



## Compositional and functional variations of oral microbiota associated with the mutational changes in oral cancer<sup>☆</sup>



Shun-Fa Yang<sup>a,b</sup>, Hsien-Da Huang<sup>c,d</sup>, Wen-Lang Fan<sup>e</sup>, Yuh-Jyh Jong<sup>c</sup>, Mu-Kuan Chen<sup>f</sup>, Chien-Ning Huang<sup>a,g</sup>, Chun-Yi Chuang<sup>h,i</sup>, Yu-Lun Kuo<sup>j</sup>, Wen-Hung Chung<sup>e,k,l,m</sup>, Shih-Chi Su<sup>e,m,\*</sup>

<sup>a</sup> Institute of Medicine, Chung Shan Medical University, Taichung 402, Taiwan

<sup>b</sup> Department of Medical Research, Chung Shan Medical University Hospital, Taichung 402, Taiwan

<sup>c</sup> Department of Biological Science and Technology, National Chiao Tung University, Hsinchu 300, Taiwan

<sup>d</sup> Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsin-Chu 300, Taiwan

<sup>e</sup> Whole-Genome Research Core Laboratory of Human Diseases, Chang Gung Memorial Hospital, Keelung 204, Taiwan

<sup>f</sup> Department of Otorhinolaryngology-Head and Neck Surgery, Changhua Christian Hospital, Changhua 500, Taiwan

<sup>g</sup> Department of Internal Medicine, Division of Endocrinology and Metabolism, Chung Shan Medical University Hospital, Taichung 402, Taiwan

<sup>h</sup> School of Medicine, Chung Shan Medical University, Taichung 402, Taiwan

<sup>i</sup> Department of Otolaryngology, Chung Shan Medical University Hospital, Taichung 402, Taiwan

<sup>j</sup> Biotools, Co., Ltd, New Taipei City 221, Taiwan

<sup>k</sup> School of Medicine, College of Medicine, Chang Gung University 333, Taoyuan, Taiwan

<sup>l</sup> Chang Gung Immunology Consortium, Chang Gung Memorial Hospital and Chang Gung University, Taoyuan, Taiwan

<sup>m</sup> Department of Dermatology, Drug Hypersensitivity Clinical and Research Center, Chang Gung Memorial Hospital, Linkou, 333, Taiwan

### ARTICLE INFO

#### Keywords:

Oral squamous cell carcinoma  
Oral microbiota  
Mutational signature  
16S rRNA

### ABSTRACT

**Objectives:** Both genetic and environmental factors are conceivably required to assess the prognosis of oral squamous cell carcinoma (OSCC), yet little is known regarding the relationship between oral microbiome and the mutational spectrum of OSCC.

**Materials and methods:** Here, we used 16S rRNA amplicon sequencing to study the composition of oral microorganisms in OSCC patients, whose cancer mutational profiles were previously defined by whole-exome sequencing, to evaluate the relationship between oral microbiome and the mutational changes in OSCC.

**Results:** Analyzing the contributions of the five mutational signatures extracted from the primary tumors revealed three groups of OSCC (mutational signature cluster, MSC1-3) that were significantly associated with demographic and clinical features. Taxonomic analysis of the predominant phyla in salivary samples showed variation in the relative abundance of *Firmicutes* and *Bacteroidetes* in the three MSC groups. In addition, significant differences in bacterial species richness (alpha diversity) and slight sample-to-sample dissimilarities in bacterial community structures (beta diversity) were noted among different MSC groups. Further, predicting the functional capabilities of microbial communities by reconstruction of unobserved states showed that many pathways related to cell motility were differentially enriched among the three MSC groups.

**Conclusion:** Collectively, these results indicate a potential association of oral microbiome with the mutational changes in OSCC.

### Introduction

Oral cancer is a prevalent malignancy globally, with a huge majority (~90%) of cases being oral squamous cell carcinoma (OSCC) [1]. In spite of advances in etiological studies and therapeutic options, the mortality of OSCC has not improved substantially over the past decades [2]. It is known that OSCC is a multifactorial neoplasm in which genetic

variants interact with environmental triggers in the predisposition to this disease [3]. Various genetic alterations that influence cell cycle, apoptosis, and DNA repair [4] alone or in combination with external risks, including human papillomavirus (HPV) infection and habitual exposure of carcinogens, such as tobacco and alcohol use and betel nut chewing [5] have been demonstrated to contribute to the etiology and pathogenesis of oral cancer. Moreover, other possible risks of oral

<sup>☆</sup> Association of oral microbiota with OSCC genomics.

\* Corresponding author at: Whole-Genome Research Core Laboratory of Human Diseases, Chang Gung Memorial Hospital, No. 200, Ln 208, Jijin 1stRd, Keelung 204, Taiwan.  
E-mail address: [ssu1@cgmh.org.tw](mailto:ssu1@cgmh.org.tw) (S.-C. Su).

tumorigenesis comprise but not limited to poor oral hygiene, periodontitis [6], and chronic microbial infections and inflammation [7].

Alterations in oral microbiome may disturb the relationship between microorganisms and humans, potentially leading to disease. Increasing evidence has indicated a role for oral microbiota in OSCC through direct metabolism of carcinogens and inflammatory effects [8]. Earlier studies of oral microbiome using culture-dependent methods examined only small numbers of species and thus provided no consensus regarding cancer-associated changes in species abundance [9,10]. Strikingly, the wide spectrum of oral microbiota is consisting of more than 600 bacterial species, the majority of which, however, are uncultivated [11]. The advent of high-throughput sequencing technology has enabled the use of recently developed, culture-independent strategies for exploring the composition of microbial ecosystems in human health [12]. The culture-independent nature of this approach does not rule out species that are difficult to grow or currently uncultivated, and its open-ended and highly sensitive nature potentiates the discovery of new taxa of low abundance that are likely associated with oral cancer.

Recent studies using the small subunit ribosomal RNA (16S rRNA) gene sequence in culture-free settings have compared the microbial communities present in oral cancer versus anatomically matched normal tissues, OSCC biopsies versus deep-epithelium swabs from matched control subjects, and saliva samples from OSCC patients with different statuses of HPV positivity and treatment [13–15]. Although mostly inconsistent, a significant reduction in the abundance of Firmicutes (particularly *Streptococcus* and *Haemophilus*) was observed in oral cancer as compared to normal samples. It is conceivable that genetic changes and environmental factors are interrelated and required to assess the prognosis of OSCC. However, at present, little is known regarding the relationship between oral microbiome and the mutational changes in OSCC. Using whole-exome sequencing, we have previously characterized the mutational catalog of oral cancer and provided clues for a molecular-based classification of OSCC patients with prognostic and therapeutic implications [16]. In this study, we analyzed the composition of oral microorganisms in OSCC patients with different mutational profiles to evaluate the potential association between the microbial flora of oral cavity and genomic alterations of oral cancer.

