

Research Article

Dysbiosis and Variation in Predicted Functions of the Granulation Tissue Microbiome in HPV Positive and Negative Severe Chronic Periodontitis

Rebecca Chowdhry,¹ Neetu Singh ,² Dinesh Kumar Sahu ,² Ratnesh Kumar Tripathi ,³ Archana Mishra,⁴ Anjana Singh,⁵ Indrashis Mukerjee,⁶ Nand Lal,¹ Madan Lal Brahma Bhatt,⁷ and Ravi Kant⁸

¹Department of Periodontology, King George's Medical University, Lucknow 226 003, India

²Department of Molecular Biology, Center for Advance Research, King George's Medical University, Lucknow 226 003, India

³Imperial Life Sciences, 463 Phase City 2, Sector 37, Gurgaon, Haryana 122001, India

⁴Department of Thoracic Surgery, King George's Medical University, Lucknow 226 003, India

⁵Department of Pulmonary and Critical Care Medicine, King George's Medical University, Lucknow 226 003, India

⁶Department of Pharmacology, King George's Medical University, Lucknow 226 003, India

⁷Vice Chancellor, King George's Medical University, Lucknow 226 003, India

⁸Director, All India Institute of Medical Sciences, Rishikesh 249 203, India

Correspondence should be addressed to Neetu Singh; neetusingh@kgmcindia.edu

Rebecca Chowdhry and Neetu Singh contributed equally to this work.

Received 13 September 2018; Revised 24 November 2018; Accepted 26 March 2019; Published 14 April 2019

Academic Editor: Jiangke Yang

Copyright © 2019 Rebecca Chowdhry et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Retrospective analysis has already shown correlation between severe Chronic Periodontitis (CP) cases with human papiloma virus (HPV). Hence, we aimed to explore deep-seated infected granulation tissue removed during periodontal flap surgery procedures for residential bacterial species between HPV+ and HVP- CP cases, which may serve as good predisposition marker for oral cancer. All CP-granulation samples showed the prominence of *Firmicutes*, *Proteobacteria*, and *Bacteroidetes* phyla with an abundance of gram negative anaerobes, except *Streptococcus*. In Beta diversity nonmetric multidimensional scaling plot, the random distribution of species was observed between HPV+ and HPV- CP granulation-samples. However, an abundance of *Capnocytophaga ochracea* was observed in HPV+ CP samples ($p < 0.05$), while *Porphyromonas endodontalis*, *Macellibacteroides fermentas*, *Treponema phagedenis*, and *Campylobacter rectus* species were highly abundant in HPV- CP samples ($p < 0.05$). The differential species richness leads altered functions related to mismatch-repair and nucleotide excision-repair and cytoskeleton-proteins. Hence, differential abundance of gram negative bacterial species between HPV+ and HPV- granulation-samples under anaerobic conditions may release virulence factors which may alter pathways favouring carcinogenesis. Hence, these species may serve as good predisposition marker for oral-cancer.

1. Introduction

Chronic periodontitis (CP), bacterially induced chronic inflammatory disease of the periodontium, is one of the 50 most prevalent disabling health conditions affecting up to 80-90% of the population worldwide and in India [1, 2]. It is the main cause of tooth loss in adults and directly linked to systemic chronic diseases, including cardiovascular disease,

diabetes, stroke, and preterm birth pneumonia and cancer [3, 4]. Risk-factors like human-papilloma-virus (HPV), smoking, and tobacco further complicate the problem. The primary etiology of periodontitis disease is plaque biofilm affecting the periodontal tissues. Pocket formation, loss of periodontal attachment, loss of alveolar bone, root furcation exposure, tooth mobility in advanced cases of bone

destruction, and infected granulation tissue formation are the main features of CP. During the periodontal flap surgery, the removing of deep seated pocket granulation tissue is the main process of treatment for filling wounds and improving the condition of new attachment.

Currently a total of 600 bacterial species are represented as periodontal pathogens that can colonize dental surfaces over (supra) and below (sub) the gingival margin and oral mucous membranes [3, 5]. Most evaluated, subgingival plaque samples from CP patients possess both periodontal pathogens (including anaerobes and facultative aerobes) [3] and commensal flora (*actinomyces*, *streptococcus*, and *staphylococcus* sp.). Dominance of periodontal pathogens mediated chronic inflammation may cause deregulation of local oral immune response. This may provide advantage to opportunistic infections which may predispose to carcinogenesis. Other sources like saliva and pooled and site-specific subgingiva have also been profiled for microbial flora. Presence of periodontal pathogens was observed in stimulated saliva and pooled subgingival plaque samples [6]. Based on above studies, we comprehend that most of the sites affected by CP have been profiled although deep-seated infected granulation tissue removed during periodontal flap surgery procedure has not been profiled for its microbiota. Importantly, granulation tissue has been suggested as a source of stem cells-involved in filling wounds and bone regeneration [7].

Additionally, deeper-pockets and higher mean alveolar bone loss mediated-CP have been shown to be associated with HPV-16 positive compared to negative base-of-tongue-SCC samples (3.90 mm vs. 2.85 mm, $P = .01$). Recent reports have also compared the abundance of microbial species between normal and oral cancer samples in significant loss of bacteria and have been observed in head and neck squamous cell carcinoma compared to control in saliva samples. They also evaluated HPV status which showed slight discrimination through beta dispersion with significant higher variants in OPSCC HPV+ samples compared to control and OSCCC HPV- samples ($p < 0.02$) [8].

Different sequencing based methods have been used to explore the microbial complexity of CP samples. However, most recent publications suggest long read sequencing technology (V1-V9 full length variable region of 16S ribosomal gene) can provide finer phylogenetic profiling [9].

Hence, we hypothesized assessing association of microbiological profile with HPV+ and HPV- CP cases (Clinical Attachment Loss = or > 5mm) through long read sequencing technology [9]. We identified differences between HPV+ and HPV- CP bacterial communities' at all taxonomic levels. Additionally, we were able to identify least expressed uncultured bacterial species at genera and species level. Major shifts at the level of genera in HPV+ versus HPV- CP samples may provide a basis for further understanding HPV microbe interactions and creation of an inflammatory niche in CP which may lead to the process of carcinogenesis.

2. Materials and Methods

2.1. Screening of Patients, Sample Collection and DNA Extraction. A total of 125 CP patients were screened and 40

patients between age of 18 and 66 years (average 37 ± 11 years) with the history of no smoking and tobacco chewing, bleeding on probing, Periodontal pockets = or > 5mm, and having Clinical attachment loss (CAL) were included in the study (Supplementary Table 1). Patients with the history of smoking, tobacco chewing, systemic disorders (diabetes, hypertension etc.), blood dyscrasias, compromised immune system, consumption of antibiotics in the past three to six months, and surgical procedures in the past three to six months were excluded from the study. The study protocol was approved by the Institutional Review Board of the King George's Medical University. Written informed consent was obtained from all patients.

After clinical examination following diagnostic parameters as reported in [10] were subjected to routine scaling and root planing [11]. After strict oral hygiene for one week, i.e., cessation of any deleterious habit, completely periodontal flap surgery was conducted in 40-CP cases. Granulation tissues were collected during flap surgery and processed for DNA extraction using Qiagen mini DNA isolation Kit. DNA recovered from the extraction process was quantified using Q5000 UV-Vis Spectrophotometer (Quawell San Jose, CA) and Qbit Fluorimeter and stored at -20°C for further analysis.

2.2. Detection of HPV-16 E6. PCR was performed for detection of HPV in all 40 CP samples using HPV 16 E6 and B-actin specific primers; HPV-16 E6 forward primer, 5'-TCA-GGACCCACAGGAGCG-3'; HPV-16 E6 reverse primer, 5'-CCTCACGTCGCAGTAACTGTTG-3' and B-actin forward primer, 5'-TCACCCACACTGTGCCCATCTACGA-3', B-actin reverse primer, 5'-CAGCGGAACCGCTCATTGCCA-ATGG-3' in triplicate [12]. Additionally, we pooled six HPV+ oral cancers and treated them as a positive control for detection of HPV.

