Identification of consensus head and neck cancer-associated microbiota signatures: a meta-analysis of 16S rRNA and The Cancer Microbiome Atlas datasets.

Kenny Yeo [1,2], Runhao Li[1,3], Fangmeinuo Wu[1,3], George Bouras[2], Linh T.H. Mai[1,2], Eric Smith[1,3], Peter-John Wormald[1,2], Rowan Valentine[2], Alkis James Psaltis[1,2], Sarah Vreugde[1,2], and Kevin Fenix[1,2]

(1) Discipline of Surgery, Adelaide Medical School, The University of Adelaide, Adelaide, SA, 5000, Australia

(2) Department of Surgery- Otolaryngology Head and Neck Surgery, The University of Adelaide and the Basil Hetzel Institute for Translational Health Research, Central Adelaide Local Health Network, Adelaide, SA, 5000, Australia

(3) Department of Haematology and Oncology, Basil Hetzel Institute for Translational Health Research and The Queen Elizabeth Hospital, Central Adelaide Local Health Network, Adelaide, SA, 5000, Australia

Corresponding author:

Kevin Fenix

Email: kevin.fenix@adelaide.edu.au

Word count: 6021

Highlights:

- The first meta-analysis of tissue microbiome in head and neck cancer containing eleven 16S ribosomal RNA and The Cancer Microbiome Atlas dataset.
- Microbiome from head and neck tissues were able to distinguish tissue types (cancer, cancer-adjacent, non-cancer) using 16S rRNA sequencing and whole genome sequencing datasets.
- Specific bacterial genera correlate with different tumour microenvironment phenotypes.
- High abundance *Fusobacterium* in tumour tissue correlates with better overall survival.

Abstract:

Objective: Multiple reports have attempted to describe the tumour microbiota in head and neck cancer. However, these have failed to produce a consistent microbiota signature which may undermine understanding the importance of bacterial-mediated effects in head and neck cancer. The aim of this study is to consolidate these datasets and identify a consensus microbiota signature in head and neck cancer.

Methods: We analysed 11 published head and neck cancer 16S ribosomal RNA microbial datasets collected from cancer, cancer-adjacent and non-cancer tissue to generate a consensus microbiota signature. These signatures were then validated using The Cancer Microbiome Atlas database.

Results: We identified unique bacteria enrichment within tissue types and correlated it with possible functional and clinical outcomes.

Conclusions: Our meta-analysis demonstrates a consensus microbiota signature for head and neck cancer, highlighting its potential importance in this disease.

Keywords: Tumour Microbiota, Head and Neck Cancer, 16s rRNA Sequencing, Meta-Analysis

52

1. Introduction

Recent studies have revealed that cancers previously thought to be sterile can contain unique microbial communities. The extent of microbial infiltration varies across different cancer types, with head and neck cancers (HNSC) containing one of the highest level of intratumoral microbial infiltrates while glioblastomas having the least amount of microbes.[1-3] This "intratumoral microbiota" can refer to bacterial infiltrates found in the extracellular matrix or within the cellular components of the tumour such as cancer, immune and stromal cells.[2] It is now widely appreciated that intratumoral bacteria can have direct and indirect effects on tumours or the tumour microenvironment (TME).[4-6] The presence of specific intratumoral bacteria has been reported to influence multiple features of tumour biology including treatment efficacy, local immune composition and activity and promoting tumour metastasis.[7-10]

Direct interaction between specific bacterial species with the tumour and the TME can induce chemoresistance, promote tumour progression, enhance therapeutic responses and modulate anti-tumour immunity through various mechanisms.[11-14] Bacteria can metabolise an active drug into its inactive form or induce autophagy in cancer cells which can promote chemoresistance.[12-14] Moreover, specific bacterial species can mount or suppress anti-tumour responses.[15-17] Most notably, *Fusobacterium nucleatum* colocalises with cancer and immune cells by binding to cell surface receptors such as Toll-like receptor 4 (TLR-4), T-cell immunoreceptor with Ig and ITIM domains (TIGIT) and Carcinoembryonic Antigen-Related Cell Adhesion Molecule 1 (CEACAM-1) receptors, or sugar groups (e.g. tumour expressed Galactose-N-acetylgalactosamine), which may then promote chemoresistance and suppress anti-tumour immunity[13, 15, 18-22]. Alternatively, *Bifidobacterium species* enhance anti-tumour immunity and efficacy of PD-1 immunotherapy responses.[8, 23, 24]