## Materials and Methods

### Subject enrollment and sample collection

The subjects were recruited from 2014 to 2015, with the approval by the institutional review board of Chung Shan Medical University Hospital, Taichung, Taiwan. Patients with any history of diabetes mellitus or immune-related diseases were excluded. All participants were free of antibiotics therapy within three months and provided informed written consent at enrollment. Of 123 eligible cases, 103 participated in the study. Salivary samples of 39 male patients, whose exome had been previously sequenced [16] and who had been neither previously treated nor proven metastatic disease at the time of diagnosis, were collected within 1–3 weeks after disease diagnosis in this study. TNM staging, mutational profiling, and data on age, gender, alcohol drinking, betel quid chewing, and cigarette smoking were assessed and reported previously [16]. Non-stimulated saliva was obtained in the morning between 8 and 10 a.m. from each patient using a prior protocol [17]. In brief, OSCC patients were asked to refrain from eating, drinking, smoking, or oral hygiene procedures for at least one hour before sample collection. Saliva was allowed to accumulate in the floor of the mouth and the subject spat it out into the preweighed test tube every 60 s for a period of 10 min. Salivary samples in sterile Falcon tubes were vortexed, then aliquoted into sterile cryogenic tubes, and stored at  $-80^{\circ}\text{C}$  until analysis.

### 16S rRNA gene amplification and sequencing

Before DNA extraction, salivary samples were centrifuged at 14,000 rpm for 2 min to pellet the bacterial communities [18]. The pellet was re-suspended and incubated with proteinase K (Sigma-Aldrich) at  $55^{\circ}\text{C}$  overnight. Bacterial genomic DNA was then isolated with a Qiamp DNA Blood Mini kit (Qiagen) according to the manufacturer's instructions and quantified by a NanoPhotometer N50 (Implen). The variable region 4 (V4) of small subunit rRNA (16S rRNA) gene was PCR-amplified using the primer set described previously [19]. Gel electrophoresis of PCR products on 2% agarose gels was performed for quality control. Amplicons were purified using the AMPure XP PCR Purification Kit (Agencourt) and then quantified using a Qubit dsDNA HS Assay Kit (Qubit) on a Qubit 2.0 Fluorometer (Qubit). To construct a library, the Illumina sequencing adapters were ligated to the purified amplicons by a second-stage PCR using the TruSeq DNA LT Sample Preparation Kit (Illumina). Purified libraries were quantified, normalized, pooled, and applied for cluster generation and sequencing on a MiSeq instrument (Illumina).

### Processing and analysis of sequence data

Paired-end reads were merged using FLASH v1.2.7 [20]. Quality filtering of reads was assessed using the QIIME 1.7 pipeline [21] and chimeric sequences were removed by UCHIME [22]. The processed sequencing reads (effective tags) were clustered into operational taxonomic units (OTU) at 97% sequence identity using the UPARSE [23], and taxonomy classification was assigned according to the information retrieved from the Greengenes database [24]. Any sequence with one-time occurrence (singletons) or detected in only one sample was filtered out, and samples with less than  $10^4$  effective tags were excluded from further analyses. To evaluate the phylogenetic relationship of different OTUs, alignment of multiple sequences was conducted using the PyNAST software v.1.2 [25] against the core-set dataset of the Greengenes database, and a phylogenetic tree was generated with the FastTree [26].

Prior to subsequent analysis of alpha and beta diversities, data regarding OTU abundance were rarefied to the minimum sequence depth to normalize the variations in sequence depth across samples. For estimating alpha diversity, species richness was evaluated by the Chao1 and abundance-based coverage estimators (ACE) indices. A rarefaction curve was generated by a random selection of certain amount of sequencing data from each sample for representing the number of the observed species, and a species accumulation curve was plotted by the occurrence rate of new OTUs (species) under continuous sampling. For evaluating beta diversity, the weighted and unweighted UniFrac parameters [27] were calculated by using the QIIME pipeline. Principal coordinate analysis (PCoA) was conducted using the weighted correlation network analysis (WGCNA), stat, and ggplot2 packages in R software by transforming a distance matrix of weighted or unweighted UniFrac parameters among samples into a new set of orthogonal axes.

Functional composition of metagenomes was predicted from 16S rRNA data by the PICRUSt software [28], the pipeline of which is composed of two workflows, gene content prediction and metagenome prediction. A table of gene copy numbers for each gene family in each sequenced bacterial and archaeal genome based on the IMG database [29] and a phylogenetic tree from the Greengenes database [24] were recomputed for gene content prediction. Subsequently, metagenome prediction was performed through multiplying the vector of gene counts for each OTU by the abundance of that OTU in each sample, and summed across all OTUs.

### Statistical analysis

The significance of association between the clusters and the clinical data was tested using Fisher's exact test. The relative abundance of the

taxa present in each cluster was compared at the phylum and genus levels using Welch's t test. Continuous variables, such as chao1 and ACE index, were analyzed using the Kruskal-Wallis test. Data were analyzed by using the SAS statistical software (Version 9.1, 2005; SAS Institute Inc., Cary, NC). Statistically significant biomarkers were evaluated by the linear discriminant analysis (LDA) of effect size (LEfSe) analysis [30], which employed the non-parametric factorial Kruskal-Wallis test, Wilcoxon rank-sum test and LDA to identify differentially abundant taxa between two metadata classes. High LDA scores reflect significantly higher abundance of certain taxa. All reported *p* values were two-tailed, and a *p* value of < 0.05 was considered significant.

## Results

### Characterization of study participants

The study cohort was composed of 39 OSCC patients, all of whom were male. The average age at disease onset was 53 (Table S1). Various anatomical sites were represented, including buccal mucosa (46.2%), tongue (20.5%), lip (15.4%), gingiva (10.3%), and others (7.6%). Of the participants profiled in this study, 87.2%, 79.5%, and 94.9% reported a history of tobacco, alcohol, and betel nut use, respectively. Surgical dissection and histological examination confirmed advanced stage III/IV in 48.8% of cases. Lymph node metastases occurred in 25.6% of patients.

