

# Dysbiotic microbiota to complement TNM staging or HPV status for prognostic indication of head and neck cancer

**Yu Dou**

Shandong University <https://orcid.org/0000-0001-7698-6791>

**Chao Ma**

Qilu Hospital of Shandong University

**Ketao Wang**

Qilu Hospital of Shandong University

**Shaohua Liu**

Qilu Hospital of Shandong University

**Jintang Sun**

Qilu Hospital of Shandong University

**Wanye Tan**

Qilu Hospital of Shandong University

**Markus Neckenig**

Shandong University

**Qingjie Wang**

Qilu Hospital of Shandong University

**Zuoqing Dong**

Qilu Hospital of Shandong University

**Wenjuan Gao**

Qilu Hospital of Shandong University

**Anwei Chen**

Qilu Hospital of Shandong University

**Dong Zhang**

Qilu Hospital of Shandong University

**Bing Wang**

Qilu Hospital of Shandong University

**Liang Shi**

Qilu Hospital of Shandong University

**Zhaodi Nan**

Qilu Hospital of Shandong University

**Dan Ai**

Qilu Hospital of Shandong University

**Wenjie Yu**

Yantai Affiliated Hospital of Binzhou Medical University

**Jia Liu**

Qilu Hospital of Shandong University

**Bingfeng Song**

Qilu Hospital of Shandong University

**Lei Zhao**

Qilu Hospital of Shandong University

**Qianqian Shao**

Qilu Hospital of Shandong University

**Yong Zhu**

Shandong University

**Tao Wang**

Qilu Hospital of Shandong University

**Jianing Wang**

Shandong University

**Xunwei Wu**

Shandong University

**Wei Hu**

Shandong University

**Fengcai Wei**

Qilu Hospital of Shandong University

**Xin Xu**

Shandong University

**Xun Qu (✉ [quxun@sdu.edu.cn](mailto:quxun@sdu.edu.cn))**

Shandong University <https://orcid.org/0000-0002-2521-3761>

---

**Research**

**Keywords:** Microbiota, Head and neck cancer, 16s rRNA, Next-generation sequence, Microbial Dysbiosis, Human Papillomavirus, Prognostic indication

**Posted Date:** September 21st, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-80076/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Microbiota has been found associated with the incidence of head and neck cancers (HNCs), however, the association of microbiota with the prognosis of HNCs remains unknown. In the present study, the relationship between tumoral microbiota and survival was examined. The prognosis predictive value of tumoral microbiota to complement classic TNM staging and Human Papillomavirus (HPV) status was also investigated.

**Results:** We conducted a retrospective study including 158 primary tumors using 16S rRNA sequencing. The tumoral microbiota in the HNC patients with poor prognosis were significantly different from that of the patients with good prognosis. A greater abundance of *Schlegelella* or *Methyloversatilis* was characterized in HNC patients with poor prognosis and a greater abundance of *Lactobacillus* or *Bacillus* was characterized in HNC patients with good prognosis. Strikingly, the microbial dysbiosis index (MDI), the ratio of above poor and good prognosis associated genera, was strongly associated with overall survival [hazard ratio (HR) 3.24; 95% confidence interval (CI), 1.88 to 5.59;  $P < 0.001$ ] and disease free survival (HR 2.11; 95% CI, 1.34 to 3.34;  $P = 0.001$ ), independently of age, TNM staging, differentiation and HPV status. Intriguingly, the combination of the MDI and HPV status significantly improved the prognostic performance of HPV status (AUC of 0.719 and 0.581 respectively,  $P = 0.024$ ). Moreover, the MDI showed higher risk for overall survival in the late stage patients (HR 3.35; 95% CI, 1.73 to 6.49;  $P < 0.001$ ) compared with the early stage patients (HR 2.41; 95% CI, 1.00 to 5.84;  $P = 0.051$ ).

**Conclusion:** Together, the MDI could serve as an applicable indicator of the prognosis of HNCs, and complements the predictive values of TNM staging and HPV status. Moreover, the findings of the higher risk of the MDI in the late stage patients supply directions for future studies that microbial modulation might be valuable in survival improvement of HNCs, especially for the late stage patients.

## Background

There are more than 834,000 new cases of head and neck cancers (HNCs) resulting in over 431,000 deaths worldwide per year[1]. Despite the development and application of improved surgical techniques, the overall five-year survival rate of HNC patients, approximately 50%, has not well been improved[2]. There is a critical and urgent need to develop effective approaches for prognostic indication to complement the efforts of traditional TNM staging, so as to stratify for precise and multi-disciplinary treatment strategies and improve the survival rate.

The oral cavity hosts a diverse community of bacteria, referred to as the oral microbiota, with more than 750 distinct bacterial taxa identified to date[3]. When the integrity of oral mucosal barrier is lost, the bacteria colonized on its surface subsequently translocate to the underlying tissues. The diversity or the composition of microbiota has been discovered to be involved in the initiation, the progression and the prognosis of various cancers, not only in gut or other cavities but also within sterile tissues[4–12]. For HNCs, there have been several studies reporting on oral microbiota[13–19] and quite limited studies

reporting on tumoral microbiota[20–25], and their associations with the incidence. However, as to the association of the microbiota with the prognosis of HNCs, no investigation has been reported.

Human papillomavirus (HPV) is an etiological factor for HNCs[26]. Patients with HPV positive HNCs generally have a better prognosis than patients with HPV negative HNCs[27–29]. However, the understanding of the distinct prognoses between HPV positive and negative HNCs is limited. It has been demonstrated that vaginal microbiota is involved in HPV acquisition and persistence and the development of cervical carcinoma[30]. However, the association of HPV status with the microbiota of HNC has not been well studied. Importantly, the complement efforts of microbiota for HPV status have not been addressed.

To gain better understanding into the association of the tumoral microbiome with overall survival and the early relapse of HNCs, we conducted a retrospective study on a followed-up cohort. Tumoral microbiota from 158 OCC and OPC patients were assessed using high-throughput sequencing of the 16S ribosomal RNA (16S rRNA) gene. The associations of the indices representing the microbial profile and the differentially abundant taxa with overall or disease free survival were analyzed. The complementary values of the microbial signatures for classical TNM staging HPV (p16) status in prognostic indication of HNCs were further determined. To understand the association of microbial signatures with prognosis of HNCs, their correlation with clinical-pathological characteristics as well as inflammatory infiltrations were analyzed.

## Methods

### Patients

A total of 158 patients admitted to Qilu Hospital of Shandong University from January 2006 to March 2015 and diagnosed with carcinoma of oral cavity (body of tongue, floor of mouth, gums, hard palate or buccal mucosa) or pharynx (root of tongue and soft palate) were included. All patients received surgical resections and were free of tumor for all margins pathologically. All patients had no history of malignance or treatments. The median age at diagnosis was 57 years old (range 17 to 80 yrs). 93 males and 65 females were included. The majority of such tumors were squamous cell carcinoma (SCC). The baselines of the clinic-pathological characteristics are shown in **Additional File 1: Table S1**. Patients were followed up until death or until May 2019 (at least 50 months for alive patients, and median 61 months for all patients).

### DNA extraction and 16S rRNA Sequencing

Bacterial genomic DNA was extracted from histology slides of tumors using TIANamp FFPE DNA Kit. The 16 s rRNA was amplified using primers (515F, 5'-GTGCCAGCMGCCGCGGTAA-3', and 806R, 5'-GGACTACHVGGGTWTCTAA T-3') targeting V4 hypervariable regions with Ion Plus Fragment Library Kit 48 rxns (Thermofisher). The amplicon libraries were conducted by incorporating adapters and sample-specific barcodes. The PCR products were purified using GeneJET (Thermo Scientific) following electrophoresis and were sequenced using Ion S5TMXL (Thermofisher).

## Sequencing and Statistical Analysis of Microbiome Data

Poor quality sequences were excluded using Cutadapt V1.9.1 to obtain raw reads. Chimeric sequences in raw reads were excluded by VSearch V2.9.0 to obtain clean reads. Filtered sequences (clean reads) were clustered into operational taxonomic units (OTUs) assuming 97% identity by Uparse V7.0.1001, and subsequently assigned to taxa according to the reference sequences of SILVA132 database. Diversity analyses were performed using QIIME (V.1.9.1). Alpha diversity (diversity within sample) was determined by observed species and Shannon index. Beta diversity (diversity between samples) was assessed by unweighted (qualitative) and weighted (quantitative) UniFrac distances and was plotted by principal coordinates analysis (PCoA). The differences in diversities were assessed by Wilcoxon test, and the similarities between groups were discriminated by permutational multivariate analysis of variance (PERMANOVA). Taxonomic differences between groups were identified using linear discriminant analysis (LDA) effect size (LEfSe). Taxa with LDA greater than 4.0 at a  $P$  value < 0.05 were considered significantly enriched. Time dependent receiver operating characteristic curve (ROC) analyses were used for the evaluation of prognosis estimating models as well as the stratification of continuous variables. Continuous variables were defined as category variables according to the optimal cutoff values of Youden's index of three-year ROC analyses (**Additional File 3: Figure S1**), except for age and tumor size categorized according to the median value, 58 years or 2.2 cm. The correlation of the observed species with differentially abundant taxa, and their correlations with the clinic-pathological characteristics were analyzed by Spearman's rank test. The prognostic values for estimating overall or disease free survival were evaluated using univariate and multivariate Cox regression. Statistical analyses and the graphs including pie charts, columns, heatmaps were carried out using R statistical software. Correlation network of the taxa at genus level was performed using Graphviz software.