The release of bacterial metabolites such as short chain fatty acids (SCFA), amino acids, vitamins and bile acids can indirectly affect the tumour and the TME.[25, 26] Butyrate, a SCFA released by anaerobic bacteria through fermentation of carbohydrates, can decrease tumour cell growth and invasion, while increasing $CD8^+$ T cell-mediated anti-tumour responses.[27-30]. However, butyrate has also been shown to have pro-tumorigenic effects by inducing senescence-associated inflammatory phenotypes and inhibiting natural killer cell functions.[31, 32] Bacteria-derived indole and its derivatives (i.e. indole-3-lactic acid) have been shown to suppress anti-tumour immunity by activating immunosuppressive tumour-associated macrophages in treatment-naïve pancreatic cancer, while improving chemotherapeutic and immune-checkpoint inhibitor efficacy in pancreatic cancer and melanoma.[33-35] Together, these studies demonstrate that the tumour microbiota can influence cancer clinical outcomes in a context-dependent manner.

There are multiple reports describing the microbiota in HNSC.[36-89] Most of these studies compared the microbiota diversity and bacterial relative abundance between cancer and healthy samples using 16S ribosomal RNA (rRNA) sequencing[36-85, 88, 89], while two studies additionally correlated the impact of the microbiota with matched transcriptome analysis.[86, 90] Samples studied include tissues, swabs, and oral fluids (saliva or oral rinse) from cancer and healthy patients. Specifically for HNSC tissue microbiota analysis, samples included cancer, cancer-adjacent (approximately > 5 mm away from the tumour), contralateral, and healthy donor tissue samples.[36-55, 57-60, 85-89] Most bacteria identified in HNSC are oral commensal bacteria from the genera *Streptococcus, Rothia, Fusobacterium, Haemophilus* and *Prevotella.*[36-38] However, changes in microbial composition have been identified when cancer samples are compared to healthy controls. In general, there was an enrichment in

Fusobacterium within cancer tissue samples, that correlated with an inflammatory phenotype.[36, 37, 47] However, inconsistencies are observed for microbes such as *Streptococcus*, *Actinomyces* and *Prevotella* warranting the need to identify a consensus microbiota signature for HNSC.[37, 38, 43, 54, 85]

In this study, we systematically reviewed the literature and performed a meta-analysis to consolidate the currently heterogenous HNSC-associated microbiota data. Selected 16s rRNA sequencing datasets were analysed consistently to minimise variability between different sample cohorts and adjusted for batch-effects.[91] These consensus HNSC-associated microbial signatures were then validated using whole genome sequencing (WGS) data from The Cancer Microbiome Atlas (TCMA).[1] Finally, we correlated the presence of different microbiota signatures with the HNSC tumour microenvironment and clinical outcomes.

2. Methods

This study was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement.[92]

2.1 Search and Study Selection

The following criteria were used to select datasets: 1) Tissue samples, 2) Presence of metadata to distinguish sample types, 3) Illumina short-read amplicon sequencing of 16S rRNA V3 to V5 primers (Figure 1). Database search was performed on 16 August 2022 and datasets after this date were not included (Supplementary Table 1). The risk of biasness assessment was conducted using RoB 2 (β v9) (Supplementary Table 1).

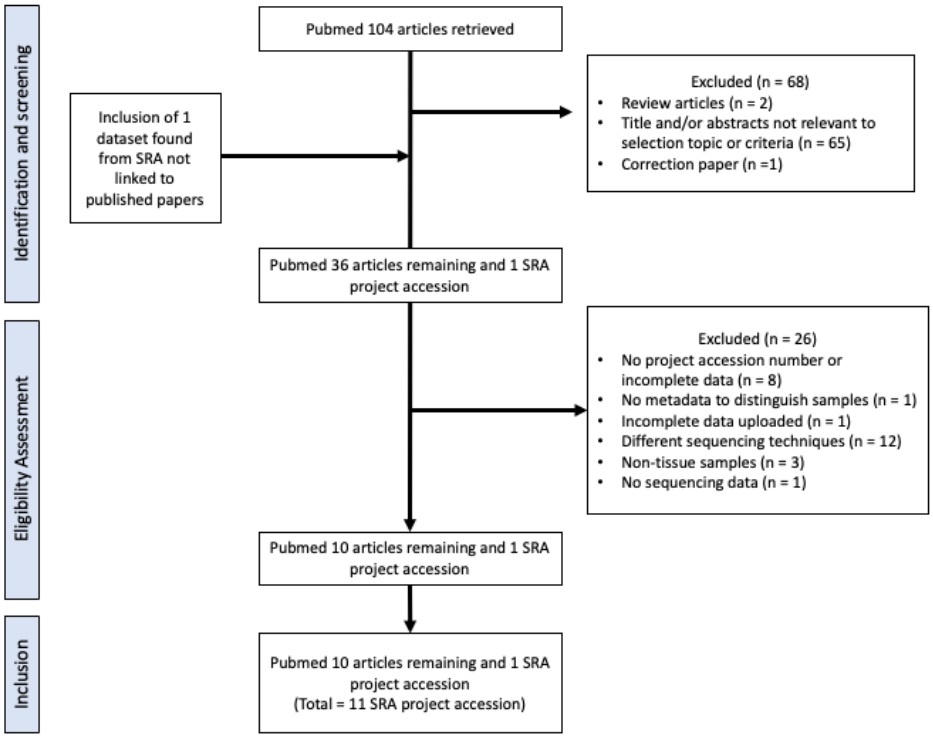


Figure 1: Study selection flow chart.