2.3. 16S rRNA Gene Amplification, Multiplexing, and PacBio Sequencing. As per [10], (a) amplification of full length 16S ribosomal gene (1464 bp) from 50 ng/ μl DNA from each CP sample using V1-V9 variable region primers; (b) barcoding of each amplified amplicon using paired 16-nt symmetric barcodes (<https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/Barcoding-with-SMRT-Analysis-2.3>) (Table 1); (c) multiplexing of barcoded-purified amplicons in a group of 5 libraries (Supplementary Table 2); and (d) Sequencing using DNA Sequencing Reagent Kit 4.0 v2 (Pacific Biosciences, USA) in a total of 24 SMRT cells (3 cells/library) using 6-hour movies data collection protocol were performed.

2.4. Raw Sequence Analysis and Metagenome Functional Predictions. The PacBio generated raw sequences were processed for RS_Read of Insert (ROI) algorithm, demultiplexed, processed, aligned, and classified at species level and generated alpha (Ace, chao, shannon index, invsimpson and sobs) and beta diversity (jclass, thetayc, nmds) indexes using Mothur (version 1.34.4) [13] against "Greengenes reference database" gg_13.8_99 as per [10]. Rarefaction curves, principal coordinates (PCoA), and nonmetric multidimensional scaling (NMDS) were also generated using PASTv3.11 [10, 14].

TABLE 1: Summary of the primer pair used to generate the 16S rDNA gene fragment fragments and the characteristics of each region in CP samples.

Region		Forward Primers (with forward barcode-Red)	Amplicon length (Without Barcode)
	F	AGRGTTYGATYMTGGCTCAG	
V1-V9	F-1	TGAGTGACGTGTAGCGAGRGTTYGATYMTGGCTCAG	1,464
	F-2	GACAGCATCTGCGCTCAGRGTTYGATYMTGGCTCAG	
	F-3	TGCGAGCGACTCTATCAGRGTTYGATYMTGGCTCAG	
	F-4	TGCTCTCGTGTACTGTAGRGTTYGATYMTGGCTCAG	
Region		Reverse Primers (with reverse barcode-Red)	
	R	RGYTACCTTGTACGACTT	
V1-V9	R-1	GCTCGACTGTGAGAGA RGYTACCTTGTACGACTT	1,464
	R-2	TGCTCGCAGTATCACA RGYTACCTTGTACGACTT	
	R-3	GCAGACTCTCACACGC RGYTACCTTGTACGACTT	
	R-4	AGACAGCATCTGCGCTCRGYTACCTTGTACGACTT	

where

F= Forward primer,

F-1= Forward barcode primer 1,

F-2= Forward barcode primer 2,

F-3= Forward barcode primer 3,

F-4= Forward barcode primer 4

R= Reverse Primer,

R-1= Forward barcode primer 1,

R-2= Forward barcode primer 2,

R-3= Forward barcode primer 3,

R-4= Forward barcode primer 4,

R = A or G,

Y = C or T,

M = A or C, and

N = any base.

Raw data analysis and metagenome functional predictions were performed as per [10]. Briefly, metagenomics functional inferences were processed through PICRUSt tool [15] using marker gene data and database of reference genomes from the 16S rRNA data and KEGG (Kyoto Encyclopedia of Genes and Genomes) database [16]; grouping of closed-reference OTUs (97% sequence identity) was performed using uclust and the greengenes reference database; OTUs abundance was normalized automatically using 16S rRNA gene copy numbers from known bacterial genomes in Integrated Microbial Genomes (IMG) [17]; predicted genes and their function were aligned to KEGG database; differences between the abundance of functional pathways and the species abundance among groups (HPV+/HPV-) were compared through software STAMP [5]; statistics and visualization of functional data like heat maps, square and PCA plots were generated displaying differences between the groups using STAMP; two-side Welch's t-test and Benjamini-Hochberg FDR correction were used in two-group analysis.

3. Results

3.1. Multiplexed Amplicon Sequencing of Full-Length 16S rRNA Gene. A total of 8 pooled and barcoded 16S rRNA amplicon libraries were prepared and sequenced on PacBio RSII system. Following sequencing and demultiplexing using Read of Insert protocol (RS_Read of Insert), 3 samples (CP25, CP28, and CP 29) were excluded from further analysis due to less number of reads. Following removal of 3 samples, a

total of 3,59,296 raw CCS reads were generated from 37 CP samples with an average accuracy of 99.36%, 12 numbers of mean passes, and a mean length of 1,402 nt. Further, removal of chimeras and CCS reads shorter than 1400 nt and longer than 1600 nt and 2,72,221 reads were obtained from 37 CP samples (Supplementary Table 3). The number of processed full-length 16S rRNA sequences per CP sample ranged from 672 to 28,780 with an average of 7,357 reads/sample (\pm SD 5117 reads) (Supplementary Table 3).

3.2. Assessment of Severe CP Associated Bacterial Diversity in Granulation Tissue Isolated during Periodontal Flap Surgery Using High-Coverage PacBio 16S rRNA Gene

3.2.1. Study Group Characteristics. In the present study, a total of 37 CP samples were collectively analyzed to profile the bacterial communities in unexplored granulation tissue of severe CP cases. The numbers of operational taxonomic units (OTUs) at a 3% dissimilarity level and the diversity estimates are listed in Supplementary Table 4. To assess the diversity of bacterial community present within each sample, a series of alpha diversity indices were calculated. Alpha-diversity measures the biological diversity of a community, taking both species richness and variance in species proportion into consideration. Using Shannon index as the metric for alpha-diversity, we found the average to be 6.84 (\pm SD 0.60) with individual samples diversity index ranging from 2.19 to 8.05 (Supplementary Table 5a). Alpha-Diversity Analysis tests ANNOVA (Repeated Measures ANNOVA) showed no

significant differences between samples equal means (Supplementary Table 5b). Tukey's pairwise test also showed no significant difference between sampled species in each sample (Supplementary Table 5c). Alpha-Diversity Analysis using MOTHUR generated rarefaction curves which nearly plateau off for the majority of the samples (Supplementary Figures 1a and 1b), indicating that sufficient sampling has been performed to capture the total diversity of the communities.

3.3. Determination of Phyla in CP Samples Identified in Tissue Samples. We found a total of 13 assigned phyla present with 3 of these dominating across all of the samples: *Proteobacteria* (31.43%), *Firmicutes* (31.32%), and *Bacteroidetes* (22.08%). Next *Spirochaetes*, *Actinobacteria*, and *Fusobacteria* showed 2.59, 3.28 and 8.53%, respectively, and other six phyla had a relative abundance lower than 1% (Supplementary Figure 2a).

3.4. Determination of Species/Phylotypes/Genus Level in CP Granulation Tissue Samples. The observed number of OTUs based on high-quality sequences processed through Mothur package against green gene_13.8_99.gg.tax database for species richness in 37 CP samples. Among these a total of 112 observational IDs were unclassified and 125 were classified at the species level (Supplementary Figure 2b).

In *Firmicutes*, *Streptococcus* showed dominance (42.5%) with unclassified *Streptococcus* ID1000186 sp. (86.9%) and *Streptococcus agalactiae* (12.0%) rest were less than 1%. Both *Veillonella* and *Selenomonas* showed 19.2% and 18.9% OTUs. Two species *parvula* and ID1004529 of *Veillonella* showed a predominance of 45.2 and 54.67 OTUs, while *Selenomonas bovis*, *Selenomonas* ID1030589, and *Selenomonas noxia* showed OTUs ranging from 21.39 to 40.62 (Supplementary Figure 2b).

