.1186/1479-5876-3-27
- 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(1):11773. doi: 10.1038/s41598-017-11779-910.1038/s41598-017-11779-9
- Al-Hebshi NN, Nasher AT, Maryoud MY, Homeida HE, Chen T, Idris AM, et al. Inflammatory bacterium featuring *Fusobacterium nucleatum* and *Pseudomonas aeruginosa* identified in association with oral squamous cell carcinoma. *Sci Rep* (2017) 7(1):1834. doi: 10.1038/s41598-017-02079-310.1038/s41598-017-02079-3
- Lee WH, Chen HM, Yang SF, Liang C, Peng CY, Lin FM, et al. Bacterial alterations in salivary microbiota and their association in oral cancer. *Sci Rep* (2017) 7(1):16540. doi: 10.1038/s41598-017-16418-x10.1038/s41598-017-16418-x
- Bornigen D, Ren B, Pickard R, Li J, Ozer E, Hartmann EM, et al. Alterations in oral bacterial communities are associated with risk factors for oral and oropharyngeal cancer. *Sci Rep* (2017) 7(1):17686. doi: 10.1038/s41598-017-17795-z10.1038/s41598-017-17795-z
- Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* (2010) 26(19):2460–1. doi: 10.1093/bioinformatics/btq461

36. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* (2011) 27(16):2194–200. doi: 10.1093/bioinformatics/btr381
37. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* (2013) 10(10):996–8. doi: 10.1038/nmeth.2604
38. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* (2010) 7(5):335–6. doi: 10.1038/nmeth.f.303
39. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* (2007) 73(16):5261–7. doi: 10.1128/AEM.00062-07
40. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* (2014) 42(Database issue):D633–42. doi: 10.1093/nar/gkt1244
41. Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res* (2017) 45(W1):W180–W8. doi: 10.1093/nar/gkx295
42. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* (2018) 46(W1):W537–W44. doi: 10.1093/nar/gky379
43. Basu S, Kumbier K, Brown JB, Yu B. Iterative random forests to discover predictive and stable high-order interactions. *Proc Natl Acad Sci U.S.A.* (2018) 115(8):1943–8. doi: 10.1073/pnas.1711236115
44. Kelly BJ, Gross R, Bittinger K, Sherrill-Mix S, Lewis JD, Collman RG, et al. Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA. *Bioinformatics* (2015) 31(15):2461–8. doi: 10.1093/bioinformatics/btv183
45. Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* (2013) 31(9):814–21. doi: 10.1038/nbt.2676
46. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* (2000) 28(1):27–30. doi: 10.1093/nar/28.1.27
47. Goodrich JK, Di Rienzi SC, Poole AC, Koren O, Walters WA, Caporaso JG, et al. Conducting a microbiome study. *Cell* (2014) 158(2):250–62. doi: 10.1016/j.cell.2014.06.037
48. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol* (2011) 12(6):R60. doi: 10.1186/gb-2011-12-6-r60
49. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* (2019) 10(1):5029. doi: 10.1038/s41467-019-13036-110.1038/s41467-019-13036-1
50. Winand R, Bogaerts B, Hoffman S, Lefevre L, Delvoeye M, Braekel JV, et al. Targeting the 16S rRNA gene for Bacterial Identification in Complex Mixed Samples: Comparative Evaluation of Second (Illumina) and Third (Oxford Nanopore Technologies) Generation Sequencing Technologies. *Int J Mol Sci* (2019) 21(1):298. doi: 10.3390/ijms21010298
51. Guerrero-Preston R, Godoy-Vitorino F, Jedlicka A, Rodriguez-Hilario A, Gonzalez 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(32):51320–34. doi: 10.18632/oncotarget.9710
52. Yang CY, Yeh YM, Yu HY, Chin CY, Hsu CW, Liu H, et al. Oral Microbiota Community Dynamics Associated With Oral Squamous Cell Carcinoma Staging. *Front Microbiol* (2018) 9:862:862. doi: 10.3389/fmicb.2018.00862
53. Su SC, Chang LC, Huang HD, Peng CY, Chuang CY, Chen YT, et al. Oral microbial dysbiosis and its performance in predicting oral cancer. *Carcinogenesis* (2020) 41:4602. doi: 10.1093/carcin/bgaa062
54. Nagy KN, Sonkodi I, Szoke I, Nagy E, Newman HN. The microflora associated with human oral carcinomas. *Oral Oncol* (1998) 34(4):304–8. doi: 10.1016/S1368-8375(98)80012-2
55. Bolz J, Dosa E, Schubert J, Eckert AW. Bacterial colonization of microbial biofilms in oral squamous cell carcinoma. *Clin Oral Investig* (2014) 18(2):409–14. doi: 10.1007/s00784-013-1007-2
56. de Andrade KQ, Almeida-da-Silva CLC, Coutinho-Silva R. Immunological Pathways Triggered by *Porphyromonas gingivalis* and *Fusobacterium nucleatum*: Therapeutic Possibilities? *Mediators Inflammation* (2019) 2019:7241312. doi: 10.1155/2019/7241312
57. Whitmore SE, Lamont RJ. Oral bacteria and cancer. *PLoS Pathog* (2014) 10(3):e1003933. doi: 10.1371/journal.ppat.1003933