2.2 Download, pre-processing, and analysis of 16S rRNA datasets

Previously published raw sequences were retrieved from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) using pysradb.[93] Samples were divided into three main groups – cancer, cancer-adjacent and non-cancer tissues. Cancer tissues are defined as tissues obtained directly from the tumour, while cancer-adjacent tissues are cancer-free regions obtained > 5mm away from cancer tissues. Non-cancer tissues are defined as tissues that were either obtained from healthy patients or contralateral tissues obtained from cancer patients. FASTQ sequences files were obtained from SRA using sratoolkit.[94] These sequences were processed using QIIME2 DADA2 denoise-paired and reads truncated using the same parameters (trim_left_f = 30, trim_left_r = 30, trunc_q = 15). Sequences from different studies were merged before bacterial Operational Taxonomic Units (OTU) classification using QIIME2 and SILVA reference database (version silva-138-99-nb-classifier).[95]

Raw microbial reads were filtered, central log-ratio (CLR) transformed and batch-adjusted using Phyloseq and MixOmics as described previously.[96-98] Microbiome datasets are inherently compositional, hence, CLR transformation addresses generates scale-invariant values which allows datasets to remain unaffected by variations in library sizes among samples.[99] Briefly, low abundance of OTUs were filtered through proportional counts of all samples (< 1%) and minimum counts per sample (< 10). Bacterial OTUs were agglomerated at the genus level before transforming into CLR for their compositional nature.[96, 98] The CLR-abundance was used for subsequent statistical and discriminant analysis. A total of 903 SRA samples from 11 projects were downloaded (Table 1).

Table 1: Study accession and sample size post-filtering

Accession number	Sample size			Primers
	Cancer	Cancer-adjacent	Non-cancer	
PRJNA412445	16	0	0	V4 -V5
PRJNA555458	0	0	4	V3 -V4
PRJNA596113	102	53	0	V3 -V4
PRJNA597251	19	20	0	V3 -V4
PRJNA666746	50	50	0	V3 -V4
PRJNA666891	7	0	10	V4
PRJNA685226	13	13	0	V3 -V4
PRJNA699728	37	0	201	V4
PRJNA803155	40	0	0	V4 -V5
PRJNA822685	75	79	0	V3 -V4
PRJNA866676	37	36	41	V3 -V4

2.3 Discriminant analysis of 16S rRNA dataset

To discriminate the microbial signature between sample types, we employed both multivariate and univariate discriminant analysis. For β-diversity analysis, CLR-abundance of all genera were ordinated using Euclidean distance and plotted on a principal component analysis (PCA) using mixOmics R package. β-diversity for each sample were calculated as distance to centroid for each tissue groups using betadisper (vegan v2.6-4). Group and pairwise permutest (vegan v2.6-4, permutations = 9999) was performed to determine if dispersions differed between sample types, while group and pairwise permutational multivariate analysis of variance (PERMANOVA) was performed using adonis2 (vegan v2.6-4, method = “euclidean”, permutation = 9999) and pairwise.adonis2 (pairwiseAdonis, method = “euclidean”, permutation = 9999) to determine statistical differences in β-diversity between groups. Other statistical test such as Analysis of similarities (ANOSIM) (vegan v2.6-4, distance = “euclidean”, permutation = 9999) and Fifty-fifty multivariate analysis of variance (FFMANOVA) (nSim = 9999) were also applied as supplementary to distinguish between sample types.[100, 101]

Multivariate sparse partial linear discriminant analysis (sPLS-DA) was applied on batch-adjusted dataset to identify discriminating genera within each sample type.[96] The Area Under Curve (AUC) of the Receiver Operating Characteristics (ROC) curve was calculated using mixOmics in Rstudio.[96, 98] The AUC value served as a quantification of the discriminatory potential between sample types. A higher AUC value, closer to 1, signified a test approaching perfection in its ability to distinguish between the samples. Heatmap of all representative bacteria in each sPLS-DA was presented with sample type clustered according to Euclidean distance and Ward's linkage.