In *Proteobacteria*, *Neisseria* showed dominance with unclassified *Neisseria* ID1007399 sp. 35.2% and *Neisseria oralis* 62.61%, and the rest were less than 1%. Both *Campylobacter* and *Sphingomonas* had 8% OTUs individually (Supplementary Figure 2b).

In *Bacteroidetes*, *Prevotella* showed dominance (43.7%) with unclassified *Prevotella* ID1000563 sp. (51.2%), *Prevotella intermedia* (20%), and *Prevotella nigrescens* (12.59%). Next dominance was of *Capnocytophaga* (27.9%) with unclassified *Capnocytophaga ochracea* sp. (67.8%) and *Capnocytophaga* ID1001498 (31.9%). 14.41% was the occurrence of *Macellibacteroides fermentans* while *Porphyromonas* with two species *endodontalis* and ID1001120 showed 4.0% richness and *Tannerella*- ID101201 5.6% (Supplementary Figure 2b).

3.5. Comparisons of Microbial Community Profiles CP Granulation Tissue Samples between HPV+ and HPV- Samples and Associated Functions. Generated OTUs through Mothur were further analysed between HPV+ and HPV- CP samples. OTU based correlation coefficient analysis showed $R^2 = 0.82$ between HPV+ and HPV- samples depicting a poor relationship between microbial abundance between the two groups (Supplementary Figure 3a). Further, nonmetric multidimensional scaling and principal component analysis of samples represented OTUs between HPV+ and HPV- samples which were similarly based on distances and showed

random distribution (Supplementary Figures 3b and 3c). Different beta-diversity analysis indexes reporting beta diversity between HPV+ and HPV- CP dataset are presented in Supplementary Table 6.

STAMP based generation of heat-map of OTUs showed, among the HPV+ samples, CP-15 (55 yrs-F), CP-19 (32 yrs-F), CP-1 (26 yrs-F), and CP-24 (39 yrs-F) were in one cluster while CP-5 (26 yrs-F), CP21(33 yrs-F), and CP-26 (48 yrs-M) showed individual existence (Supplementary Figure 4). PICRUST analysis of normalized OTUs showed the formation of three distinct clusters among the HPV+ samples. Among the HPV+ samples, high and significant ($p < 0.05$) abundance of *Bacteroidetes* phyla-*Capnocytophaga ochracea* was observed, while *Bacteroidetes* phyla-*Porphyromonas endodontalis*, *Macellibacteroides fermentans*, and *Spirochaetes* phyla-*Treponema phagedenis* and *Proteobacteria* phyla-*Campylobacter rectus* species were highly abundant in HPV- samples (Figure 1).

Functional predictions of the observed species abundance showed the formation of two distinct clusters HPV+ CP and HPV- CP cases in heat-map profiles (Supplementary Figure 5a), with correlation coefficient analysis $R^2 = 0.99$. As per Figure 2, difference in mean proportion ranging 0.0-0.1% included highly significant penicillin and cephalosporin biosynthesis ($p = 3.69e^{-3}$) and significant-sulphur relay system ($p = 0.01$), ascorbate and aldarate metabolism (0.02), polyaromatic hydrocarbon degradation ($p = 0.04$), and propanoate metabolism ($p = 0.04$). Only phosphotransferase system (PTS) showed mean proportion between 0.1 and 0.2% with significance value of $p = 0.03$. Among HPV- sample the difference in mean proportion ranged between 0.0 to 0.1 including highly significant-methane metabolism ($p = 7.57e^{-4}$), transcription machinery ($p = 1.61e^{-3}$), epithelial cell signalling in *helicobacter pylori* ($4.66e^{-3}$), mismatch repair ($p = 6.80e^{-3}$) histidine metabolism ($p = 8.11e^{-3}$) and significant-carbon fixation pathways in prokaryotes ($p = 0.017$), alanine, aspartate, and glutamate metabolism ($p = 0.01$), plant pathogen interaction ($p = 0.03$), nucleotide excision repair ($p = 0.04$), and cytoskeleton proteins ($p = 0.04$), while propanoate metabolism, sulphur relay system ascorbate, and aldarate metabolism, phosphotransferase system (PTS), and polycyclic aromatic hydrocarbon degradation related microbial functions were highly represented in HPV+ samples in comparison to HPV- samples. Among the HPV- samples, a significantly high ($p > 0.05$) prevalence of microbial communities was observed in methane metabolism, transcription machinery, mismatch repair, histidine metabolism, nucleotide excision repair, cytoskeleton proteins, etc. (Figure 2).

3.6. Comparisons of Microbial Community Profiles CP Granulation Tissue Samples between Male and Female Samples and Associated Functions. Generated OTUs through Mothur were also analyzed between male and female-CP samples. OTU based correlation coefficient analysis showed $R^2 = 0.96$ between male and female-CP samples depicting a poor relationship between microbial abundance between the two groups (Supplementary Figure 6a). STAMP based generation of heat-map of OTUs had, among the male and female-CP,

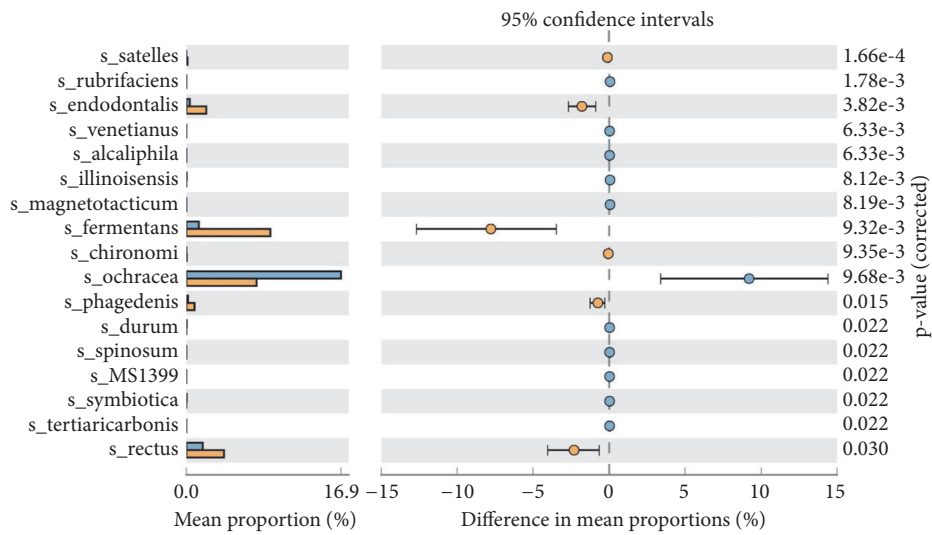


FIGURE 1: Mean proportion of bacterial species abundance between HPV+ (blue bar) and HPV- (yellow bar) CP samples. The significant differences observed between the two groups at 95% confidence level and $p < 0.05$ are reported.