## Immunohistochemistry Of P16

Deparaffinized tissue sections were incubated with monoclonal antibody against human p16 (BD Biosciences, CA, USA) at 1:50 dilution at 4 °C overnight, followed by a biotin-streptavidin horseradish peroxidase detection system (SP-9000, SPlink Detection Kit, ZSGB-bio, Beijing, China) following the manufacturer's instructions. The expression of p16 was quantified according to staining intensity (on a scale of 0–3) and positivity (0, 5, 25, 50 or 75%) of the strongest signal of each core in tumor cells.

## Klintrup-Mäkinen score

Overall inflammatory infiltration (lymphoid cells, neutrophilic and eosinophilic granulocytes) was assessed according to the Klintrup-Mäkinen (K-M) score examined in haematoxylin and eosin stained sections using a four-degree scale as previously described[31, 32]. Briefly, 0, 1, 2 and 3 indicated none, poorly, moderately and highly inflammatory infiltration respectively.

## Results

## 1. Profile and composition of the tumoral microbiota differ in poor prognosis patients and good prognosis patients

The microbiota of 158 tumors from 158 HNC patients were analyzed by next generation sequencing of the 16S rRNA gene. An average of 75730 clean reads per sample and a total of 19484 OTUs in all samples were obtained. To evaluate the profile of tumoral microbiota between distinct prognoses, we measured the observed species and Shannon index representing microbial  $\alpha$ -diversity and unweighted and weighted UniFrac distances representing microbial  $\beta$ -diversity between the patients with overall survival (OS  $\leq$  3 yrs or  $>$  3 yrs), or between the patients with or without early relapse (relapse  $\leq$  2 yrs or  $>$  2 yrs). As shown in Fig. 1a, the observed species in patients with poor overall survival (OS  $\leq$  3 yrs) were lower than that in the patients with overall survival longer than three years ( $P= 0.041$ ). Other indices of  $\alpha$ -diversity were shown in **Additional File 2: Table S2**. Furthermore, unweighted UniFrac distances amongst the patients with early or late relapse exhibited significant difference ( $P= 0.030$ ) (Fig. 1g). PERMANOVA indicated significant differences of compositions of the tumoral microbiota based on both unweighted and weighted UniFrac distances (Fig. 1b-c and Fig. 1g-h).

To identify the most relevant taxa responsible for the differences between distinct prognoses, LEfSe was conducted. In total, 16 taxa, including four genera, were characterized differentially abundant in patients with distinct prognoses. The phylum *Firmicutes*, the class *Bacilli*, the order *Bacillales*, *Lactobacillales* and *Bacteroidales*, the family *Bacillaceae* and *Lactobacillaceae* and the genera *Bacillus* and *Lactobacillus* were enriched in patients with longer OS or late relapse. By contrast, the phylum *Proteobacteria*, the class *Gammaproteobacteria*, the order *unidentified Gammaproteobacteria*, the families *Rhodocyclaceae* and *Burkholderiaceae* and the genera *Methyloversatilis* and *Schlegelella* were enriched in patients with poor overall survival or early relapse (Fig. 1d-e and Fig. 1i-j).

This is a novel finding reveals that not only the structure but also the composition of the tumoral microbiota between HNC patients with distinct prognoses are different for the first time. However, questions, such as how the microbial structure and the composition interact, whether these microbial signatures associate with the patient outcomes, and whether they could complement the prognostic values of TNM staging or HPV status, address the following analyses.

## 2. Poor prognosis patients lost the balance of tumoral microbiota

To reveal the interaction of the profile and the composition of the tumoral microbiota between distinct prognoses of HNCs, we reconstructed the correlation networks of the relative abundances of top 100 genera for poor prognosis (OS  $\leq$  3 yrs) and good prognosis (OS  $>$  3yrs) respectively. The good prognosis network showed that *Bacillus* was positively correlated with most the other genera of *Firmicutes*, *Bacteroidetes* and *Proteobacteria*, and negatively correlated with both *Methyloversatilis* and *Schlegelella* (Fig. 2b). However, in the poor prognosis network, *Bacillus* lost the most correlations and only retained a positive correlation with *Bacteroides*, on the other hand, both *Methyloversatilis* and *Schlegelella* lost most of the negative correlations, but retained their positive interactions within *Proteobacteria* (Fig. 2a).

The correlation of the observed species representing the richness of the microbiota and the 16 differentially abundant taxa was further analyzed. As shown in Fig. 3a, the poor prognosis differential taxa were positively correlated with each other within phylum *Proteobacteria*, while the good prognosis differential taxa were positively correlated with each other within phylum *Firmicutes* or *Bacteroidetes*. The greater abundance of the poor prognosis associated taxa was correlated with a lower observed species, whereas, the greater abundance of the good prognosis associated taxa was correlated with a higher observed species (Fig. 3a). On the other hand, the greater abundance of *Bacillus* and/or *Lactobacillus* contributed 82.75% to a good richness of microbiota, whereas, the greater abundance of *Methyloversatilis* and/or *Schlegelella* contributed to 59.52% of poor richness of microbiota (Fig. 3b).

Considering the interaction of the differential taxa, we used the ratio of total abundances of *Shlegelella* and *Methyloversatilis* and total abundances of *Bacillus* and *Lactobacillus*, and defined as microbial dysbiosis index (MDI) representing the imbalance of the microbiota between distinct prognoses [33]. The MDI demonstrated significant differences between the patients with OS  $\leq$  3 yrs and the patients with OS  $>$  3 yrs as expected ( $P < 0.001$ , Fig. 3c). Notably, the MDI showed an inverse correlation with the observed species ( $r = -0.232$ ,  $P = 0.004$ ) and a direct correlation with the weighted UniFrac Distances ( $r = 0.157$ ,  $P = 0.001$ ) as shown in Fig. 3d-f.

Together, the results suggested that a greater abundance of *Bacillus* and *Lactobacillus* promoted a richer microbiota. On the other hand, a greater abundance of *Methyloversatilis* and *Schlegelella* impeded the richness of the microbiota which might result in an imbalance. These also indicated that the MDI, the ratio of above taxa, could not only represent the compositional imbalance, but also reflect the structural imbalance to a certain extent.

### 3. Microbial dysbiosis index is associated with HNC prognosis independently

Next, we seek to investigate the association of the microbial dysbiosis index and the prognosis of HNC. Firstly, we evaluated discriminating value of the MDI in prognosis indication of HNCs using time dependent receiver operating characteristics (ROC) analysis. The MDI exhibited an excellent performance in estimating overall survival with an AUC of 0.708 for three-year survival estimation (Fig. 4a), and was much improved compared to the use of observed species or single taxa (**Additional File 3: Figure S1**). In addition, a higher cutoff value applied for the MDI presented an even better discrimination for one-year overall survival (AUC of 0.769, Fig. 4a).

Secondly, we performed Kaplan Meier analyses to investigate the association of the MDI with survival of HNCs. Median overall survival of all the 158 patients in this cohort was 61 months. The median overall survival of the patients with the high MDI (46 months with 95% CI of 33 months to not reached) was higher compared with that of the patients with the low MDI (not reached, log-rank  $P < 0.001$ , Fig. 4b). Consistently, median disease free survival of the patients with the high MDI patients (43 months with 95% CI of 35 to 83 months) was higher compared with that of the patients with the low MDI (not reached, log-rank  $P < 0.001$ , Fig. 4c).

Thirdly, we performed Cox regression analyses of the microbial and the clinic-pathological characteristics before multivariate Cox regression. The MDI and nearly all differentially abundant taxa (except for *Bacillus*), but not the observed species, were associated with overall survival. Lymph node metastasis, TNM staging and p16 status were associated with overall survival, age and differentiation were also associated with disease free survival in univariate Cox regression (**Additional File 3: Figure S2**). In multivariate Cox regression, the MDI was a prognostic indicator of both overall survival and disease free survival characteristics independent of age, TNM staging, differentiation and p16 status (HR 3.24; 95% CI, 1.88 to 5.59;  $P < 0.001$  and HR 2.11; 95% CI, 1.34 to 3.34;  $P = 0.001$ , respectively; Fig. 4d). We further evaluated discriminating value of the risk score of above multivariate Cox model (a factor considering the effects of MDI, age, TNM staging, differentiation and p16 status) in prognosis indication of HNCs using ROC analysis. The risk score exhibited an improved performance in estimating overall survival (AUC of 0.831 for one-year survival and 0.793 for three-year survival; Fig. 4e).