Univariate Kruskal-Wallis test with Bonferroni multiple comparisons test was also performed to determine microbial genera differences between sample types using microbiomeMarker in Rstudio v3.3.0, followed by a post-hoc Wilcoxon test (Mann-Whitney test) with Bonferroni-Dunn multiple comparison test to determine differences between groups (cancer– cancer-adjacent, cancer – non-cancer, non-cancer – cancer-adjacent). Additionally, Wilcoxon matched-pairs signed rank test with Bonferroni-Dunn multiple comparison test was also performed on paired cancer and cancer-adjacent samples.

2.4 Functional profiling analysis of 16S rRNA datasets in different sample types

To predict the microbial functions of genera detected from 16S rRNA sequencing between each tissue sample type, Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUST2) from QIIME2 was applied on raw 16S rRNA reads using MetaCyc database.[102, 103] Functional abundance was processed and analysed similarly as described for raw microbial reads. Univariate Kruskal-Wallis test and post-hoc Wilcoxon test was performed as previously described to compare differences between groups.

2.5 Reanalysis of tissue microbiome data from TCMA

Decontaminated microbial read count derived from The Cancer Genome Atlas (TCGA) HNSC whole genome sequences were obtained from TCMA repository.[1] Data from a total of 177 cancer (TCGA annotation: primary tumour) and 22 cancer-adjacent (TCGA annotation: solid tumour normal) tissues were obtained from TCMA repository (n = 22 paired cancer and cancer-adjacent samples). Similar to 16S rRNA pre-processing, read counts were agglomerated to the genus before CLR transformation as described in 2.2. As samples were already pre-processed in the TCMA dataset, no further filtering or batch adjustment was required. Microbiome statistical analysis were performed similarly as 16S sequencing datasets. Metadata were obtained from cBioPortal for Cancer Genomics.[104]

2.6 Microbiome correlation analysis with tumour microenvironment and survival analysis

The TME immune subtype and 29 functional gene expression signatures (FGES) scores were previously described by Bagaev et al. (2021) using transcriptomics datasets from TCGA.[105] The 29 FGES represents the major functional components and immune, stromal, and other cellular populations of the tumour.[105] Pearson's correlation test was applied to determine the correlation between FGES scores and selected bacteria genera. The four TME immune subtypes were – Desert (D), Fibrotic (F), Immune-enriched (IE), Immune-enriched/Fibrotic (IE/F) (Described in Supplementary Table 2).[105] Specifically, tissues with IE and IE/F phenotype contains high T-cell infiltration, while D and F phenotypes have low T-cell infiltration (Supplementary Table 2).[105] Using a cut-off of high (top 35th percentile) and low (bottom 35th percentile) CLR-abundance, the proportion of each patient within the four TME subtypes were determined, and survival analysis was performed. Since there were 153 TCGA-HNSC samples with both FGES/TME subtypes and microbiome datasets, these

samples were used for subsequent correlation and survival analysis. Chi-squared (χ^2) test was performed in Prism9 to determine association between high/low bacterial genera CLR-abundance and proportion of patients within each tumour subtype.

2.7 Statistical analysis

For comparisons made between all unpaired tissue groups, Kruskal-Wallis test with Bonferroni's multiple comparison was used for comparisons made between all tissue groups unless stated otherwise. Post hoc Wilcoxon matched pairs signed rank test with Bonferroni's multiple comparison was used to compare differences between unpaired tissue samples. For all paired cancer and cancer-adjacent samples, Wilcoxon matched-pairs signed rank test was performed. Univariate and multivariate Cox proportional hazard model was performed using survminer in Rstudio v3.3.0. Statistical analysis was performed using RStudio v3.3.0 and Prism9.