### Bacterial profiling by 16S rRNA gene sequencing

A total of 19,01,329 sequencing reads, ranging from 6781 to 1,77,645 per sample, was generated from the V4 hypervariable region of 16S rRNA gene. After strict quality and size filtering, eight samples were excluded from further analyses due to low read numbers, such that a total of 1050 OTUs was assigned (per sample range 283–583) (Table S2). Rarefaction curves show that a plateau of species richness (up to 400 OTUs) was achieved in approximately 16,000 reads per sample (Fig. S1). To control the sample heterogeneity, roughly 16,000 reads were used hereafter as the minimum sampling depth to evaluate the diversity. In addition, the shape of the species accumulation curve generated from our dataset reveals that the community was well sampled as the specimens we collected contained considerable information about total species richness (Fig. S2). Taxonomic analyses identified a predominance of phylum *Firmicutes* (37.7%; that is, 37.7% of the total number of sequences obtained), *Bacteroidetes* (20.3%), *Proteobacteria* (16.5%), and *Actinobacteria* (10.4%), with less frequent presence of *Fusobacteria* (6.8%) in the saliva of OSCC patients (Fig. S3). At the lower taxonomic level, the most abundant genera were *Streptococcus* (16.2%), *Prevotella* (13.4%), *Neisseria* (9.1%), and *Veillonella* (8.9%), with lower numbers of *Rothia* (5.4%) and *Porphyromonas* (4.2%).

### Microbial composition and abundance categorized by the distributions of mutational signatures within OSCC

To explore the potential relationship between bacterial flora of oral cavity and mutational changes in oral cancer, we categorized OSCC patients based on the distributions of five mutational signatures extracted from the dataset of somatic variations in the primary tumors of our cohort [16]. 31 OSCC tumors were classified into three groups (mutational signature cluster, MSC1-3) and a singleton based on the intensity of the mutational signatures operative in each tumor using an unsupervised hierarchical clustering method (Fig. 1A). Tumors in the MSC3 group (*n* = 11) predominantly presented at the late stage (*p* = .006) and with a greater tumor size (*p* = .021) (Fig. 1B). The proportion of patients with a history of smoking in MSC3 was lower than that for the other groups (*p* = .021). The MSC1 group (*n* = 9) was characterized by frequent *TP53* mutations (*p* = .021). In addition, the MSC2 group (*n* = 10) exhibited a late histological stage (*p* = .017) and

was associated with recurrent mutations in key driver genes (*FAT1*, *FZR1*, and *AXIN1*) related to the Wnt pathway (*p* = .032). These findings suggest that molecular classification of OSCC according to the intensity of mutational patterns in oral cancer reflects clinicopathological parameters and etiological backgrounds.

Further analysis of bacterial composition and abundance categorized by the distributions of mutational signatures in OSCC revealed that the relative abundance of many taxa at the phylum and genus levels differed among the three MSC groups (Fig. 1C, Fig. 1D, and Fig. S4). Specifically, for phyla with a relative abundance higher than 1%, patients in the MSC2 group harbored higher levels of *Firmicutes* than did those in the other groups (Fig. 1C and Fig. S5). Significant differences in relative abundance of phylum *Bacteroidetes* and *Synergistetes* were observed between the MSC2 and MSC3 groups. At the lower taxonomic level, comparisons of dominant genera (> 1%) showed that *Selenomonas* and *Rothia* distinguish MSC2 from MSC1 and MSC3, respectively (Fig. 1D and Fig. S6). In addition, a significantly higher prevalence of genus *Capnocytophaga* was found in MSC3 than that in the other groups.

We also sought to predict the biomarker for each category by taking statistical significance and biological consistency into consideration using linear discriminant analysis (LDA) of effect size (LEfSe). Accordingly, we observed a profusion of *Firmicutes* in the saliva of the MSC2 over the MSC1 group (Fig. 2). On the contrary, samples in the MSC1 group had significant enrichment for phylum *Actinobacteria*, class *Actinobacteria*, and order *Actinomycetales* as compared with those in the MSC2 group.

### Diversity of microbiota associated with molecular OSCC subgroups

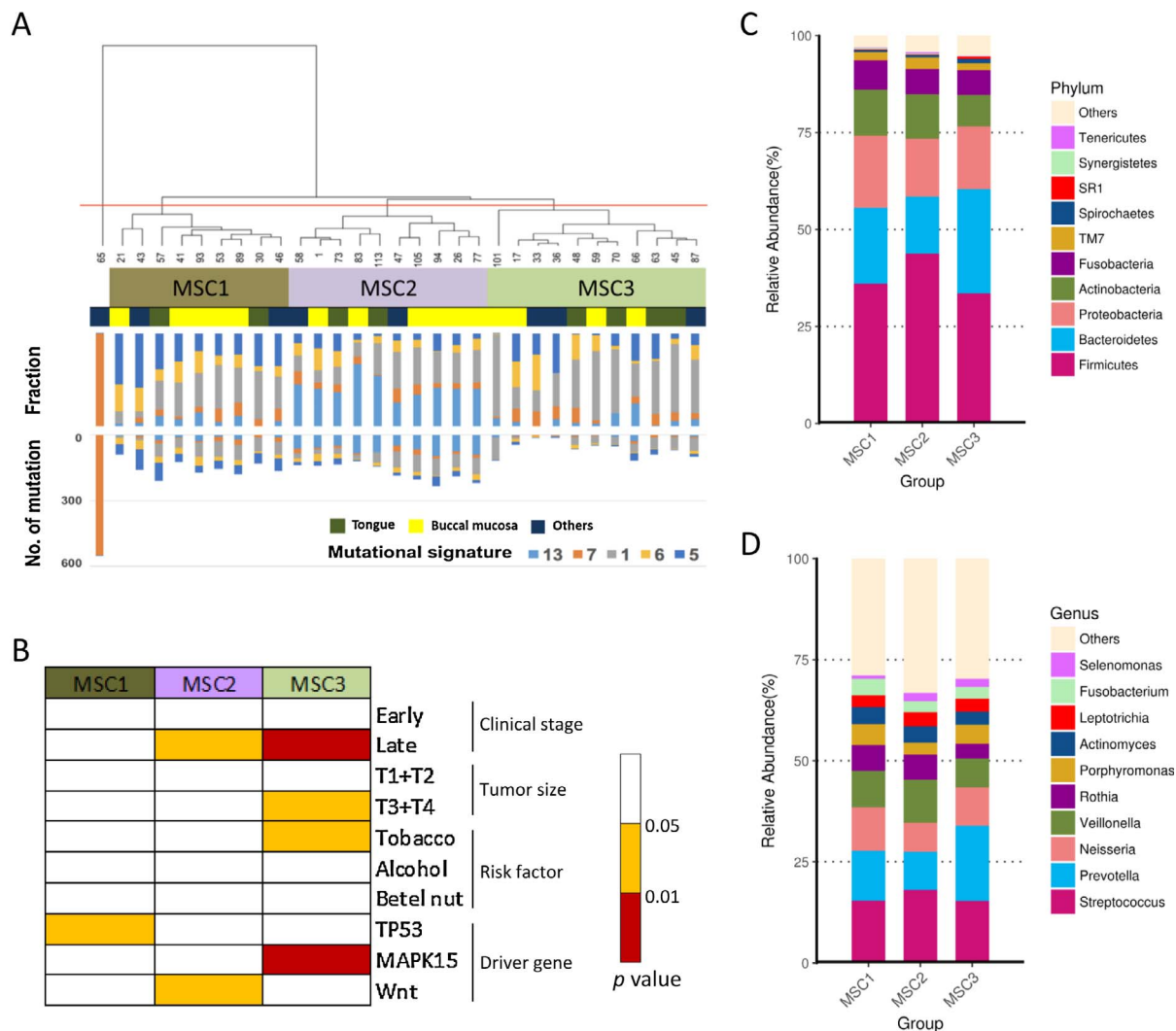
In addition to changes in relative abundance of different taxa, we also evaluated alterations in microbial diversity among three molecular OSCC subgroups. Significant differences in bacterial species richness and evenness (alpha diversity) were detected between the MSC2 and MSC3 groups but not between subgroups assigned with regard to the mutation rate or the status of *TP53* alterations (Fig. 3). Although not statistically significant, a marginal decrease in alpha diversity in the MSC1 group compared to the MSC2 was also observed. Moreover, analyses of sample-to-sample dissimilarities in bacterial community structures (beta diversity) demonstrated a slight discrimination among three MSC groups, especially between MSC2 and MSC3, but not between tumors positive and negative for *TP53* mutations (Fig. 4).