In addition, we reviewed 12 articles describing differentially enriched taxa in oral microbiota and tissue microbiota between HNCs and controls (**Additional File2: Table S3**). Although the taxa differentially enriched in the cases and controls (or tumor tissue and adjacent/contradictory controls) were discriminating, we analyzed all above taxa succeeded to be annotated in our study and identified the genera *Bacteroides*, *Enterococcus* (enriched in oral cavity of the cases), *Rothia* and *Veillonella* (enriched in the tissue of the cases) to be associated with good overall survival (**Additional File 3: Figure S3**). These results indicated that the risking taxa of incidence might not associate with poor prognosis, or vice versa.

Collectively, these data proved that the microbial dysbiosis index exhibited an excellent performance for survival prediction compared to the use of single taxon including the differentially enriched taxa and the HNC incidence associated taxa. The MDI showed a strong and an independent association with the overall survival of HNCs comparable to the TNM staging. The risk score combining the effects of the MDI, age, TNM staging, differentiation and p16 status, is an excellent prognostic indicator of HNCs.

#### **4. Microbial dysbiosis index complement TNM staging or HPV status in prognostic prediction of HNCs**

TNM staging and HPV status have been generally accepted as prognostic indicators of HNCs, therefore, we performed following analyses to investigate whether MDI could complement TNM staging and HPV status. Firstly, we compared the MDI between early stage and late stage patients (scatters in Fig. 5a), as well as between HPV negative and HPV positive patients (scatters in Fig. 5b). Although no differences were observed, higher MDI distributions were shown in early stage patients and HPV positive patients. Cochrane-Mantel-Haenszel tests demonstrated that the high MDI was associated with poor overall survival (OS  $\leq 3$  yrs) in both early TNM patients and late TNM patients or both p16 negative patients and p16 positive patients (pies in Fig. 5a and 5b).

Secondly, we conducted stratified Kaplan-Meier analyses and subgroup Cox regressions. The late stage patients with the high MDI exhibited significant poor overall survival (median of 30.5 months, log-rank  $P < 0.001$ , Fig. 5c). Interestingly, the survivals of the high MDI patients and the low MDI patients were significantly different in the late stage patients ( $P < 0.001$ ), but the difference in the early stage patients

was not significant ( $P=0.066$ ). On the other hand, the survivals of the early stage patients and the late stage patients were significantly different in the patients with high MDI ( $P=0.028$ ), but the difference in the patients with low MDI was not significant ( $P=0.107$ ). In terms of p16 status stratification, the survivals of high MDI patients and low MDI patients were significantly different in both the p16 positive patients ( $P=0.010$ ) and the p16 negative patients ( $P<0.001$ ). However, no significant differences were observed between the p16 positive patients and the p16 negative patients in the patients with high MDI ( $P=0.100$ ) or in the patients with low MDI ( $P=0.053$ ). Consistently, in the subgroup Cox regression analysis, the MDI showed higher risk in the late stage patients (HR 3.35; 95% CI, 1.73 to 6.49;  $P<0.001$ ) compared with the early stage patients (HR 2.41; 95% CI, 1.00 to 5.84;  $P=0.051$ ) (Fig. 5e). And the MDI showed both high risks in the p16 positive patients (HR 4.77; 95% CI, 1.43 to 15.88;  $P=0.011$ ) and the p16 negative patients (HR 3.37; 95% CI, 1.88 to 6.03;  $P<0.001$ ). These results indicated that the MDI show more distinctive efforts in the late stage patients. Furthermore, these results also suggested that the MDI might showed more association with the overall survival compared to p16 status.

Thirdly, we evaluated discriminating value of the MDI in three-year overall survival prediction of HNCs using ROC analysis (Fig. 5f-5g). The MDI showed significant higher AUC compared to p16 status (AUC of 0.708 and 0.581 respectively,  $P=0.050$ ), but not compared to TNM staging (AUC of 0.708 and 0.644 respectively,  $P=0.362$ ). Consistently, p16 status together with the MDI exhibited an improved performance compared to p16 status only (AUC of 0.719 and 0.581 respectively,  $P=0.024$ ). However, the improvement of TNM staging together with the MDI was not significant compared to TNM staging only (AUC of 0.730 and 0.644 respectively,  $P=0.171$ ). TNM staging, p16 status together with the MDI exhibited an excellent performance with an AUC of 0.736 and a sensitivity of 0.971. These results suggested that the MDI could complement the prognostic indication of HNCs for p16 status to a considerable extent, but a limited extent for TNM staging.

## 5. Tumor behavior and inflammatory infiltration modulation might contribute to the association of microbiota and the prognosis of HNCs

To seek to understand how the microbiota influences the outcomes of HNCs, we analyze the correlation of the microbial signature and the clinic-pathological characteristics. We found that the greater abundances of *Gammaproteobacteria* and *Shlegelella* were correlated with elder age, and a greater abundance of *Proteobacteria* was correlated with lower differentiation, whereas, *Lactobacillales* was correlated with smaller tumor size (Fig. 6a). Furthermore, significant microbial differences were observed in the patients with differing age, lymph node metastasis or differentiation, but not in the patients with differing gender, smoking or drinking status, tumor site, histological type or tumor size (**Additional File 3: Figure S4-5**). Essentially, the differential taxa enriched in poor prognosis, elder age, positive lymph nodes or late stage belong to the phylum *Proteobacteria*, by contrast, the taxa differentially enriched in good prognosis, negative lymph nodes or early stage belong to the phylum *Firmicutes* (**Additional File 3: Figure S6**). These findings suggested the microbiota might influence the patient outcomes through affecting the tumor behaviors, such as lymph node metastasis, potentially.

To investigate whether the bacteria affect the outcomes of HNCs through tumor microenvironment modulation, we analyze the correlation of the microbial signature with the overall inflammatory infiltration indicated by Klintrup-Mäkinen (K-M) score [31, 32]. Greater abundances of *Methyloversatilis* and/or *Schlegelella*, and less abundances *Bacillus* and/or *Lactobacillus* were distributed in none inflammatory infiltrated patients (Fig. 6b). No differences were found between the K-M score and the abundances of above differentially abundant taxa ( $P=0.109$  for *Lactobacillus*, Fig. 6c-f). However, higher *Lactobacillus* was observed in highly inflammatory infiltration compared to none inflammatory infiltration ( $P=0.042$ , Fig. 6d). As illustrated in the Alluvial plots in Fig. 6g-6j, the inflammatory infiltration did not affect the association of the differentially abundant taxa with the overall survival, except for *Lactobacillus*. These results suggested that *Lactobacillus* promoted the inflammatory infiltration to a certain extent, which might contribute to its association with the outcomes of HNCs. However, the other differentially abundant taxa, especially the risk taxa, *Methyloversatilis* and *Schlegelella*, might not affect the outcomes of HNCs through the modulation of inflammatory infiltration.

## Discussion

In this study, by profiling 16S rRNA genes using nextgeneration sequencing, we explored the tumoral microbiota associated with poor prognosis and good prognosis in large-scale and in-depth analyses. For the first time, we revealed the structural differences of tumoral microbiota in poor prognosis patients and good prognosis patients, with a reduced richness in poor prognosis patients. This is consistent with the findings of the gut microbiota in the recipients of allogeneic hematopoietic stem cell transplantation[6]. In composition, we identified nine good prognosis associated taxa and seven poor prognosis associated taxa. *Gammaproteobacteria*, which is negatively associated with the prognosis of pancreatic ductal adenocarcinoma[7], was also found associated with the prognosis of HNCs in our study. However, *Fusobacterium*, reported to be associated with periodontitis and pancreatic cancer[10, 34], showed no differential enrichment in HNC patients with distinct overall survival or early relapse.

The poor prognosis associated taxa, *Schlegelella*, has recently been identified in dentin caries lesions due to the advances of the next generation sequencing [35, 36]. This suggested a potential connection of the microbial translocation post mucosal integrity loss and the HNCs related with residual root or crown resulting from caries. Intriguingly, *Methyloversatilis* broadly distributes in the natural environment with the capabilities of methylotroph of methanol or methylamine as well as utilization of other organic acids such as alcohols and aromatic compounds [37, 38]. We have known that the Warburg effect of cancer cells needs large amount of carbon sources. Therefore, further studies investigating the involvement of the metabolites from *Methyloversatilis* in the growth and the metabolism of cancer cells may explain its association with prognosis and would be a stepping-stone for the understanding of HNCs. In accordance with previous studies of vaginal microbiota, *Lactobacillus*, associated with the early regressions of cervical intraepithelial neoplasia[39], was found associated with good prognosis of HNCs in our study. Together with the essential role of *Lactobacillus* in the balance of microbiota, the findings suggested that *Lactobacillus* depletion and microbial dysbiosis by antibiotic abuse in clinic might associate with poor prognosis.