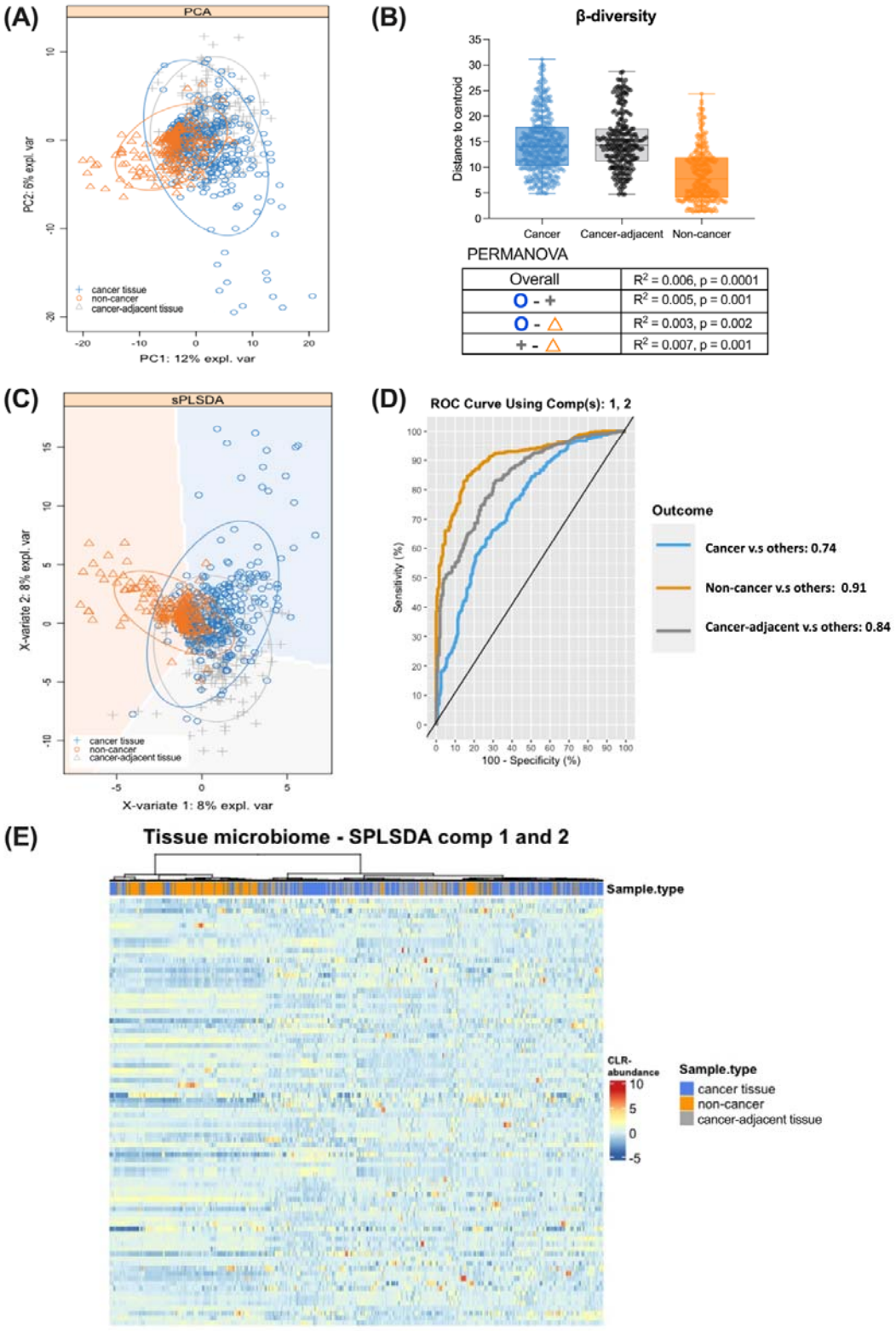
3 Results

3.1 Multivariate analysis identifies homogenous microbial abundance and functions between cancer and cancer-adjacent samples, contrasting to non-cancer samples.

The 16S rRNA amplicon datasets were obtained for 903 head and neck tissue types (396 cancer, 251 cancer-adjacent, and 256 non-cancer) from 11 studies.[37, 38, 41-45, 87-89] Following sample processing and aggregation of 16S data at the genus level, a total of 177 distinct bacterial genera were identified. Differences in the microbiota and β-diversity between tissue types were assessed using PCA and PERMANOVA test (Figure 2A-2B). The β-diversity index was calculated for cancer (14.6 ± 5.7) and cancer-adjacent (15.0 ± 5.5) tissues, revealing similar levels of β-diversity. In contrast, non-cancer tissues (8.61 ± 5.3) exhibited lower β-diversity (PERMANOVA – Overall R2 = 0.006, $p < 0.0001$) (Figure 2B). Post-hoc pairwise test identified significant differences in β-diversity between cancer and non-cancer ($R^2 = 0.003$, $p = 0.002$), cancer and cancer-adjacent ($R^2 = 0.005$, $p < 0.001$), and non-cancer and cancer-adjacent samples ($R^2 = 0.007$, $p < 0.001$) (Figure 2B). These findings were consistent with additional multivariate and univariate statistical analysis, ANOSIM ($R = 0.027$, $p = 0.002$) and FFMANOVA ($p < 0.0001$) (Supplementary Table 3).

Multivariate sparse partial least squares discriminant analysis (sPLS-DA) identified 116 representative bacterial genera in sPLS-DA component 1 and 2 which were discriminant between tissue types (Figure 2C-E). The AUC values were computed for different sample comparisons: cancer versus others (AUC = 0.74, $p < 0.05$), non-cancer versus others (AUC = 0.91, $p < 0.05$), and cancer-adjacent versus others (AUC = 0.84, $p < 0.05$). These results demonstrate that sPLS-DA components 1 and 2 (Figure 2D) can effectively differentiate between tissue types. Lastly, majority of cancer and cancer-adjacent samples clustered

273 together and were distinct from non-cancer samples, as determined by Euclidean distance
274 metric (Figure 2E).