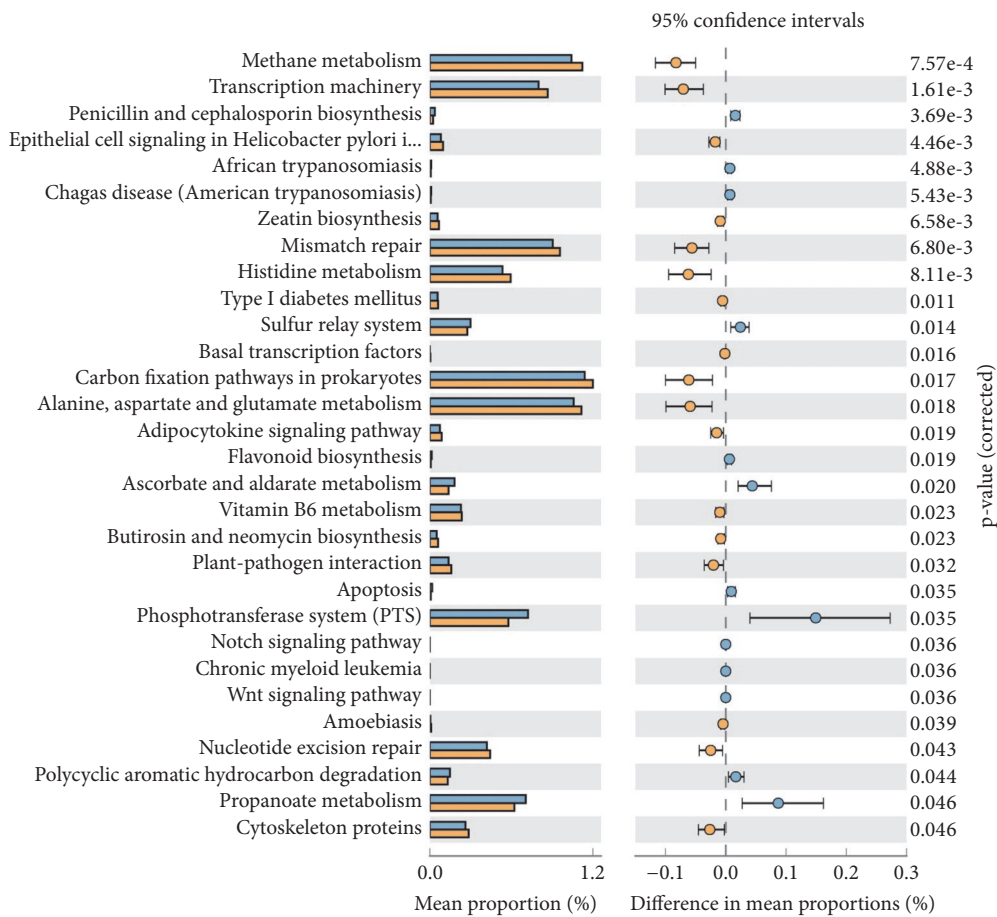


FIGURE 2: Mean proportion of bacterial species functional predictions between HPV+ HPV+ (blue bar) and HPV- (yellow bar) CP samples. The significant differences observed between the two groups at 95% confidence level and $p < 0.05$ are reported.

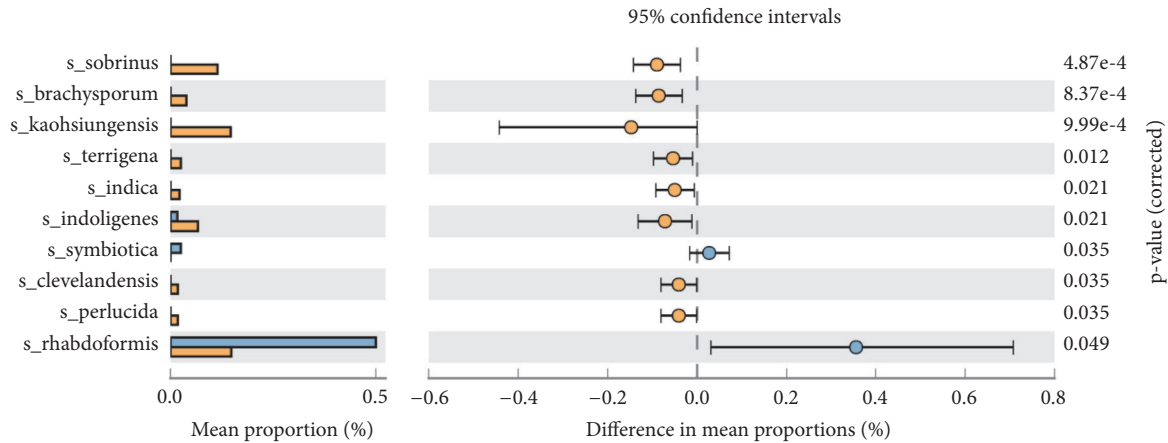


FIGURE 3: Mean proportion of bacterial species abundance between male (blue bar) and female (yellow bar) samples in CP samples. The significant differences observed between the two groups at 95% confidence level and $p < 0.05$ are reported.

a distributive type (Supplementary Figure 6b). PICRUST analysis of normalized OTUs showed a significant dominance of *Desulfohalobium rhabdoformis* ($p = 0.05$; difference in mean proportion 0.2-0.4%) and *Serratia symbiotica* in males ($p = 0.04$; difference in mean proportion 0.0-(-0.2%)) and *Pseudoxanthomonas kaohsiungensis* ($p = 9.99 \times 10^{-4}$; difference in mean proportion 0.0-0.2%) in females (Figure 3).

Functional predictions of the observed species abundance showed the formation of two distinct clusters (pathway based-y-axis) in heat-map profiles (Supplementary Figure 7a), with correlation coefficient analysis showing $R^2 = 0.998$. Difference in mean proportion between male and female (ranged between 0.00 and 0.05) showed high representation of pathways of I cluster-lipopolysaccharide biosynthesis ($p = 0.05$), alanine, aspartate, and glutamate metabolism ($p = 0.04$), arginine and proline metabolism ($p = 0.04$), and carbon fixation pathways in prokaryotes ($p = 0.11$). Among females, a significantly high ($p < 0.05$) prevalence of benzoate degradation ($p = 0.13$), ascorbate and aldurate metabolism ($p = 0.23$), bacterial toxins ($p = 0.29$), valine, leucine, and isoleucine degradation ($p = 0.44$), and synthesis and degradation of ketone bodies pathways ($p = 0.44$) was observed (Figure 4).

3.7. Comparisons of Microbial Community Profiles in CP Granulation Tissue Samples between Samples Aged above 40 and below 40 and Associated Functions. Generated OTUs through Mothur were also analyzed between above 40- and below 40-year-old CP patients samples. OTU based correlation coefficient analysis showed $R^2 = 0.96$ between above 40- and below 40-year-old CP patients samples depicting a poor relationship between both kinds of microbial abundance between the two groups (Supplementary Figure 8a). Further, beta diversity nonmetric multidimensional scaling plot represented OTUs between above 40- and below 40-year-old CP patient's samples which were similarly based on distances and showed random distribution.

STAMP based generation of heat-map of OTUs had, among the above 40 and below 40-year-old CP patient's samples, a distributive type (Supplementary Figure 8b).

PICRUST analysis of normalized OTUs showed a significant dominance of *Streptococcus sorbinus* ($p = 3.86 \times 10^{-3}$, difference in mean proportion 0.0-0.2), *Serratia symbiotica* ($p = 0.037$, difference in mean proportion 0.0-0.2) in those above 40 years of age, and *Pseudoxanthomonas mexicana* ($p = 9.99 \times 10^{-4}$, difference in mean proportion 0.0-(-0.4)) in those below 40 years of age (Figure 5).

Functional predictions of the observed species abundance showed the formation of two distinct clusters in heat-map profiles (Supplementary Figure 8c), with correlation coefficient analysis of $R^2 = 0.99$. Amino sugar and nucleotide sugar metabolism, porphyrin and chlorophyll metabolism, fructose and mannose metabolism, galactose metabolism, cysteine and methionine, ABC-transporters, and transporters were highly represented in those above 40 compared to those below 40 years of age ($p < 0.05$). Among those below 40 years of age, a significantly high ($p > 0.05$) prevalence of secretion system and bacterial secretion system pathways ($p = 0.03$) was observed (Figure 6).