We integrated the above genera most relevant to the prognosis of HNCs to calculate the microbial dysbiosis index. It not only reflected the compositional imbalance of the microbiota between poor prognosis and good prognosis, but also showed high correlations with the diversities of the microbiota. The MDI exhibited an improved capacity to discriminate poor prognosis compared to the use of single differentially abundant taxa. This suggests that the microbiota as a whole rather than individual taxa, including differentially enriched taxa as well as the HNC incidence associated taxa, played more important role in the prognosis of HNCs. Notably, risk factors for cancer initiation may not associated with cancer development and outcome, such as *Helicobacter Pylori*, the risk factor for gastric cancers, has been proven not to affect the prognosis or a beneficial prognostic indicator in different studies[40–42]. Our results suggested the HNC incidence risk taxa might not associate with poor prognoses, or vice versa. Therefore, the need to evaluate microbiological prognostic indicators, other than incidence risk indicators, is much essential. In our study, the risk scores combining the MDI, age, TNM staging, differentiation and p16 status demonstrated an excellent performance with an AUC nearly 80%.

Classical TNM staging is still the most important strategy for prognostic evaluation of HNCs in clinic. HPV positive patients have also been generally accepted to have a good prognosis. In our study, we found the MDI showed a comparable hazard ratio with the TNM staging. The late stage patients with the high MDI exhibited extremely reduced overall survival, by contrast, the late stage patients with the low MDI exhibited comparable overall survivals with the early stage patients. These findings not only supported the complementary value of the MDI for prognostic indication, but also suggested that microbial modulation might improve the overall survival of the late stage HNC patients potentially. In terms of HPV status, our results demonstrated a much stronger association of the MDI with the prognosis of HNCs compared to p16 status. The MDI could significantly improve the performance of p16 status in three-year survival indication of HNCs. These implicate more attention to be given to HPV positive HNCs with a high MDI microbiota for therapy or follow up strategies. On the other hand, vaginal microbiota has been proven to influence HPV acquisition and persistence and the development of cervical cancer[43]. Together with the microbial discriminations between HPV positive and HPV negative patients of HNCs in our study and other reports[14], these suggested that the distinct prognoses between HPV positive and negative HNCs may possibly result from the contribution of microbiota and provide directions for further studies.

Next, we found correlations of the differentially abundant taxa with the clinic-pathological characteristics. On the other hand, we found overlaps of the differentially abundant taxa in distinct prognosis and the clinic-pathological characteristics. These findings suggested that the microbiota might affect the clinic-pathological characteristics, such as tumor behaviors, leading to distinction of prognoses. We also took the tumor microenvironment into consideration. However, no correlations of the microbial signatures with the overall inflammatory infiltration based on Klintrup-Mäkinen (K-M) score were found, except for *Lactobacillus*. These results suggested that the inflammatory infiltration modulated by *Lactobacillus* might be involved in the progression of HNCs to a certain extent.

## Conclusion

In this study, we have demonstrated that the structure and the composition of the tumor microbiota in the HNC patients with poor prognosis are significantly different from that of the patients with good prognosis. Microbial dysbiosis index, the ratio representing the compositional imbalance of the whole microbiota, has shown strong association with overall and disease free survival independently of age, TNM staging, differentiation and HPV status. It could significantly improve the predictive performance of HPV status, indicating the important role of tumoral microbiota in the distinct prognoses between HPV positive patients and HPV negative patients. Unexpectedly, we found that the MDI severely affected the overall survival of the late stage patients. This suggests that tumoral microbiota might not only be used as a prognostic indicator, but also its modulation might be valuable in survival improvement of the late stage patients.

## Declarations

### Ethics approval and consent to participate

The study was approved from Ethics Committee of Qilu Hospital of Shandong University.

### Consent for publication

Not applicable.

### Availability of data and materials

The sequencing data that support the results of this study have been deposited in “NCBI” with the BioProject accession code “PRJNA624194” (<https://www.ncbi.nlm.nih.gov/sra/PRJNA624194>).

### Competing Interests

The authors declare no competing interests.

### Funding

This research was supported by the grants from the National Natural Science Foundation of China (NSFC) awarded to Prof. Xun Qu (No.81772879) and Dr. Yu Dou (No.81902770).

### Author's contributions

XQ and XX conceived and designed the project; KTW, SHL, WYT, ZQD, AWC, DZ, BW, LS, TW performed the surgeries and collected the clinical characteristics data; YD, KTW, YZ and WYT completed the follow-ups; CM, ZDN, DA WJG, and JL performed the immunohistochemistry and histological analyses; YD, JTS, QJW, BFS, LZ, QQS, WJY and JNW analyzed the data; XWW, WH and FCW gave constructive guides for

the design and analyses; YD and MN wrote the paper. All authors read and approved the final version of the article.

## Acknowledgements

Not applicable.

## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.
2. Cohen EE, LaMonte SJ, Erb NL, Beckman KL, Sadeghi N, Hutcheson KA, Stubblefield MD, Abbott DM, Fisher PS, Stein KD, et al. American Cancer Society Head and Neck Cancer Survivorship Care Guideline. *CA Cancer J Clin*. 2016;66(3):203-39.
3. Hooper SJ, Wilson MJ, Crean SJ. Exploring the link between microorganisms and oral cancer: a systematic review of the literature. *Head Neck*. 2009;31(9):1228-39.
4. Dzutsev A, Badger JH, Perez-Chanona E, Roy S, Salcedo R, Smith CK, Trinchieri G. Microbes and Cancer. *Annu Rev Immunol*. 2017;35:199-228.
5. Pflug N, Kluth S, Vehreschild JJ, Bahlo J, Tacke D, Biehl L, Eichhorst B, Fischer K, Cramer P, Fink A-M, et al. Efficacy of antineoplastic treatment is associated with the use of antibiotics that modulate intestinal microbiota. *Oncoimmunology*. 2016;5(6):e1150399-e99.
6. Taur Y, Jenq RR, Perales M-A, Littmann ER, Morjaria S, Ling L, No D, Gobourne A, Viale A, Dahi PB, et al. The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. *Blood*. 2014;124(7):1174-82.
7. Geller LT, Barzily-Rokni M, Danino T, Jonas OH, Shental N, Nejman D, Gavert N, Zwang Y, Cooper ZA, Shee K, et al. Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science (New York, NY)*. 2017;357(6356):1156-60.
8. Iida N, Dzutsev A, Stewart CA, Smith L, Bouladoux N, Weingarten RA, Molina DA, Salcedo R, Back T, Cramer S, et al. Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science (New York, NY)*. 2013;342(6161):967-70.
9. Naito S, Koga H, Yamaguchi A, Fujimoto N, Hasui Y, Kuramoto H, Iguchi A, Kinukawa N, Kyushu University Urological Oncology G. Prevention of recurrence with epirubicin and lactobacillus casei after transurethral resection of bladder cancer. *The Journal of urology*. 2008;179(2):485-90.
10. Mitsuhashi K, Noshō K, Sukawa Y, Matsunaga Y, Ito M, Kurihara H, Kanno S, Igarashi H, Naito T, Adachi Y, et al. Association of *Fusobacterium* species in pancreatic cancer tissues with molecular features and prognosis. *Oncotarget*. 2015;6(9):7209-20.
11. Heshiki Y, Vazquez-Urbe R, Li J, Ni Y, Quainoo S, Imamovic L, Li J, Sørensen M, Chow BKC, Weiss GJ, et al. Predictable modulation of cancer treatment outcomes by the gut microbiota. *Microbiome*.

2020;8(1):28.