### Functional prediction of bacterial communities related to mutational changes in OSCC

To gain an insight into the functionality of oral microbiota in OSCC biology, we inferred the functional profile of microbial communities by phylogenetic reconstruction of unobserved states (PICRUSt) [28]. We found that many pathways relevant to cell motility were differentially enriched among the three MSC subgroups (Fig. 5A). Notably, as for MSC2 and MSC3, between which bacterial composition and diversity differed profoundly, the genetic markers with significant discriminative power in many pathways assigned to environmental information processing, including membrane transport (transporters, phosphotransferase system, and ABC transporters), signal transduction (two-component system), and signaling molecules and interaction (cellular antigens) were detected (Fig. 5B). Collectively, data shown in the present study indicate a potential association of oral microbiome with the mutational changes in OSCC.

## Discussion

Disturbance in the relationship between microbiota and humans is known to be central to various physiological and pathological conditions. Recent studies have demonstrated associations of host genetics with the intestinal microbiome, revealing another level of host-microbe



**Fig. 1.** Taxonomic profiles, at the phyla and genus levels, of salivary samples from OSCC patients classified by the contributions of mutational signatures. (A) Five mutational signatures were previously deciphered in the OSCC cohort [16] and corresponding to the updated consensus signatures (#1, 5, 6, 7, and 13) [33]. Tumor classification (mutational signature cluster, MSC) based on the contributions of mutational signatures. The fraction and number of mutations attributed to each signature in each tumor, together with location of tumors, are represented by the colored bars below the dendrogram. (B) Clinicopathological and etiological characteristics associated with each MSC group. (C–D) The distribution of top 10 phyla (C) and top 10 genera (D) detected in the OSCC subgroups.

interaction [31,32]. In the present study, we showed changes in abundance and diversity of oral microorganisms among different OSCC subgroups clustered by the distribution of mutational patterns. We also unraveled that microbial gene pathways were differentially enriched in salivary samples of OSCC patients with distinct mutational profiles.

All cancers carry somatic mutations that were accumulated through multiple mutational processes. Different mutational processes give rise to unique combinations of mutation types, termed mutational signatures [33]. Our previous large-scale analysis of OSCC genomes detected five mutational signatures operative in our cohort [16]. Here, we showed that changes in composition and abundance of oral microbiota were associated with OSCC subgroups categorized by the distributions of mutational signatures. It is known that microorganisms and their products activate fibroblasts and immune cells to generate reactive oxygen species that elicit DNA damage in epithelial cells [34]. In addition, microbial genotoxins (cytotoxic distending toxin, typhoid toxin, and colibactin) or metabolites (hydrogen sulfide, ammonia and fatty acids) may directly target to DNA and induce mutations. These suggest that oral microbiota, as an exogenous trigger, may be involved in certain mutational processes in oral tumorigenesis, ultimately contributing to differential intensities of mutational patterns in oral cancer. Notably, such host-microbe relationship we observed among different

mutational signature clusters was not seen while tumors were classified with regard to the mutation rate and the status of *TP53* alterations (Fig. 3 and 4), implying that the genetic alterations mediated by oral microbiota may be induced at a relatively even rate across samples and dispersed in various tumor suppressor genes and proto-oncogenes.

Consistent with other reports [13,14], the most predominant phyla of oral microbiota detected in the present study were *Firmicutes* (37.7%), *Bacteroidetes* (20.3%), *Proteobacteria* (16.5%), and *Actinobacteria* (10.4%). This is not in concordance with the finding from a recent study [15] where *Fusobacteria* was shown to be the most abundant taxa at the phylum level in OSCC biopsies. In addition to a potential regional impact derived from OSCC etiology, such difference in bacterial associations with oral cancer may be partly accounted for by the types of samples analyzed. Moreover, the relative abundance of many taxa was found to vary within the three OSCC subgroups (Fig. 2 and Fig. S4–6). Among these, *Actinobacteria* was identified as a biomarker for the MSC1 group by LEfSe in our dataset (Fig. 2). In a recent investigation on the oral mucosal and salivary microbial communities, members of *Actinobacteria*, especially the *Rothia* spp., have been found to be associated with Behçet's syndrome, a multisystemic immune-related condition characterized by recurrent oral ulceration [35]. Additionally, significant differences in relative levels of phylum