275

Figure 2: Multivariate discriminant analysis (sPLS-DA and PERMANOVA) of tissue 16S rRNA microbiota to discriminant between cancer, cancer-adjacent and non-cancer tissues. (A) Principal coordinates analysis (PCA) plot of tissue CLR-abundance microbiota based on Euclidean distance. (B) Dispersion of β-diversity (top-right panel) for each sample type, with error bar representing 95% confidence interval. PERMANOVA test was performed with bacterial genera as variable for sample types. (C) sPLS-DA sample plot of 16S rRNA tissue microbiota. Ellipse displays 95% confidence interval for each sample group. The batch-adjusted normalized abundance of tissue microbiota from 16S amplicon sequencing was compared between cancer, cancer-adjacent and non-cancer tissue samples. sPLS-DA identified 116 bacterial genera on component 1 and 2. (D) ROC curve and AUC values determined from sPLS-DA analysis was used to access discriminatory potential of sPLS-DA component 1 and 2. (E) Heatmap representing 86 bacterial genera after sPLS-DA discriminant analysis. Each column and row represent a unique sample and bacterial genera respectively, with OTUs clustered based on Euclidean distance and Ward linkage method.

3.2 Univariate analysis identifies differences in microbial abundance and functions between sample types.

Next, unpaired univariate analysis was applied to determine the differences between tissue types. Out of the 177 bacterial genera, 33 were identified as significantly different among tissue types using Kruskal-Wallis test ($P_{adjust} < 0.05$) (Supplementary Table 4). Notably, 18 of these were also identified as representative bacterial genera in sPLS-DA discriminant analysis (Supplementary Table 4). These 33 genera are denoted as bacterial genera of interest (Supplementary Table 4). The top 20 differentially abundant genera, based on the effect size (η^2), are presented in Figure 3. Post-hoc unpaired Wilcoxon test with Bonferroni-Dunn's multiple comparison test was performed on these genera to determine the mean differences in the central log ratio transform (CLR) abundance between tissue types (Figure 3A,

Supplementary Table 5). Since most published studies compared cancer to non-cancer, or cancer to cancer-adjacent tissues, we performed post-hoc test for these comparisons (Figure 3B). We identified 27 out of 33 genera as significantly different (P_{adjust} (#) < 0.05) between cancer and non-cancer tissues (Figure 3A-3B, Supplementary Table 5). Non-cancer tissues contained more *Fretibacterium* (CLR-abundance diff. = 1.42, SE = 0.12), *Stenotrophomonas* (CLR-abundance diff. = 0.80, SE = 0.12) and *Tannerella* (CLR-abundance diff. = 0.71, SE = 0.10), while cancer tissue had a greater CLR-abundance of *Neisseria* (CLR-abundance diff. = 2.32, SE = 0.15), *Capnocytophaga* (CLR-abundance diff. = 2.02, SE = 0.15), and *Streptococcus* (CLR-abundance diff. = 1.98, SE = 0.19) (Figure 3A-3B). *Capnocytophaga* abundance in cancer tissues was consistent to previous findings[46, 57, 85], while contradicting findings were identified for the abundance for *Streptococcus*[38, 41, 52, 57, 85] and *Fusobacterium*[38, 41, 42, 52, 57, 58].

For cancer and cancer-adjacent tissue, 13 out of 33 bacterial genera were significantly different (post-hoc unpaired Wilcoxon test P_{adjust} (*) < 0.05) (Figure 3A and 3C, Supplementary Table 5). Similar to many studies, *Fusobacterium* (CLR-abundance diff. = 1.11, SE = 0.20) displayed significantly higher CLR-abundance in cancer tissue than cancer-adjacent tissue, while *Rothia* (CLR-abundance diff. = 0.92, SE = 0.18), *Stenotrophomonas* (CLR-abundance diff. = 1.33, SE = 0.15) and *Serratia* (CLR-abundance diff. = 0.70, SE = 0.12) had higher CLR-abundances in cancer-adjacent tissue than cancer tissue (Figure 3A).[36, 37, 43, 45, 52, 55, 58, 59]. Additionally, we found that *Prevotella* was elevated in cancer tissue as compared to cancer-adjacent tissues.[43, 45, 52, 55, 58] Unlike previous studies, we did not observe any significant differences in *Streptococcus* abundance between cancer and cancer-adjacent tissues.[36, 37, 45, 51, 52, 55, 59]