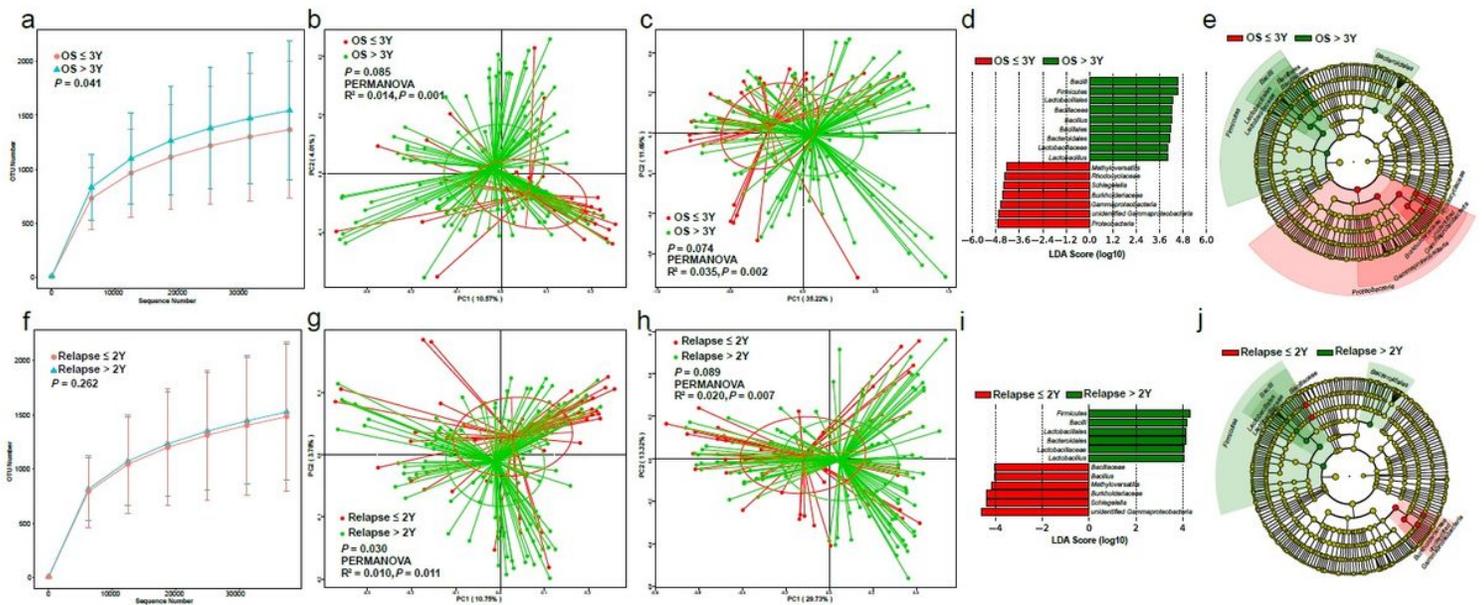
12. Yamamura K, Baba Y, Nakagawa S, Mima K, Miyake K, Nakamura K, Sawayama H, Kinoshita K, Ishimoto T, Iwatsuki M, et al. Human Microbiome *Fusobacterium Nucleatum* in Esophageal Cancer Tissue Is Associated with Prognosis. *Clin Cancer Res*. 2016;22(22):5574-81.
13. Hayes RB, Ahn J, Fan X, Peters BA, Ma Y, Yang L, Agalliu I, Burk RD, Ganly I, Purdue MP, et al. Association of Oral Microbiome With Risk for Incident Head and Neck Squamous Cell Cancer. *JAMA Oncol*. 2018;4(3):358-65.
14. Guerrero-Preston R, Godoy-Vitorino F, Jedlicka A, Rodriguez-Hilario A, Gonzalez H, Bondy J, Lawson F, Folawiyo O, Michailidi C, Dziedzic A, 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.
15. Guerrero-Preston R, White JR, Godoy-Vitorino F, Rodriguez-Hilario A, Navarro K, Gonzalez H, Michailidi C, Jedlicka A, Canapp S, Bondy J, et al. High-resolution microbiome profiling uncovers *Fusobacterium nucleatum*, *Lactobacillus gasseri/johnsonii*, and *Lactobacillus vaginalis* associated to oral and oropharyngeal cancer in saliva from HPV positive and HPV negative patients treated with surgery and chemo-radiation. *Oncotarget*. 2017;8(67):110931-48.
16. Yang CY, Yeh YM, Yu HY, Chin CY, Hsu CW, Liu H, Huang PJ, Hu SN, Liao CT, Chang KP, et al. Oral Microbiota Community Dynamics Associated With Oral Squamous Cell Carcinoma Staging. *Front Microbiol*. 2018;9:862.
17. 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.
18. Lee WH, Chen HM, Yang SF, Liang C, Peng CY, Lin FM, Tsai LL, Wu BC, Hsin CH, Chuang CY, et al. Bacterial alterations in salivary microbiota and their association in oral cancer. *Sci Rep*. 2017;7(1):16540.
19. Pushalkar S, Mane SP, Ji X, Li Y, Evans C, Crasta OR, Morse D, Meagher R, Singh A, Saxena D. Microbial diversity in saliva of oral squamous cell carcinoma. *FEMS Immunol Med Microbiol*. 2011;61(3):269-77.
20. Schmidt BL, Kuczynski J, Bhattacharya A, Huey B, Corby PM, Queiroz EL, Nightingale K, Kerr AR, DeLacure MD, Veeramachaneni R, et al. Changes in abundance of oral microbiota associated with oral cancer. *PLoS One*. 2014;9(6):e98741.
21. Al-Hebshi NN, Nasher AT, Idris AM, Chen T. Robust species taxonomy assignment algorithm for 16S rRNA NGS reads: application to oral carcinoma samples. *J Oral Microbiol*. 2015;7:28934.
22. Pushalkar S, Ji X, Li Y, Estilo C, Yegnanarayana R, Singh B, Li X, Saxena D. Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. *BMC Microbiol*. 2012;12:144.
23. Wang H, Funchain P, Bebek G, Altemus J, Zhang H, Niazi F, Peterson C, Lee WT, Burkey BB, Eng C. Microbiomic differences in tumor and paired-normal tissue in head and neck squamous cell

- carcinomas. *Genome Med.* 2017;9(1):14.
24. Shin JM, Luo T, Kamarajan P, Fenno JC, Rickard AH, Kapila YL. Microbial Communities Associated with Primary and Metastatic Head and Neck Squamous Cell Carcinoma - A High Fusobacterial and Low Streptococcal Signature. *Sci Rep.* 2017;7(1):9934.
  25. Banerjee S, Tian T, Wei Z, Peck KN, Shih N, Chalian AA, O'Malley BW, Weinstein GS, Feldman MD, Alwine J, et al. Microbial Signatures Associated with Oropharyngeal and Oral Squamous Cell Carcinomas. *Sci Rep.* 2017;7(1):4036.
  26. Chi AC, Day TA, Neville BW. Oral cavity and oropharyngeal squamous cell carcinoma—an update. *CA Cancer J Clin.* 2015;65(5):401-21.
  27. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, Westra WH, Chung CH, Jordan RC, Lu C, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med.* 2010;363(1):24-35.
  28. Maxwell JH, Grandis JR, Ferris RL. HPV-Associated Head and Neck Cancer: Unique Features of Epidemiology and Clinical Management. *Annu Rev Med.* 2016;67:91-101.
  29. Lin BM, Wang H, D'Souza G, Zhang Z, Fakhry C, Joseph AW, Drake VE, Sanguineti G, Westra WH, Pai SI. Long-term prognosis and risk factors among patients with HPV-associated oropharyngeal squamous cell carcinoma. *Cancer.* 2013;119(19):3462-71.
  30. Mitra A, MacIntyre DA, Marchesi JR, Lee YS, Bennett PR, Kyrgiou M. The vaginal microbiota, human papillomavirus infection and cervical intraepithelial neoplasia: what do we know and where are we going next? *Microbiome.* 2016;4(1):58-58.
  31. Klintrup K, Mäkinen JM, Kauppila S, Väre PO, Melkko J, Tuominen H, Tuppurainen K, Mäkelä J, Karttunen TJ, Mäkinen MJ. Inflammation and prognosis in colorectal cancer. *Eur J Cancer.* 2005;41(17):2645-54.
  32. Park JH, Richards CH, McMillan DC, Horgan PG, Roxburgh CSD. The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. *Annals of oncology : official journal of the European Society for Medical Oncology.* 2014;25(3):644-51.
  33. Ferreira RM, Pereira-Marques J, Pinto-Ribeiro I, Costa JL, Carneiro F, Machado JC, Figueiredo C. Gastric microbial community profiling reveals a dysbiotic cancer-associated microbiota. *Gut.* 2018;67(2):226-36.
  34. Galvao-Moreira LV, da Cruz MC. Oral microbiome, periodontitis and risk of head and neck cancer. *Oral Oncol.* 2016;53:17-9.
  35. Xu H, Hao W, Zhou Q, Wang W, Xia Z, Liu C, Chen X, Qin M, Chen F. Plaque bacterial microbiome diversity in children younger than 30 months with or without caries prior to eruption of second primary molars. *PLoS One.* 2014;9(2):e89269.
  36. Simón-Soro A, Mira A. Solving the etiology of dental caries. *Trends Microbiol.* 2015;23(2):76-82.
  37. Kittichotirat W, Good NM, Hall R, Bringel F, Lajus A, Médigue C, Smalley NE, Beck D, Bumgarner R, Vuilleumier S, et al. Genome sequence of *Methyloversatilis universalis* FAM5T, a methylotrophic

representative of the order Rhodocyclales. J Bacteriol. 2011;193(17):4541-2.