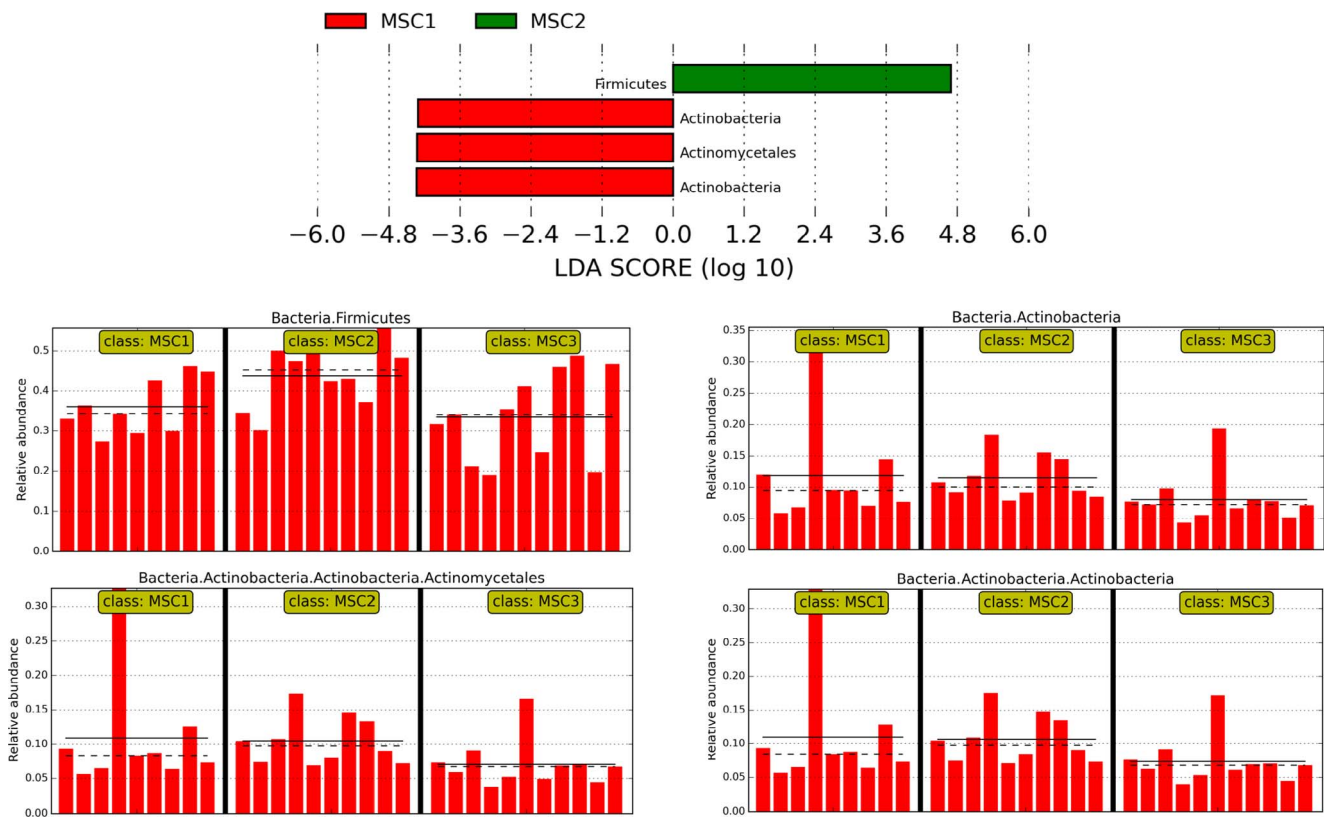


Fig. 2. Determination of bacterial biomarkers for each molecular subgroup. Taxa that best characterize each OSCC subgroups were identified by using linear discriminant analysis of effect size (LEfSe) on OTU tables (Upper panel). The relative abundance of each taxon across all samples is shown in the lower panel where the straight and dotted lines plot the means and medians of the relative abundance, respectively, in each subgroup.

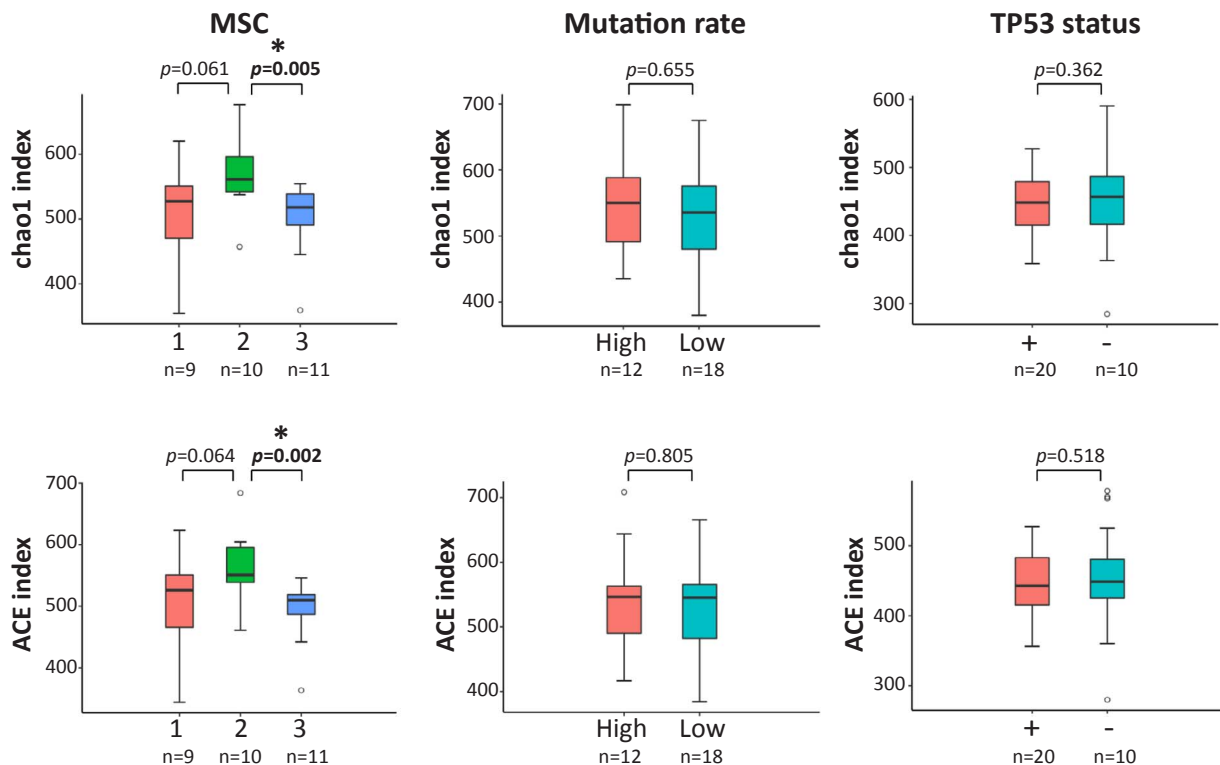


Fig. 3. Comparisons of bacterial species richness in saliva of OSCC patients categorized by the distribution of mutational signatures (MSC), mutation rate, and status of TP53 alterations. Bacterial species richness was estimated by the Chao1 and ACE indexes. The box-plot shows the median, the 25th, and the 75th percentile in each molecular classification. The boundary between the high and low mutation rate group is the average of somatic mutations in our cohort (mean = 154), \*  $p < .05$ , Kruskal-Wallis test.