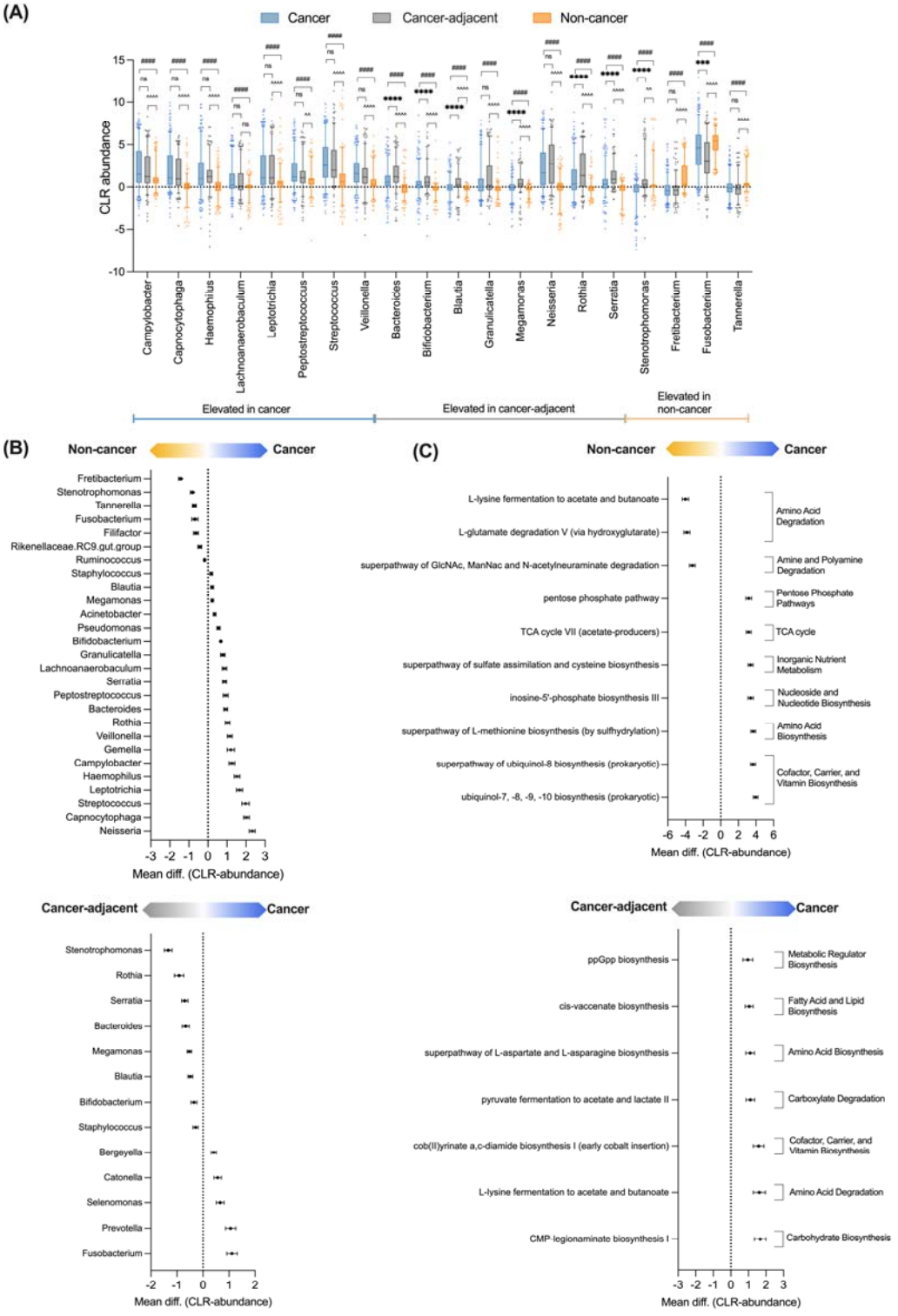
Lastly, 28 of the 33 top bacterial genera were significantly different (post-hoc unpaired Wilcoxon test $P_{adjust} < 0.05$) when comparing non-cancer to cancer-adjacent tissue samples (Figure 3A, Supplementary Table 5). Genera *Neisseria* (CLR-abundance diff. = 2.83, SE = 0.19), *Rothia* (CLR-abundance diff. = 1.95, SE = 0.16) and *Streptococcus* (CLR-abundance diff. = 1.95, SE = 0.16) were higher in CLR-abundance in cancer-adjacent, while *Fusobacterium* (CLR-abundance diff. = 1.80, SE = 0.18) and *Prevotella* (CLR-abundance diff. = 1.28, SE = 0.20 were greater in CLR-abundance in non-cancer tissue (Figure 3A).