38. Smalley NE, Taipale S, De Marco P, Doronina NV, Kyrpides N, Shapiro N, Woyke T, Kalyuzhnaya MG. Functional and genomic diversity of methylotrophic Rhodocyclaceae: description of *Methyloversatilis discipulorum* sp. nov. Int J Syst Evol Microbiol. 2015;65(7):2227-33.
39. Mitra A, MacIntyre DA, Ntritsos G, Smith A, Tsilidis KK, Marchesi JR, Bennett PR, Moscicki AB, Kyrgiou M. The vaginal microbiota associates with the regression of untreated cervical intraepithelial neoplasia 2 lesions. Nat Commun. 2020;11(1):1999.
40. Santos RS, Lourenço JE, Herbella FA, Del Grande JC, Patti MG. *Helicobacter pylori* has no influence on distal gastric cancer survival. Arq Gastroenterol. 2011;48(2):109-11.
41. Varga MG, Wang T, Cai H, Xiang YB, Gao YT, Ji BT, Pawlita M, Waterboer T, Zheng W, Shu XO, et al. *Helicobacter pylori* Blood Biomarkers and Gastric Cancer Survival in China. Cancer Epidemiol Biomarkers Prev. 2018;27(3):342-44.
42. Wang F, Sun GP, Zou YF, Zhong F, Ma T, Li XQ, Wu D. *Helicobacter pylori* infection predicts favorable outcome in patients with gastric cancer. Curr Oncol. 2013;20(5):e388-95.
43. Mitra A, MacIntyre DA, Marchesi JR, Lee YS, Bennett PR, Kyrgiou M. The vaginal microbiota, human papillomavirus infection and cervical intraepithelial neoplasia: what do we know and where are we going next? Microbiome. 2016;4(1):58.

## Figures



**Figure 1**

Profile and composition of tumoral microbiota in patients with distinct prognoses a-e The profile and composition of tumoral microbiota in the patients with poor overall survival (OS ≤ 3Y) or good overall

survival (OS > 3Y). f-j The profile and composition of tumoral microbiota in the patients with early relapse ( $\leq 2Y$ ) or late relapse (> 2Y) or non-relapse. Rarefaction curves of observed OTUs (a, f) represent  $\alpha$ -diversity of tumoral microbiota. PCoA plots of unweighted (b, g) and weighted (c, h) UniFrac distances represent  $\beta$ -diversity of tumoral microbiota. Wilcoxon test is used to analyze the significance of differences, and PERMANOVA is used to discriminate the similarities. Differentially abundant taxa are identified by LefSe and shown in LDA scores (d and i) and cladogram (e and j).

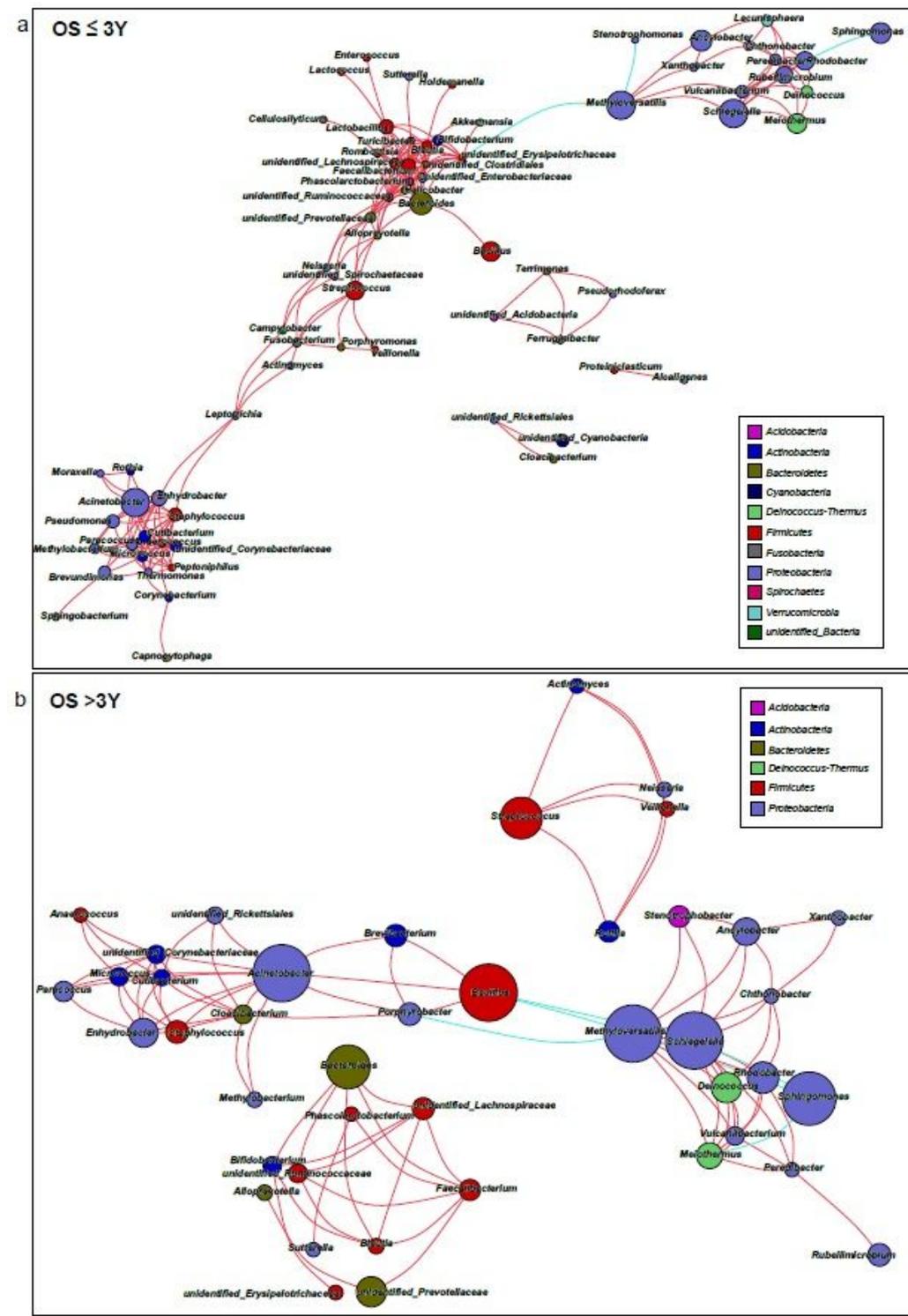
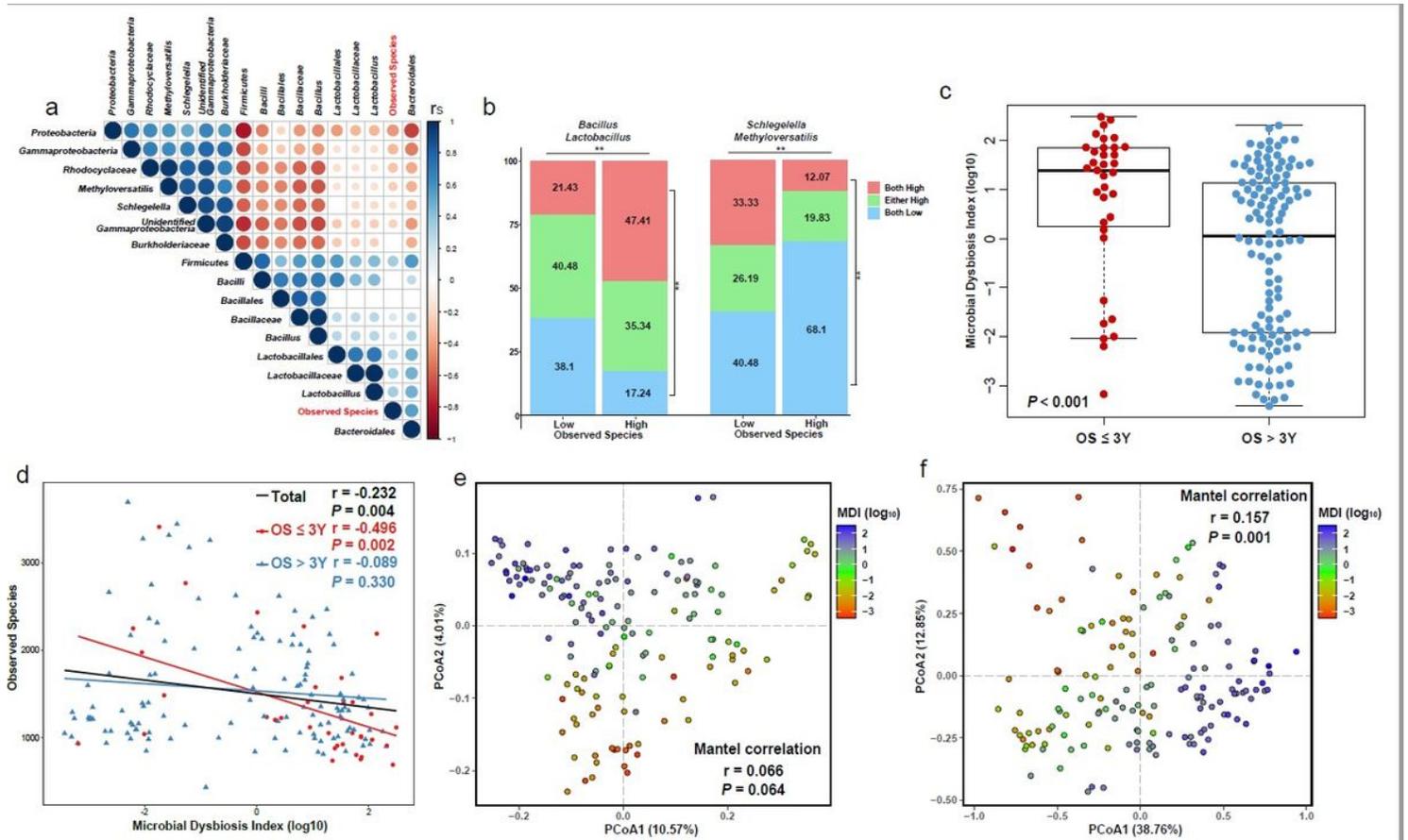