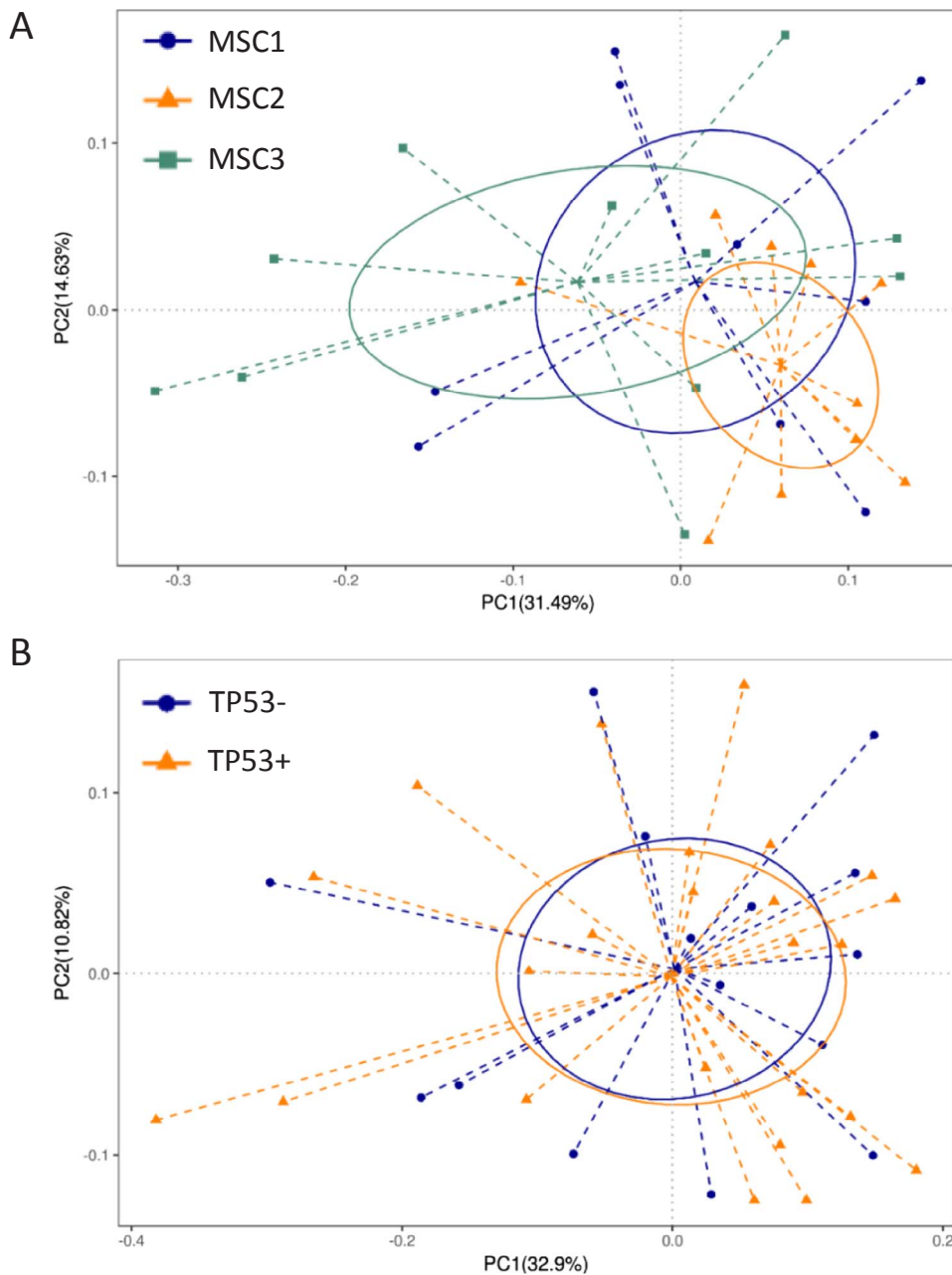


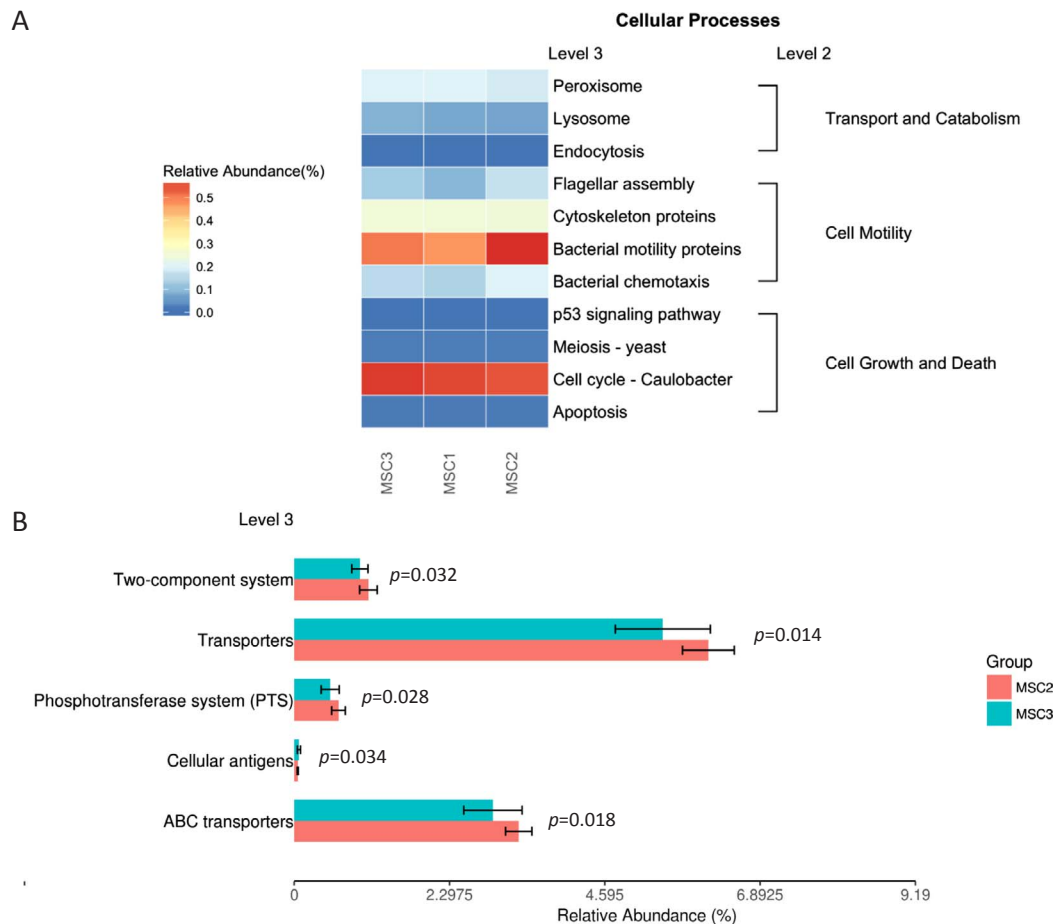
Fig. 4. Discrimination among different molecular groups of OSCC. Principal coordinate analysis (PCoA) based on Weighted UniFrac distance between samples categorized by clustering of mutational signatures (A) and status of TP53 alterations (B). The percentage in each axis indicates how much variation explained.