4. Discussion

The recent study on profiling of coral-associated bacterial communities [7] very well demonstrated the application of PacBio sequencing technology for microbial diversity analysis. In the present study, we utilized one-step PCR procedure to amplify 16S gene with bar-coded, HPLC purified primers as it was aimed to amplify only a single target harbouring complete V1-V9 region. Using SMRT technology, we identified OTUs based on high-quality sequences processed through Mothur package against green gene_13.8.99.gg.tax database for species richness in 37 CP granulation tissue samples. Among these a total of 112 observational IDs were unclassified and 125 were classified at the species level (Supplementary Figures 2a and 2b; Supplementary Table 4). These unclassified species may represent microbial communities' specific for granulation tissues for which full-length OTU IDs have not been characterized in green gene database at the species level. Although some reports have used a different database for comprehensive taxonomic identification, the limitation in

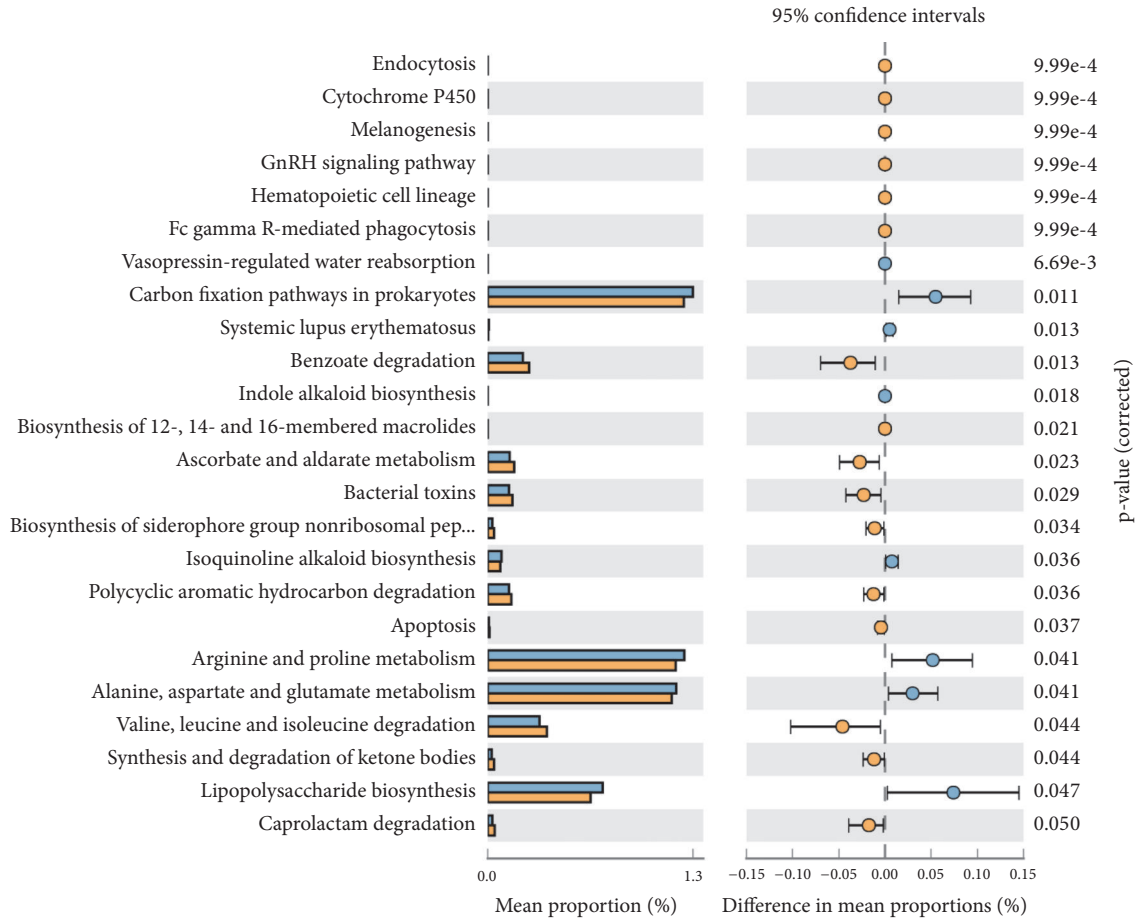


FIGURE 4: Mean proportion of bacterial species functional predictions between male (blue bar) and female (yellow bar) samples in CP samples. The significant differences observed between the two groups at 95% confidence level and $p < 0.05$ are reported.

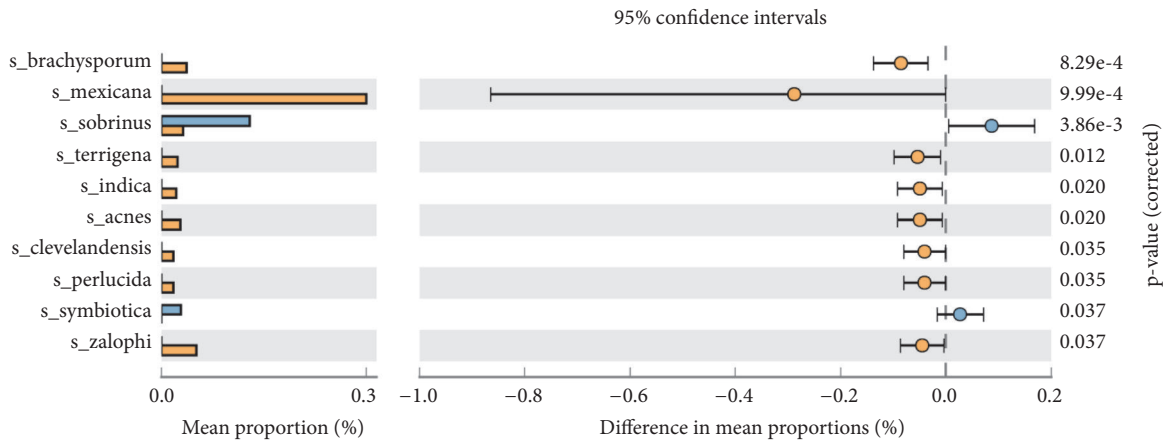


FIGURE 5: Mean proportion of bacterial species abundance between samples above 40 (blue bar) and below 40 years (yellow bar) age in CP patients. The significant differences observed between the two groups at 95% confidence level and $p < 0.05$ are reported.

our study was that PICRUST utilizes green gene database based OTU IDs for functional classification. Thus, we cannot annotate these unclassified species.

Under normal conditions, microorganism reported in the healthy sub- and supragingival plaque showed a decrease of *Firmicutes* compared to other superficial oral regions

like hard palate, buccal mucosa, keratinized gingiva, saliva, and tongue. Genera-wise *Streptococcus* showed dominance in the sub- and supragingival plaque with an abundance of *Veillonella*, *Prevotella*, *Neisseria*, *Fusobacterium*, *Actinomyces*, *Leptotricha*, and *Cornybacteria*. The above evidence suggests decrease of *Firmicutes* was observed as depth increases in