Figure 2

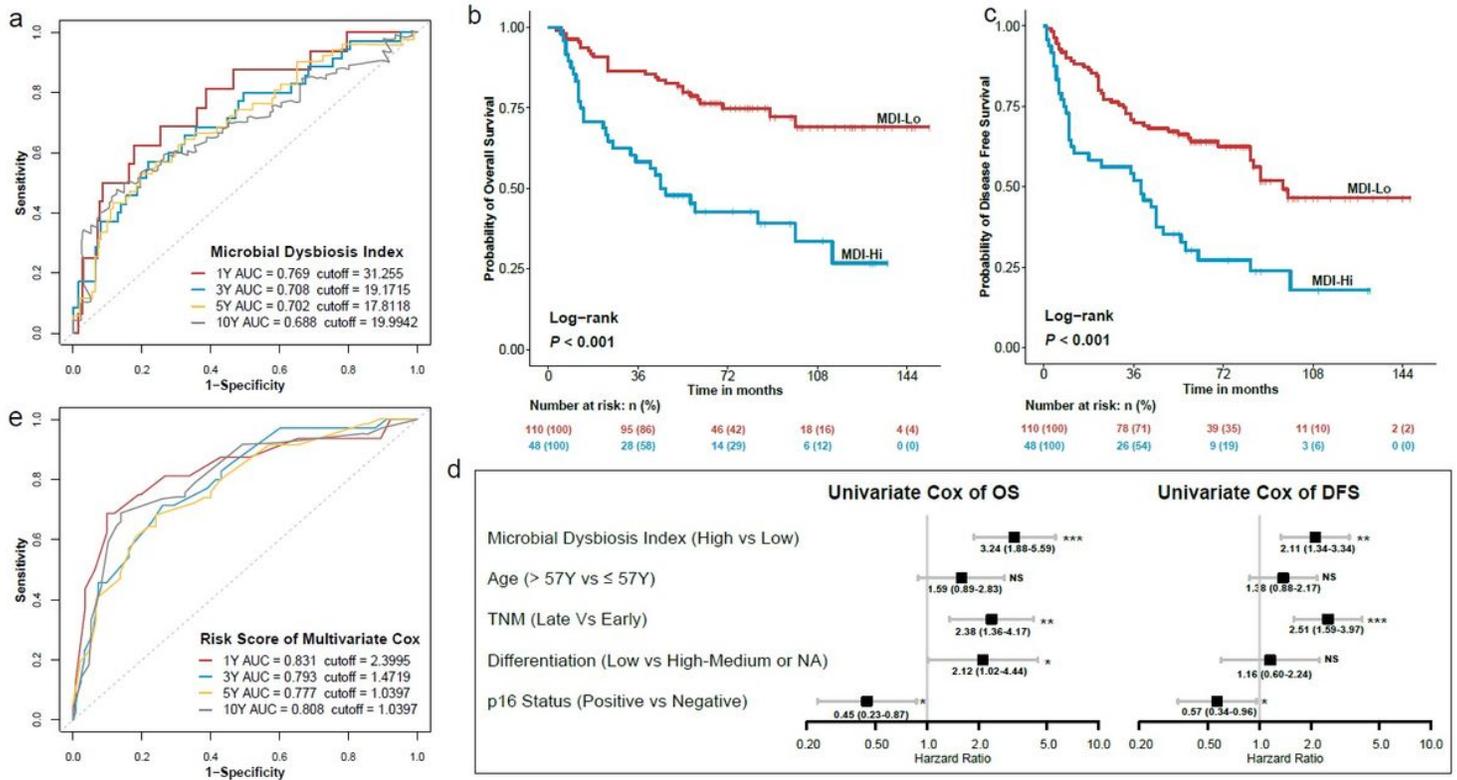
Correlation networks of the genera in patients with distinct overall survival a Network in the patients with  $OS \leq 3Y$ . b Network in the patients with  $OS > 3Y$ . Each node represents a genus, and the size is proportional to the mean relative abundance. The edges represent the correlation analyzed using Spearman's rank test. Red edges represent positive correlation, and blue edges represent negative correlation. Only the nodes and the edges with correlation coefficient  $> 0.6$ ,  $P < 0.05$  and relative abundance  $> 0.005\%$  are shown.



**Figure 3**

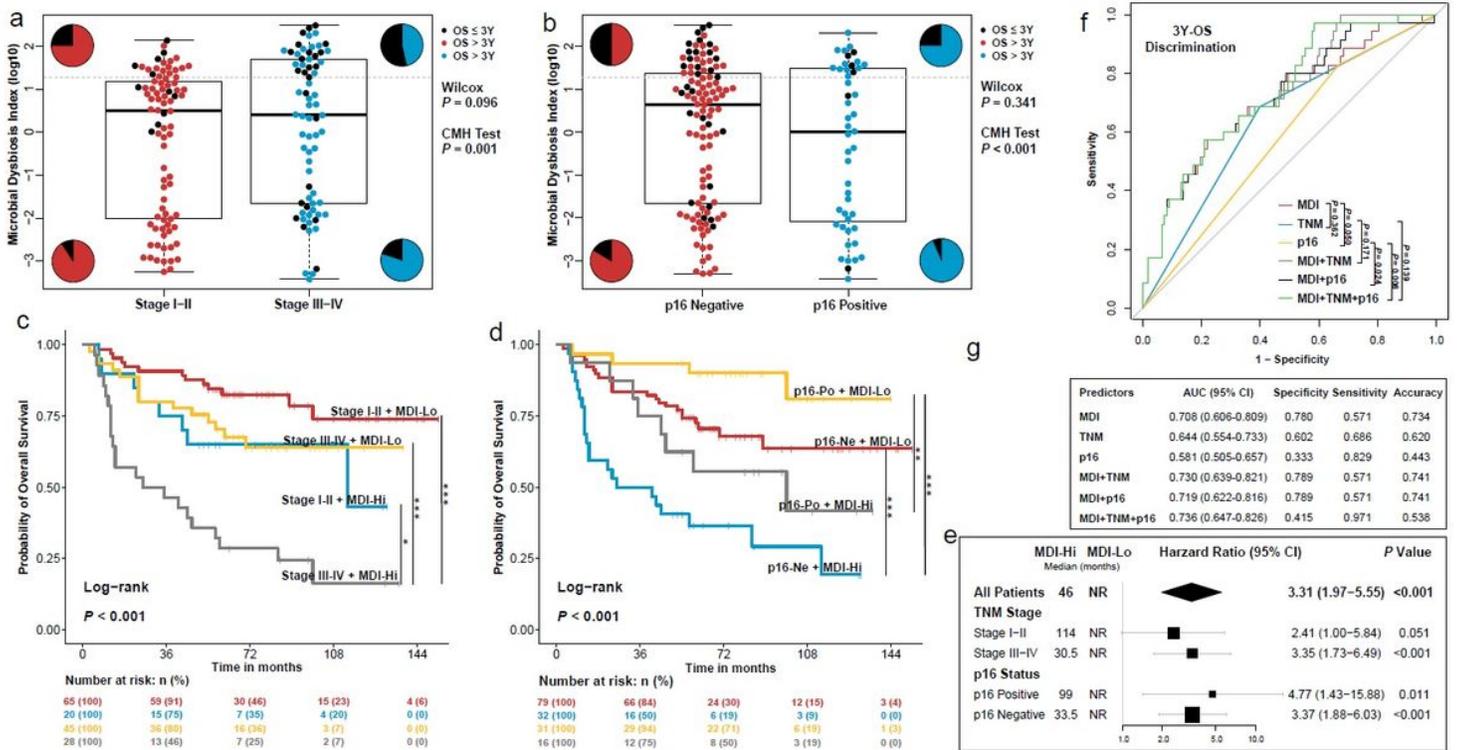
Correlation of differentially abundant taxa, microbial dysbiosis index and microbial diversities a A heatmap illustrates the correlation of observed species and the relative abundances of the differentially abundant taxa for  $P < 0.05$  (FDR adjusted) in Spearman's rank test. Hierarchical clustering of relative abundances is based on the Ward.D method. b Columns show the proportions of poor prognosis associated genera, Schlegellella and/or Methyloversatilis, and good prognosis associated genera, Bacillus and/or Lactobacillus, in the contribution of high or low observed species. Chi square test is used to analyze the significance of differences, \*\* indicates  $P < 0.01$ . c Beeswarm plot illustrates the microbial dysbiosis index (MDI) between the patients with poor overall survival ( $OS \leq 3Y$ ) or good overall survival ( $OS > 3Y$ ). Wilcoxon test is used to analyze the significance of differences. d Scatter plots illustrates Spearman's correlation between MDI and observed species. e-f Principal coordinate analysis (PCoA) plot of the unweighted (e) and weighted (f) UniFrac distance coloured by increasing MDI. Mantel correlations controlled with 104 permutations are used to compare distances. MDI is determined as the ratio between

the total abundance of genera increased in poor prognosis (*Schlegelella* and *Methyloversatilis*) and the total abundance of genera decreased in poor prognosis (*Bacillus* and *Lactobacillus*). Log transformed MDI values are used before plotted (c-f).