58. Hajishengallis G, Lamont RJ. Beyond the red complex and into more complexity: the polymicrobial synergy and dysbiosis (PSD) model of periodontal disease etiology. *Mol Oral Microbiol* (2012) 27(6):409–19. doi: 10.1111/j.2041-1014.2012.00663.x
59. Zhang L, Liu Y, Zheng HJ, Zhang CP. The Oral Microbiota May Have Influence on Oral Cancer. *Front Cell Infect Microbiol* (2019) 9:476:476. doi: 10.3389/fcimb.2019.00476
60. 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) 8:267:267. doi: 10.3389/fcimb.2018.00267
61. Pavlopoulou A, Savva GD, Louka M, Bagos PG, Vorgias CE, Michalopoulos I, et al. Unraveling the mechanisms of extreme radioresistance in prokaryotes: Lessons from nature. *Mutat Res Rev Mutat Res* (2016) 767:92–107. doi: 10.1016/j.mrrev.2015.10.001
62. Zhang X, 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) 10:1439:1439. doi: 10.3389/fmicb.2019.01439
63. Chang C, Geng F, Shi X, Li Y, Zhang X, Zhao X, et al. The prevalence rate of periodontal pathogens and its association with oral squamous cell carcinoma. *Appl Microbiol Biotechnol* (2018) 103(3):1393–404. doi: 10.1007/s00253-018-9475-610.1007/s00253-018-9475-6
64. Yu T, Guo F, Yu Y, Sun T, Ma D, Han J, et al. *Fusobacterium nucleatum* Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. *Cell* (2017) 170(3):548–63 e16. doi: 10.1016/j.cell.2017.07.008
65. Komiya Y, Shimomura Y, Higurashi T, Sugi Y, Arimoto J, Umezawa S, et al. Patients with colorectal cancer have identical strains of *Fusobacterium nucleatum* in their colorectal cancer and oral cavity. *Gut* (2018) 68(7):1335–7. doi: 10.1136/gutjnl-2018-316661
66. Siddiqui A, Ceppi P. A non-proliferative role of pyrimidine metabolism in cancer. *Mol Metab* (2020) 35:100962. doi: 10.1016/j.molmet.2020.02.005
67. Dietlein F, Thelen L, Reinhardt HC. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. *Trends Genet* (2014) 30(8):326–39. doi: 10.1016/j.tig.2014.06.003
68. Bradner JE, Hnisz D, Young RA. Transcriptional Addiction in Cancer. *Cell* (2017) 168(4):629–43. doi: 10.1016/j.cell.2016.12.013
69. Lee LJ, Papadopoli D, Jewer M, Del Rincon S, Topisirovic I, Lawrence MG, et al. Cancer Plasticity: The Role of mRNA Translation. *Trends Cancer* (2020) 7(2):134–45. doi: 10.1016/j.trecan.2020.09.005
70. Berdis AJ. Inhibiting DNA Polymerases as a Therapeutic Intervention against Cancer. *Front Mol Biosci* (2017) 4:78:78. doi: 10.3389/fmolb.2017.00078
71. Yost S, Stashenko P, Choi Y, Kukuruzinska M, Genco CA, Salama A, et al. Increased virulence of the oral microbiome in oral squamous cell carcinoma revealed by metatranscriptome analyses. *Int J Oral Sci* (2018) 10(4):32. doi: 10.1038/s41368-018-0037-710.1038/s41368-018-0037-7
72. Miller CS, Johnstone BM. Human papillomavirus as a risk factor for oral squamous cell carcinoma: a meta-analysis, 1982–1997. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* (2001) 91(6):622–35. doi: 10.1067/moe.2001.115392
73. Prabhu SR, Wilson DF. Human papillomavirus and oral disease - emerging evidence: a review. *Aust Dent J* (2013) 58(1):2–10; quiz 125. doi: 10.1111/adj.12020
74. Gupta S. Role of human papillomavirus in oral squamous cell carcinoma and oral potentially malignant disorders: A review of the literature. *Indian J Dent* (2015) 6(2):91–8. doi: 10.4103/0975-962X.155877
75. Feller L, Khammissa RA, Wood NH, Lemmer J. Epithelial maturation and molecular biology of oral HPV. *Infect Agent Cancer* (2009) 4:16. doi: 10.1186/1750-9378-4-16
76. Scheifele C, Reichart PA. [Oral leukoplakia in manifest squamous epithelial carcinoma. A clinical prospective study of 101 patients]. *Mund Kiefer Gesichtschir* (1998) 2(6):326–30. doi: 10.1007/s100060050081

77. Sudbo J, Reith A. Which putatively pre-malignant oral lesions become oral cancers? Clinical relevance of early targeting of high-risk individuals. *J Oral Pathol Med* (2003) 32(2):63–70. doi: 10.1034/j.1600-0714.2003.00054.x
78. van der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management. *Oral Oncol* (2010) 46(6):423–5. doi: 10.1016/j.oraloncology.2010.02.016
79. Licitra L, Perrone F, Bossi P, Suardi S, Mariani L, Artusi R, et al. High-risk human papillomavirus affects prognosis in patients with surgically treated oropharyngeal squamous cell carcinoma. *J Clin Oncol* (2006) 24(36):5630–6. doi: 10.1200/JCO.2005.04.6136

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Sarkar, Malik, Laha, Das, Bunk, Ray, Chatterjee and Saha. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.