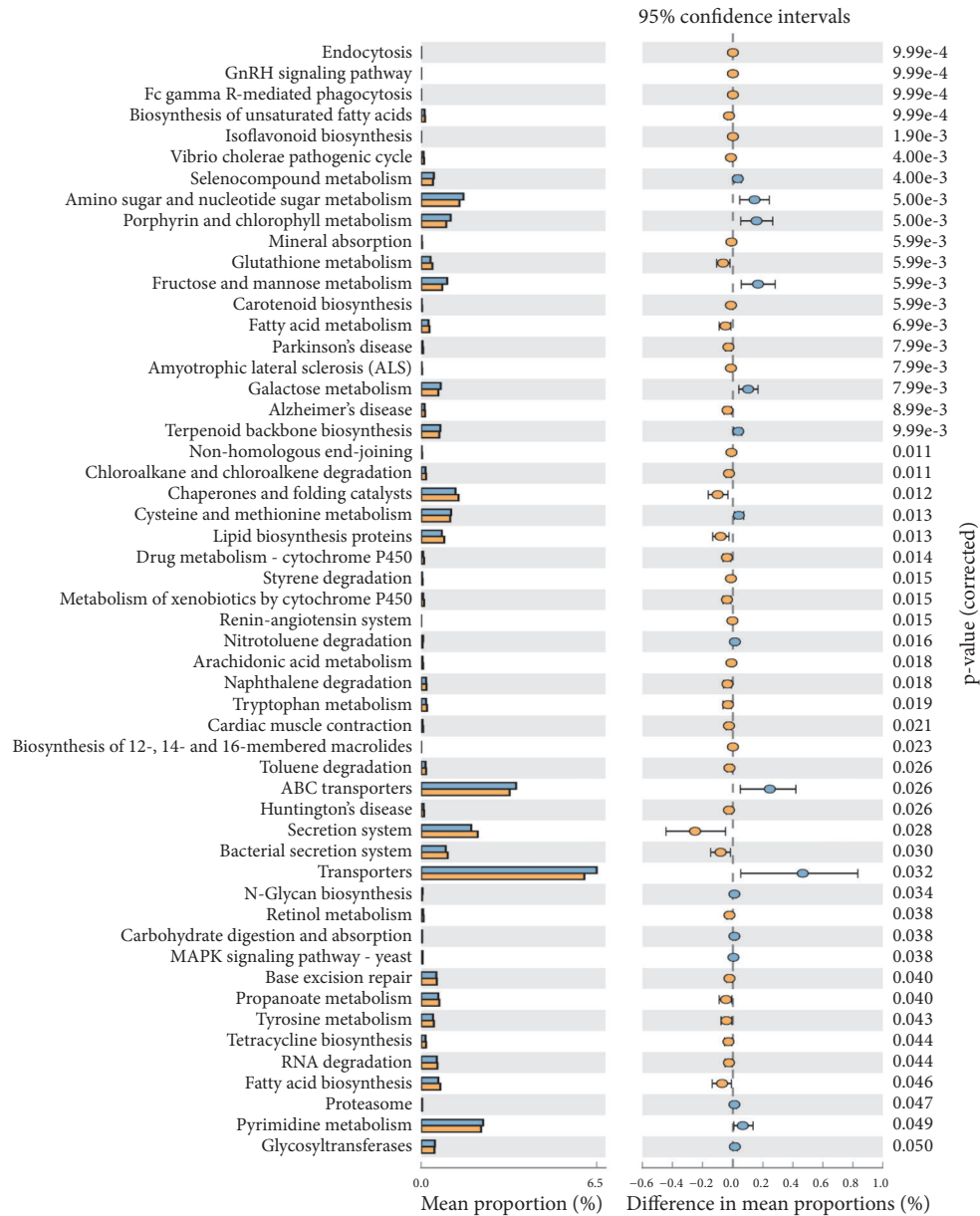


FIGURE 6: Mean proportion of bacterial species functional predictions between samples above 40 (blue bar) and below 40 years (yellow bar) age in CP samples. The significant differences observed between the two groups at 95% confidence level and $p < 0.05$ are reported.

healthy samples [18]. Significant high relative abundance of *Synergistete* phyla [19] and *Firmicutes* phyla has also been reported under CP conditions compared to healthy conditions by Galimans [20]. Additionally, Phyla *Proteobacteria* and *Bacteroidetes*, as well as genera, *Prevotella*, *Porphyromonas*, and *Treponema*, showed high abundance across all subgingiva of severe chronic diseased samples [21, 22]. Popova et al. [3] reviewed and suggested *Porphyromonas gingivalis*, *Treponema denticola*, and *Tannerella forsythia* form a consortium in the subgingival biofilm and main periodonto-pathogenic bacteria [3], while other predominant species in the disease process were *Aggregatibacter actinomycetemcomitans*, *Fusobacterium nucleatum*, *Prevotella*

intermedia, *Campylobacter rectus*, *Peptostreptococcus migros*, and *Eikenella corrodens* [3].

In our study, all deep seated-CP granulation tissue samples showed the prominence of *Firmicutes*, *Proteobacteria*, and *Bacteroidetes* (Supplementary Figure 2a). Most of the genera and identified species in the above phylum were gram-negative anaerobes [*Firmicutes*-*Veillonella parvula* (45.2%) and *Veillonella* ID1004529 (54.67%); *Selenomonas* ID1030589 (21.39%), and *Selenomonas noxia* (40.62%); *Proteobacteria*-*Neisseria* ID1007399 sp. 35.2% and *Neisseria oralis* 62.61%; *Bacteroidetes*-*Prevotella* ID1000563 sp. (51.2%), *Prevotella intermedia* (20%) and *Prevotella nigrescens* (12.59%) and *Capnocytophaga ochracea* sp. (67.8%) and

Capnocytophaga ID1001498 (31.9%) (Supplementary Figure 2b)] [19]. Although the species *Selenomonas* and *Veillonella species* in phyla-Firmicutes; *Neisseria species* in Phyla Proteobacteria; *Capnocytophaga* and *Prevotella species* in phyla *Bacteroidetes* showed equivalent abundance as earlier reported [18] and some of the species have been reported in children with dental caries [23] except Firmicutes-*Streptococcus* (ID1000186-86.9%, *agalactiae*-12%,) species were gram-positive (Supplementary Figure 2b). This corroborates the fact that gram-negative bacteria play important role in tissue loss in CP, most likely due to their production of various virulence factor including collagenase, protease, and lipopolysaccharide endotoxins [3]. Additionally, high dipeptidase activity has been reported in *Proteobacteria* and *Bacteroidetes* [24] which may further cause tissue damage.

Random distribution of microbial species was observed in HPV+ and HPV- CP samples (Supplementary Figures 2a and 3b) with significant dominance of *Capnocytophaga ochracea* in HPV+ samples (Figure 1). *Capnocytophaga ochracea*, a capnophilic aerotolerant anaerobe, has been reported to ferment hexoses and disaccharides into their respective phosphoesters. During this process, phosphoenol-pyruvate is used as a source of energy to regulate phosphotransferase pathway (PTS-in our study difference in mean proportion for PTS pathway was 0.1-0.2%; $p=0.035$) (Figure 2), which transports sugar across bacterial cell membrane [25]. Also, *C. ochracea* have adhesion sites (*L. rhamnose* sensitive, N acetylneuraminic acid sensitive sites) which help it to bind to other bacterial species like *Streptococcus* and *Fusobacteria* [26]. This coadhesion of species contributes to early dental plaque formation and may play important role in causing periodontitis as suggested by Popova et. 2013 [3]. Dominance of *C. ochracea* in HPV+ sample may further explain the higher mean alveolar loss (3.90 mm) in HPV+ tongue SQCC-CP cases compared to 2.85 mm in HPV- tongue SQCC-CP cases (a retrospective study by Guerrero-Preston et al., 2016) [8]. On comparing HPV+ and HPV- cases, NER, mismatch repair pathway showed a significant decrease in HPV+ samples suggesting the higher formation of DNA adducts, which may lead to the process of carcinogenesis.

Members of *Porphyromonas genera*, *Macellibacteroides fermentans* and *Porphyromonas endodontalis*, *Campylobacter rectus*, and *Treponema phagedenis* showed prominence in HPV- CP samples (Figure 1). *Porphyromonas* sp. and *Campylobacter* sp. are responsible for beta-lactamase and bone loss activities, respectively [2]. *Porphyromonas endodontalis*, gram negative black-pigmented anaerobes, and *Macellibacteroides fermentans* (obligate anaerobe) have also been shown to be associated with CP [2, 27]. *Porphyromonas* species do not ferment carbohydrates; instead, they use proteinaceous substrates as carbon and energy source and also possess dipeptidyl peptidase (DPP) 7 peptidase activity [8] which may lead to tissue loss in CP. *Campylobacter rectus* has also been reported as periodontal pathogens which invade into deeper tissue. It also shows commensalism with *Porphyromonas gingivalis* [3]. Both *Porphyromonas endodontalis* and *Campylobacter rectus* have been reported in orange

complex and similar to the red complex have been shown to have a significant association with increasing pocket depth [19, 20].