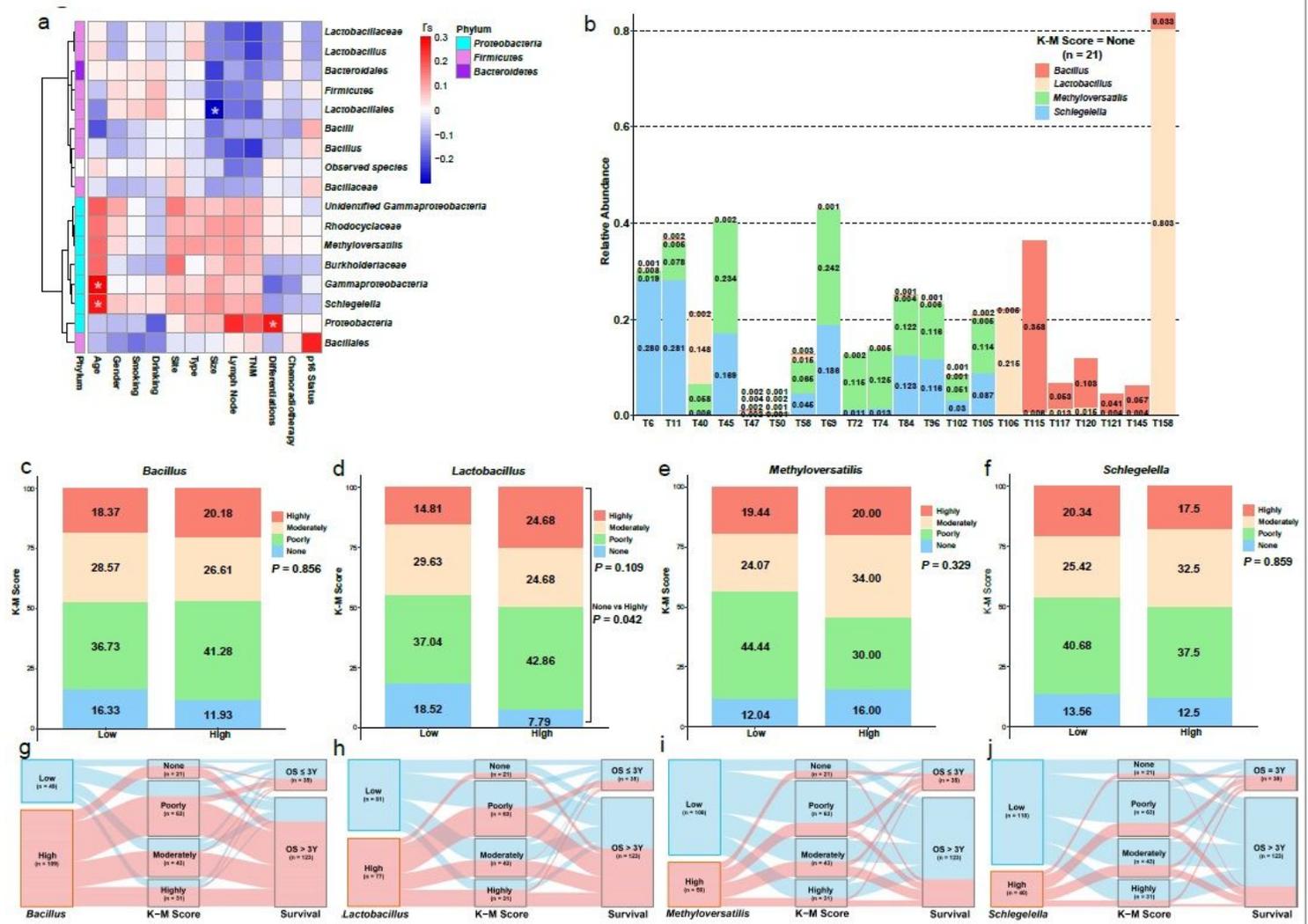
**Figure 4**

Association of microbial dysbiosis with overall or disease free survival of head and neck cancer a Time dependent ROC analysis evaluates the discriminatory potential of MDI in prognostic indication of head and neck cancer. The one-year, three-year, five-year and ten year areas under the curve (AUCs) and the cutoff values are indicated. b-c Kaplan-Meier curves illustrates overall survival (b) and disease free survival curve (c) probability of patients with high or low MDI. Statistical significance is assessed by log-rank test. d A forest plot illustrates the hazard ratio of MDI and the clinic-pathological characteristics including age, TNM stage, differentiation and p16 status in the multivariate Cox regression of overall survival and disease free survival. \* indicates  $P < 0.05$ , \*\* indicates  $P < 0.01$ , \*\*\* indicates  $P < 0.001$ . Univariate Cox regression of observed species, differentially abundant taxa and risk taxa ratio, as well as clinic-pathological characteristics are shown in Additional File 3: Figure S2. e Time dependent ROC analysis evaluates the discriminatory potential of risk score from multivariate Cox regression of overall survival in prognostic indication of head and neck cancer.



**Figure 5**

Complementary features of microbial dysbiosis for TNM stage and HPV status in prognostic indication of HNCs a-b Beeswarm plots show the log transformed MDI values between the early and late patients (a) or between p16 positive and negative patients (b). MDI of the patients with OS  $\leq$  3Y are indicated in black. Pie charts in the four quadrants indicated the proportions of above patients, such as the pie in left upper quadrant represents the proportion of the patient with OS  $\leq$  3Y or OS > 3Y in early stage or p16 negative patients with high MDI. The high or low MDI is discriminated by the cutoff value of ROC analysis indicated by the grey dashed lines. Wilcoxon test and Cochran-Mantel-Haenszel (CMH) test are used to analyze the significances of the differences for MDI values (scatters) and stratified proportions (pies) respectively. c-d Kaplan-Meier curves illustrates overall survival probability of the early or late stage patients (c) and p16 positive or negative patients (d) with high or low MDI. Statistical significance is assessed by log-rank test. P values of pairwise comparisons are FDR adjusted. \* indicates  $P < 0.05$ , \*\* indicates  $P < 0.01$ , \*\*\* indicates  $P < 0.001$  e A forest plot shows the hazard ratios of MDI in TNM stage and p16 status subgroups for overall survivals. f-g ROC analysis evaluates the discriminatory potential of the MDI, TNM stage, p16 status, MDI together with TNM stage, MDI together with p16 status, as well as MDI together with TNM stage and p16 status in three-year overall survival prediction of head and neck cancer. P16 negative is considered as “1” in this ROC analysis.



**Figure 6**

Association of differentially abundant taxa with clinic-pathological characteristics and inflammatory infiltration. a A heatmap of Spearman's correlation coefficient ( $r_s$ ) illustrates the correlation of observed species or the differentially abundant taxa with the clinic-pathological characteristics. Hierarchical clustering of relative abundances is performed using Euclidean distance based on the Ward.D method. \* indicates  $P < 0.05$  (FDR adjusted). b Columns show the distribution of *Bacillus*, *Lactobacillus*, *Methyloversatilis* and *Schlegelella* in the patients with none inflammatory infiltration (K-M score = 0). c-f Columns show the proportion of K-M score (none, poorly, moderately or highly inflamed) in the patients with low or high abundance of *Bacillus* (c), *Lactobacillus* (d), *Methyloversatilis* (e) and *Schlegelella* (f). P values of chi square test are indicated. g-j Alluvial plots for visualization of the association of *Bacillus* (g), *Lactobacillus* (h), *Methyloversatilis* (i) and *Schlegelella* (j) with inflammatory infiltration or three-year overall survival. The nodes represent the proportions of the according categories.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile1TableS1Baseline.docx](#)
- [AdditionalFile2TableS2TableS3.docx](#)
- [AdditionalFile2TableS3.pdf](#)
- [AdditionalFile3Supplementaryfiguresandlegends.pdf](#)