*Synergistetes* were observed between the MSC2 and MSC3 groups (Fig. S5). Association of salivary *Synergistetes* with periodontitis has been documented [36], as periodontitis is considered a risk factor for oral cancer [6]. Of note, *Capnocytophaga* species, whose levels in the MSC3 group were higher than that in the other groups (Fig. S6), were known to potentially cause bloodstream infections in patients with severe oropharyngeal mucositis or periodontal disease [37]. Intriguingly, although less abundant, salivary samples of the MSC3 group were more highly colonized with *Bifidobacterium* spp. compared to the MSC1 group (MSC3, 0.116%; MSC1, 0.027%;  $p = .029$ ), while oral administration of *Bifidobacterium* has been shown to promote antitumor immunity and facilitate the efficacy of immunotherapy with check-point inhibitors by altering the composition of intestinal microorganisms [38]. Our results together with findings from others implicate a usefulness of oral microbiota in not only characterization of OSCC etiology but also stratification of OSCC patients with differential therapeutic efficacy.

Previous investigations showed changes in alpha and beta diversity of bacterial communities between salivary specimens of OSCC patients

and that from control subjects [13,15,39], whereas the bacterial diversity of tissue biopsy or swab samples between OSCC and controls remained similar [14,40]. Here, we observed alterations in bacterial diversity in saliva obtained from OSCC patients with distinct mutational profiles (Fig. 3 and 4), suggesting a better discriminative potential for saliva in analyzing OSCC-associated oral microbiota in comparison with other sample types. Nevertheless, difference in microbial diversity was not seen in saliva from OSCC patients categorized by mutation rate, the status of TP53 mutation, or the location of oral cancer (Fig. S7). In addition, several functional pathways related to cell motility were predicted to be differentially enriched among the three MSC subgroups (Fig. 5A). Bacterial flagella were recognized as key inflammatory structures [41], and defects in bacterial chemotaxis were shown to dampen the inflammation [42]. These restate a role of oral microbiota in regulating cancer-related inflammation and connect genetic variations of OSCC to oral flora-mediated immunity.

To explore the relationship between oral microorganisms and the genomic landscapes of OSCC, additional work is needed to address



**Fig. 5.** Prediction of microbial gene functions among distinct mutational signature clusters of OSCC. (A) Pathway enrichment for KEGG cellular processes in three MSC groups. (B) Difference in relative abundance of predicted microbial genes related to environmental information processing between MSC2 and MSC3 using Student's t test.

several limitations of the present study. One issue is the limited sample size, which likely underestimated the potential associations of salivary microbiome with the alteration status of a given driver gene or pathway in oral cancer. In addition, hyposalivation often occurs in patients with oral diseases [43] and could shift the quality of saliva towards a more acidogenic microflora [44]. The association of changes in the oral microbiota with hyposalivation was shown to be related to the reason for the hyposalivation rather than to the magnitude of reduced salivary flow rate [45]. A relation between salivary flow and oral symptoms in patients with advanced terminal cancer has been observed [46]. Yet, salivary flow rate was not taken in consideration in the present study and its variations could be a potential caveat of the current study design. Also unavailable was the status of oral hygiene and periodontitis. Mounting evidence has indicated the potential association of OSCC risk with oral health parameters, such as tooth loss, cavity, gum bleeding, tartar deposits, and periodontitis [47–49]. These OSCC-associated oral conditions collectively perturb the finely-tuned equilibrium of the oral ecosystem and may be the additional confounding factors in our analysis. Another drawback is that the functional capacities of bacterial communities were inferred based on 16S rRNA gene sequencing. Further defining the network of diverse pathways employed in OSCC-associated oral microbiota will require a comprehensive analysis of whole metagenome in future investigations. Moreover, our data failed to define whether oral microbiome has altered during the development of OSCC or after the formation of tumors.

Overall, we showed that oral microbiota is compositionally and functionally associated with the mutational changes in oral cancer. Our study provides potential avenues for the understanding of oral tumorigenesis and the identification of OSCC patients who could likely

benefit from individual treatment regimens through the integration of cancer genomics and metagenomics.

## Acknowledgements

We thank Dr. Wen-Hsiung Li, Academia Sinica, Taipei, Taiwan, for reviewing this manuscript. This study was supported by research grants from Health and welfare surcharge of tobacco products, Ministry of Health and Welfare, Taiwan (MOHW103-TD-B-111-08, MOHW104-TDU-B-212-124-005, MOHW105-TDU-B-212-134002, MOHW106-TDU-B-212-144005), from the Ministry of Science and Technology, Taiwan (NSC-102-2314-B-040-008-MY3) and from Chung Shan Medical University Hospital, Taiwan (CSH-2017-E-001-Y2) to SFY, and from Chang Gung Memorial Hospital (CMRPG2F0421) to SCS. We are grateful to Yu-Jr Lin from the Chang Gung Memorial Hospital, Linkou, Taiwan (CIRPD1D0031) for the assistance in statistics. Special thanks to the Research Services Center For Health Information, Chang Gung University, Taoyuan, Taiwan.

## Role of funding source

The research sponsors were not involved in study design, the collection, analysis and interpretation of data, the writing of the report, and the decision to submit the article for publication.

## Conflict of interest

None declared.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.oraloncology.2017.12.005>.