Based on nonmetric multidimensional scaling plot (Supplementary Figure 6a) and heat map profile (Supplementary Figure 6b), we observed no species difference between male and female CP granulation tissue which corroborated with human oral microbiome study which showed no specific difference between female and males under normal conditions. However, through Picrust analysis, we identified *Pseudoxanthomonas kaohsiungensis* which showed prevalence in CP females (Figure 3) and *Desulfobulbus rhabditiform* (gram negative) in granulation tissue of male CP patients. *Desulfobulbus rhabdoformis* (gram negative) in granulation tissue of male CP patients which has been previously associated with periodontitis in the CP-smokers (Figure 3) [27, 28] and is also associated with poor response to periodontal therapy [27]. Also inorganic sulfate has been determined by controlled-flow anion chromatography in human saliva and sweat may be a substrate for *Desulfobulbus rhabdoformis*[24] and the main product of reduction is hydrogen sulphide which is potentially toxic.

Due to enhanced abundance of *Desulfobulbus rhabdoformis*, upregulation of Lipopolysaccharide biosynthesis ($p=0.047$), protein metabolism including alanine, aspartate, and glutamate metabolism ($p=0.041$), and arginine and proline metabolism as well as carbon fixation pathways ($p=0.011$) was observed both in our (Figure 4) and Bizzarro's studies [28].

Although another species, *Pseudoxanthomonas mexicana*, identified in below 40 years of age group (Figure 5) has been reported to possess dipeptidyl aminopeptidase [23], which may play important role in tissue loss in CP, we also identified upregulation of benzoate degradation, ascorbate and aldorate metabolism, bacterial toxins, valine, leucine, and isoleucine degradation, synthesis, and degradation of ketone bodies pathways in females and of secretion system and bacterial secretion system, pathways in below 40 years of age group (Figure 6).

Hence, we conclude that dominance of gram negative species in granulation tissue of severe CP cases with peptidase activity in deep seated granulation tissue of severe CP cases may be responsible for increased depth of periodontal pocket. They also release virulence factors which utilize the multitude of methods to invade mammalian host, damage tissue sites, and thwart the immune system. This has also been suggested through other studies which showed a link between proteobacteria and dysbiosis, to explain the potential relationship between inflammation and gene level differences in gut microbial functions [29]. Further, an association of HPV with granulation tissue-microbiota under anaerobic conditions enhances fermentation of hexoses using PEP (phosphoenolpyruvate) as a source of energy may synergize various processes and may create an environment favouring carcinogenesis. Hence, granulation tissue may serve as good predisposition marker for oral cancer.

Data Availability

The Sequence files and metadata for all samples used to support the findings of this study have been deposited in the NCBI repository (SRA Accession: SRR7088851, SRR7088868, SRR7088872, SRR7088845, SRR7088847, SRR7088852, SRR7088855, SRR7088842, SRR7088841, SRR7088844, SRR7088870, SRR7088859, SRR7088854, SRR7088863, SRR7088869, SRR7088856, SRR7088864, SRR7088866, SRR7088853, SRR7088857, SRR7088858, SRR7088867, SRR7088865, SRR7088860, SRR7088862, SRR7088861, SRR7088871, SRR7088846, SRR7088843, SRR7088849, SRR7088848 and SRR7088850 under Bio Project (PRJNA451246) and Biosample (SAMN08966100).

Additional Points

Originality Significance Statement. The work carried out in this manuscript is original, describing Dysbacteriosis induced severe Chronic Periodontitis (CP) cases which have been correlated with human papiloma virus (HPV) causing oral cancer. Highly abundant *Capnocytophaga ochracea* in HPV+ CP samples compared to *Porphyromonas endodontalis*, *Macellibacteroides fermentas*, and *Treponema phagedenis* and *Campylobacter rectus* species in HPV- samples lead to downregulation of mismatch-repair and nucleotide excision-repair, with cytoskeleton-proteins creating an environment favoring carcinogenesis. The abundance of above species in HPV+ granulation-tissue may serve as good predisposition marker for oral-cancer.

Conflicts of Interest

The authors also declare no conflicts of interest.

Authors' Contributions

Neetu Singh and Ravi Kant conceived and designed the experiments. Rebecca Chowdhry and Nand Lal provided clinically diagnosed samples. Archana Mishra, Dinesh Kumar Sahu, Neetu Singh, Anjana Singh, and Rebecca Chowdhry were responsible for sample collection and nucleic acid isolation and performed the experiment. Neetu Singh, Dinesh Kumar Sahu, and Ratnesh Kumar Tripathi analyzed the data. Neetu Singh wrote and discussed the paper. Madan Lal Brahma Bhatt and Ravi Kant contributed reagents/materials/analysis tools. Neetu Singh, Dinesh Kumar Sahu, Ratnesh Kumar Tripathi, Ravi Kant, and Madan Lal Brahma Bhatt reviewed the paper. Rebecca Chowdhry and Neetu Singh have equal contributions.

Acknowledgments

The authors thank Research Cell-KGMU for providing funds for accomplishment of this project; this work was supported by "Research Cell, King George's Medical University Fund."

Supplementary Materials

Supplementary Table 1: Selection of CP patients and their status related to Human Papiloma Virus (HPV+/HPV-) used in present study. Supplementary Table 2: Multiplexing with PacBio Barcoded Universal Primers in CP samples. Supplementary Table 3: The number of processed full-length 16S rRNA sequences per CP samples ranged from 672 to 28,780 with an average of 7,357 reads/sample. Supplementary Table 4: Observed number of OTUs at a 3% dissimilarity level and the diversity estimates. Supplementary Table 5a: Alpha-Diversity values of Shannon index for individual CP samples. Supplementary Table 5b: Alpha-Diversity Analysis tests ANNOVA (Repeated Measures ANNOVA) equal means. Supplementary Table 5c: Tukey's pairwise test showing no significant difference between sampled species in each sample. Supplementary Table 6: Different indexes reporting Beta-Diversity between HPV+ and HPV- CP dataset. Supplementary Figure 1a: Rarefaction Curves calculated at 3% dissimilarity compare the number of reads with the number of operational taxonomic units (OTUs) found in the DNA from the 40 CP samples. The curves for individual samples which nearly plateaued off for the majority of the samples. Supplementary Figure 1b: Sample rarefaction (Mao tau) curve of Alpha-Diversity for species richness estimation ($p < 0.05$) as a function of number of samples derived from Shannon index values of individual samples. Supplementary Figure 2a: Diversity of bacterial phyla identified in granulation tissue of CP samples. Supplementary Figure 2b: Diversity of bacterial species in CP samples identified in granulation tissue. Supplementary Figure 3a: Scatter plot and correlation estimates of bacterial species abundance between HPV+ and HPV- samples in CP samples. Supplementary Figure 3b: Beta-Diversity nonmetric multidimensional scaling plot based on Theta YC two axis coordinate values obtained through Mothur using Weighted UniFrac algorithm, showing distance or dissimilarity between each sample pair. Supplementary Figure 3c: Principal Coordinates Analysis (PCoA) plot from ThetaYC values of beta diversity distance matrix, where coordinate 1 (X-axis) shows the maximum amount of variation present in the samples, with respect to coordinate 2 (Y-axis). Representation of OTU's between HPV+ and HPV- samples is randomly distributed. Supplementary Figure 4: Heat-map profile of abundance distribution of identified species in CP samples. Supplementary Figure 5a: Scatter plot and correlation estimates of bacterial species functional predictions between HPV+ and HPV- samples in CP samples. Supplementary Figure 6a: Scatter plot and correlation estimates of bacterial species abundance between male and female samples in CP samples. Supplementary Figure 6b: Heat-map profile of abundance distribution of identified species in male and female samples in CP samples. Supplementary Figure 7a: Heat-map profile of functional predictions of identified species in male and female samples in CP samples. Supplementary Figure 7b: Mean proportion of bacterial species abundance between male and female CP samples. The significant differences observed between the two groups at 95% confidence level and $p < 0.05$ are reported. Supplementary Figure 8a: Scatter plot and correlation estimates of

bacterial species abundance between above and below 40 years of age samples in CP samples. Supplementary Figure 8b: Heat-map profile of abundance distribution of identified species in above and below 40 years of age samples in CP samples. Supplementary Figure 9a: Heat-map profile of functional predictions of identified species in above and below 40 years of age samples in CP samples. Supplementary Figure 9b: Mean proportion of bacterial species abundance between above and below 40 years of age samples in CP samples. The significant differences observed between the two groups at 95% confidence level and $p < 0.05$ are reported. (Supplementary Materials)

References

- [1] K. Benachinmardi, J. Nagamoti, S. Kothiwale, and S. Metgud, "Microbial flora in chronic periodontitis: Study at a tertiary health care center from North Karnataka," *Journal of Laboratory Physicians*, vol. 7, no. 1, p. 49, 2015.
- [2] M. E. Kirst, E. C. Li, B. Alfant et al., "Dysbiosis and alterations in predicted functions of the subgingival microbiome in chronic periodontitis," *Applied and Environmental Microbiology*, vol. 81, no. 2, pp. 783–793, 2015.
- [3] V. Galimanas, M. Hall, N. Singh et al., "Bacterial community composition of chronic periodontitis and novel oral sampling sites for detecting disease indicators," *Microbiome*, vol. 2, no. 1, p. 32, 2014.
- [4] F. A. Scannapieco, A. P. Dasanayake, and N. Chhun, "Does periodontal therapy reduce the risk for systemic diseases?" *Dental Clinics of North America*, vol. 54, no. 1, pp. 163–181, 2010.
- [5] F. E. Dewhirst, T. Chen, J. Izard et al., "The human oral microbiome," *Journal of Bacteriology*, vol. 192, no. 19, pp. 5002–5017, 2010.
- [6] D. Belstrom, M. L. Sembler-Møller, M. A. Grande et al., "Microbial profile comparisons of saliva, pooled and site-specific subgingival samples in periodontitis patients," *PLoS ONE*, vol. 12, no. 8, p. e0182992, 2017.
- [7] L. Diaz-Flores Jr., R. Gutierrez, J. F. Madrid, H. Varela, F. Valladares, and L. Diaz-Flores, "Adult stem cells and repair through granulation tissue," *Frontiers in Bioscience*, vol. 14, no. 4, pp. 1433–1470, 2009.
- [8] R. Guerrero-Preston, F. Godoy-Vitorino, A. Jedlicka et al., "16S rRNA amplicon sequencing identifies microbiota associated with oral cancer, human papilloma virus infection and surgical treatment," *Oncotarget*, vol. 7, no. 32, pp. 51320–51334, 2016.
- [9] P. D. Schloss, M. L. Jenior, C. C. Koumpouras, S. L. Westcott, and S. K. Highlander, "Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system," *PeerJ*, vol. 4, p. e1869, 2016.
- [10] R. Chowdhry, N. Singh, D. K. Sahu et al., "16S rRNA long-read sequencing of the granulation tissue from nonsmokers and smokers-severe chronic periodontitis patients," *BioMed Research International*, vol. 2018, Article ID 4832912, 10 pages, 2018.
- [11] S. H. Wei and K. P. Lang, "Periodontal epidemiological indices for children and adolescents: I. gingival and periodontal health assessments," *Journal of Pediatric Dentistry*, vol. 3, no. 4, pp. 353–360, 1981.
- [12] B. J. Keijser, E. Zaura, S. M. Huse et al., "Pyrosequencing analysis of the oral microflora of healthy adults," *Journal of Dental Research*, vol. 87, no. 11, pp. 1016–1020, 2008.
- [13] P. D. Schloss, S. L. Westcott, T. Ryabin et al., "Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities," *Applied and Environmental Microbiology*, vol. 75, no. 23, pp. 7537–7541, 2009.
- [14] Ö. Hammer, D. A. T. Harper, and P. D. Ryan, "PAST: paleontological statistics software package for education and data analysis," *Palaeontologia Electronica*, vol. 4, no. 1, pp. 1–9, 2001.
- [15] M. G. I. Langille, J. Zaneveld, J. G. Caporaso et al., "Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences," *Nature Biotechnology*, vol. 31, no. 9, pp. 814–821, 2013.
- [16] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: new perspectives on genomes, pathways, diseases and drugs," *Nucleic Acids Research*, vol. 45, no. 1, pp. D353–D361, 2017.
- [17] A. Y. Chuang, T. C. Chuang, S. Chang et al., "Presence of HPV DNA in convalescent salivary rinses is an adverse prognostic marker in head and neck squamous cell carcinoma," *Oral Oncology*, vol. 44, no. 10, pp. 915–919, 2008.
- [18] T. Chen, W. H. Yu, J. Izard, O. V. Baranova, A. Lakshmanan, and F. E. Dewhirst, "The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information," *Database*, vol. 2010, Article ID baq013, 2010.
- [19] S. S. Socransky, A. D. Haffajee, M. A. Cugini, C. Smith, and R. L. Kent Jr., "Microbial complexes in subgingival plaque," *Journal of Clinical Periodontology*, vol. 25, no. 2, pp. 134–144, 1998.
- [20] P. S. Kumar, A. L. Griffen, J. A. Barton, B. J. Paster, M. L. Moeschberger, and E. J. Leys, "New bacterial species associated with chronic periodontitis," *Journal of Dental Research*, vol. 82, no. 5, pp. 338–344, 2003.
- [21] V. T. Dosseva-Panova, C. L. Popova, and V. E. Panov, "Subgingival microbial profile and production of proinflammatory cytokines in chronic periodontitis," *Folia Medica*, vol. 56, no. 3, pp. 152–160, 2014.
- [22] Y. Tsai, C. Lin, J. Hou et al., "Effects of high di(2-ethylhexyl) phthalate (DEHP) exposure due to tainted food intake on prepubertal growth characteristics in a Taiwanese population," *Environmental Research*, vol. 149, pp. 197–205, 2016.
- [23] P. S. Kumar, A. L. Griffen, M. L. Moeschberger, and E. J. Leys, "Identification of candidate periodontal pathogens and beneficial species by quantitative 16S clonal analysis," *Journal of Clinical Microbiology*, vol. 43, no. 8, pp. 3944–3955, 2005.
- [24] T. Z. DeSantis, P. Hugenholtz, N. Larsen et al., "Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB," *Applied and Environmental Microbiology*, vol. 72, no. 7, pp. 5069–5072, 2006.
- [25] R. Calmes, G. W. Rambicure, W. Gorman, and T. T. Lillich, "Energy metabolism in Capnocytophaga ochracea," *Infection and Immunity*, vol. 29, no. 2, pp. 551–560, 1980.
- [26] E. I. Weiss, J. London, P. E. Kolenbrander, A. S. Kagermeier, and R. N. Andersen, "Characterization of lectinlike surface components on Capnocytophaga ochracea ATCC 33596 that mediate coaggregation with gram-positive oral bacteria," *Infection and Immunity*, vol. 55, no. 5, pp. 1198–1202, 1987.
- [27] T. Chiranjeevi, O. H. Prasad, U. Prasad et al., "Identification of microbial pathogens in periodontal disease and diabetic patients of south indian population," *Bioinformation*, vol. 10, no. 4, pp. 241–244, 2014.
- [28] S. Bizzarro, B. G. Loos, M. L. Laine, W. Crielaard, and E. Zaura, "Subgingival microbiome in smokers and non-smokers

in periodontitis: An exploratory study using traditional targeted techniques and a next-generation sequencing,” *Journal of Clinical Periodontology*, vol. 40, no. 5, pp. 483–492, 2013.

- [29] P. H. Bradley and K. S. Pollard, “Proteobacteria explain significant functional variability in the human gut microbiome,” *Microbiome*, vol. 5, no. 1, article no. 36, 2017.