## Reference

- Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, et al. Cancer statistics, 2008. *CA Cancer J Clin* 2008;58:71–96.
- Zini A, Czerninski R, Sgan-Cohen HD. Oral cancer over four decades: epidemiology, trends, histology, and survival by anatomical sites. *J Oral Pathol Med: Off Publ Int Assoc Oral Pathol Am Acad Oral Pathol* 2010;39:299–305.
- Krishna Rao SV, Mejia G, Roberts-Thomson K, Logan R. Epidemiology of oral cancer in Asia in the past decade—an update (2000–2012). *Asian Pacific J Cancer Prevent: APJCP* 2013;14:5567–77.
- Scully C, Field JK, Tanzawa H. Genetic aberrations in oral or head and neck squamous cell carcinoma (SCCHN): 1. Carcinogen metabolism, DNA repair and cell cycle control. *Oral Oncol* 2000;36:256–63.
- Scully C, Bagan J. Oral squamous cell carcinoma overview. *Oral Oncol* 2009;45:301–8.
- 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 Biomark Prevent: Publ Am Assoc Cancer Res Cospons Am Soc Prevent Oncol* 2009;18:2406–12.
- Hooper SJ, Wilson MJ, Crean SJ. Exploring the link between microorganisms and oral cancer: a systematic review of the literature. *Head Neck* 2009;31:1228–39.
- Wang L, Ganly I. The oral microbiome and oral cancer. *Clin Lab Med* 2014;34:711–9.
- Hooper SJ, Crean SJ, Lewis MA, Spratt DA, Wade WG, Wilson MJ. Viable bacteria present within oral squamous cell carcinoma tissue. *J Clin Microbiol* 2006;44:1719–25.
- Nagy KN, Sonkodi I, Szoke I, Nagy E, Newman HN. The microflora associated with human oral carcinomas. *Oral Oncol* 1998;34:304–8.
- Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner AC, Yu WH, et al. The human oral microbiome. *J Bacteriol* 2010;192:5002–17.
- Human Microbiome Project. C. A framework for human microbiome research. *Nature* 2012;486:215–21.
- 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;7:51320–34.
- Schmidt BL, Kuczynski J, Bhattacharya A, Huey B, Corby PM, Queiroz EL, et al. Changes in abundance of oral microbiota associated with oral cancer. *PLoS One* 2014;9:e98741.
- Al-Hebshi NN, Nasher AT, Maryoud MY, Homeida HE, Chen T, Idris AM, et al. Inflammatory bacteriome featuring *Fusobacterium nucleatum* and *Pseudomonas aeruginosa* identified in association with oral squamous cell carcinoma. *Sci Rep* 2017;7:1834.
- Su SC, Lin CW, Liu YF, Fan WL, Chen MK, Yu CP, et al. Exome sequencing of oral squamous cell carcinoma reveals molecular subgroups and novel therapeutic opportunities. *Theranostics* 2017;7:1088–99.
- Li Y, St John MA, Zhou X, Kim Y, Sinha U, Jordan RC, et al. Salivary transcriptome diagnostics for oral cancer detection. *Clin Cancer Res: Off J Am Assoc Cancer Res* 2004;10:8442–50.
- Weng SL, Chiu CM, Lin FM, Huang WC, Liang C, Yang T, et al. Bacterial communities in semen from men of infertile couples: metagenomic sequencing reveals relationships of seminal microbiota to semen quality. *PLoS One* 2014;9:e110152.
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 2011;108(Suppl 1):4516–22.
- Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;27:2957–63.
- 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;7:335–6.
- Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;27:2194–200.
- Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 2013;10:996–8.
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 2006;72:5069–72.
- Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 2010;26:266–7.
- Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;26:1641–50.
- Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. *ISME J* 2011;5:169–72.
- 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;31:814–21.
- Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, et al. IMG: the integrated microbial Genomes database and comparative analysis system. *Nucl Acids Res* 2012;40:D115–22.
- Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60.
- Turpin W, Espin-Garcia O, Xu W, Silverberg MS, Kevans D, Smith MI, et al. Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat Genet* 2016;48:1413–7.
- Wang J, Thingholm LB, Skievcicene J, Rausch P, Kummen M, Hov JR, et al. Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat Genet* 2016;48:1396–406.
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature* 2013;500:415–21.
- Schwabe RF, Jobin C. The microbiome and cancer. *Nat Rev Cancer* 2013;13:800–12.
- Seoudi N, Bergmeier LA, Drobniewski F, Paster B, Fortune F. The oral mucosal and salivary microbial community of Behcet's syndrome and recurrent aphthous stomatitis. *J Oral Microbiol* 2015;7:27150.
- Marchesan JT, Morelli T, Moss K, Barros SP, Ward M, Jenkins W, et al. Association of synergistetes and cyclodipeptides with periodontitis. *J Dent Res* 2015;94:1425–31.
- Martino R, Ramila E, Capdevila JA, Planes A, Rovira M, Ortega M, et al. Bacteremia caused by capnocytophaga species in patients with neutropenia and cancer: results of a multicenter study. *Clin Infect Dis: Off Publ Infect Dis Soc Am* 2001;33:E20–2.
- Sivan A, Corrales L, Hubert N, Williams JB, Aquino-Michaels K, Earley ZM, et al. Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science* 2015;350:1084–9.
- Pushalkar S, Mane SP, Ji X, Li Y, Evans C, Crasta OR, et al. Microbial diversity in saliva of oral squamous cell carcinoma. *FEMS Immunol Med Microbiol* 2011;61:269–77.
- Pushalkar S, Ji X, Li Y, Estilo C, Yegnanarayana R, Singh B, et al. Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. *BMC Microbiol* 2012;12:144.
- Rastogi D, Ratner AJ, Prince A. Host-bacterial interactions in the initiation of inflammation. *Paediatr Respirat Rev* 2001;2:245–52.
- Williams SM, Chen YT, Andermann TM, Carter JE, McGee DJ, Ottemann KM. *Helicobacter pylori* chemotaxis modulates inflammation and bacterium-gastric epithelium interactions in infected mice. *Infect Immun* 2007;75:3747–57.
- Meurman JH, Gronroos L. Oral and dental health care of oral cancer patients: hyposalivation, caries and infections. *Oral Oncol* 2010;46:464–7.
- Tschoppe P, Wolgin M, Pischon N, Kielbassa AM. Etiologic factors of hyposalivation and consequences for oral health. *Quintessence Int* 2010;41:321–33.
- Almstah IA, Wikstrom M, Stenberg I, Jakobsson A, Fagerberg-Mohlin B. Oral microbiota associated with hyposalivation of different origins. *Oral Microbiol Immunol* 2003;18:1–8.
- Chausu G, Bercovici M, Dori S, Waller A, Taicher S, Kronenberg J, et al. Salivary flow and its relation with oral symptoms in terminally ill patients. *Cancer* 2000;88:984–7.
- Guha N, Boffetta P, Wunsch Filho V, Eluf Neto J, Shangina O, Zaridze D, et al. Oral health and risk of squamous cell carcinoma of the head and neck and esophagus: results of two multicentric case-control studies. *Am J Epidemiol* 2007;166:1159–73.
- Talamini R, Vaccarella S, Barbone F, Tavani A, La Vecchia C, Herrero R, et al. Oral hygiene, dentition, sexual habits and risk of oral cancer. *Br J Cancer* 2000;83:1238–42.
- Divaris K, Olshan AF, Smith J, Bell ME, Weissler MC, Funkhouser WK, et al. Oral health and risk for head and neck squamous cell carcinoma: the Carolina head and neck cancer study. *Cancer Causes Control: CCC* 2010;21:567–75.