To provide functional insights to microbial abundance between cancer tissues and other tissue types, we applied Picrust2 to predict possible differences in MetaCyc pathway functional CLR-abundance.[102] After filtering low abundant functional pathways, we identified a total of 365 MetaCyc pathways. Using Kruskal-Wallis test, 162 MetaCyc pathways were identified as significantly different among sample types ($P_{adjust} < 0.05$) (Supplementary Table 5). Post-hoc analysis identified 129/162 and 7/162 pathways that were significantly different between cancer – non-cancer, and cancer – cancer-adjacent tissues comparisons respectively (Supplementary Table 6).

Cancer tissues, when compared to non-cancer tissues, were enriched in pathways involving the synthesis of ubiquinol, L-methionine, inosine-5'-phosphate and cysteine and metabolic pathways such as TCA cycle and pentose phosphate pathway, while non-cancer tissues were enriched in the degradation of L-lysine, L-glutamine, N-Acetylglucosamine (GlcNac), N-acetylmannosamine (ManNac), and N-acetylneuraminate (Figure 3C). Cancer tissues were more similar to cancer-adjacent tissues, albeit enrichment was identified in pathways involving biosynthesis of ppGpp (guanosine pentaphosphate and tetraphosphate), cis-vaccenate, L-asparatate, L-asparagine, cob(II)yrinate a,c-diamide and CMP-legionaminate,

352 and enrichment in pathways involving degradation of pyruvate and L-lysine, when compared

353 to cancer-adjacent tissues (Figure 3C).



354

Figure 3. Comparison of bacterial CLR-abundance and functional prediction between sample types. (A) Top 20 bacterial genera (based on effect size) in CLR-normalized abundances between sample groups using Kruskal-Wallis test with Bonferroni's multiple comparison. 33 out of 177 genera were identified as significantly different ($P_{adjust} < 0.05$) using Kruskal-Wallis test. Post-hoc Wilcoxon test with Bonferroni-Dunn's multiple comparison was performed to identify group-wise differences between Cancer – Non-cancer (#), Cancer – Cancer-adjacent (*), Non-cancer – Cancer-adjacent (^). Post-hoc unpaired Wilcoxon test with Bonferroni-Dunn's multiple comparison for (B) bacterial genera and (C) functional CLR-abundance for Cancer – Non-cancer (Top panel), and Cancer – Cancer-adjacent (Bottom panel).

3.3 Paired cancer and cancer-adjacent tissues display similar bacterial abundance differences using multiple sequencing techniques.

To understand microbial abundance differences between cancer tissue and cancer-adjacent tissue within the same patients, we performed Wilcoxon matched-pairs signed rank test to identify changes in microbial diversity and abundance within paired tissue samples in the 16S rRNA datasets. Similar to unpaired data analysis, no significant differences in microbial β-diversity was identified between the patient's paired cancer and cancer-adjacent tissues (Supplementary Figure 2).

However, 76 bacterial genera were significantly different between paired tissue samples (Figure 4A, Supplementary Table 7). Bacterial genera with the greatest differences in CLR-abundance were then identified by using a cut-off of > 0.4 and < -0.4 (Figure 3A). Using this cut-off, we found that *Fusobacterium, Prevotella, Alloprevotella, Catonella, Selenomonas and Treponema* were elevated in cancer tissue vs cancer-adjacent tissue, while *Stenotrophomonas, Rothia, Granulicatella, Serratia, Anoxybacillus, Actinomyces and*

Bacteroides were greater in cancer-adjacent tissue compared to cancer tissue (Figure 4A-4B). Similarly, nine of these bacteria were also found to be significantly different in unpaired tissue analysis (Supplementary Table 4 and 7). Contrary to published studies on unpaired samples, *Streptococcus*, an abundant oral commensal, was not significantly different in our paired sample analysis.[36, 37, 45, 51, 52, 55, 59]

To validate this finding, we probed the publicly available TCMA dataset, a repository containing microbiota reads derived from WGS of tissue samples.[1] Similar to the 16S rRNA dataset, we observed that cancer tissues from TCMA displayed significantly ($p < 0.05$) higher CLR-abundance for genera *Fusobacterium, Selenomonas* and *Treponema*, while *Rothia* and *Actinomyces* were elevated ($p < 0.05$) in cancer-adjacent tissues (Figure 4D). In the TCMA dataset, *Anoxybacillus, Serratia,* and *Stenotrophomonas* were not present due to pre-analysis filtering, while no significant differences in CLR-abundance were observed for *Prevotella*, *Catonella*, *Alloprevotella*, and *Bacteroides* (Figure 4C). Notably, similar trend in CLR-abundance between cancer and cancer-adjacent samples was still observed for *Prevotella*, *Catonella*, and *Alloprevotella* in TCMA dataset. Overall, 16S rRNA and TCMA WGS dataset showed similar trend for most bacteria genera, regardless of sequencing techniques.

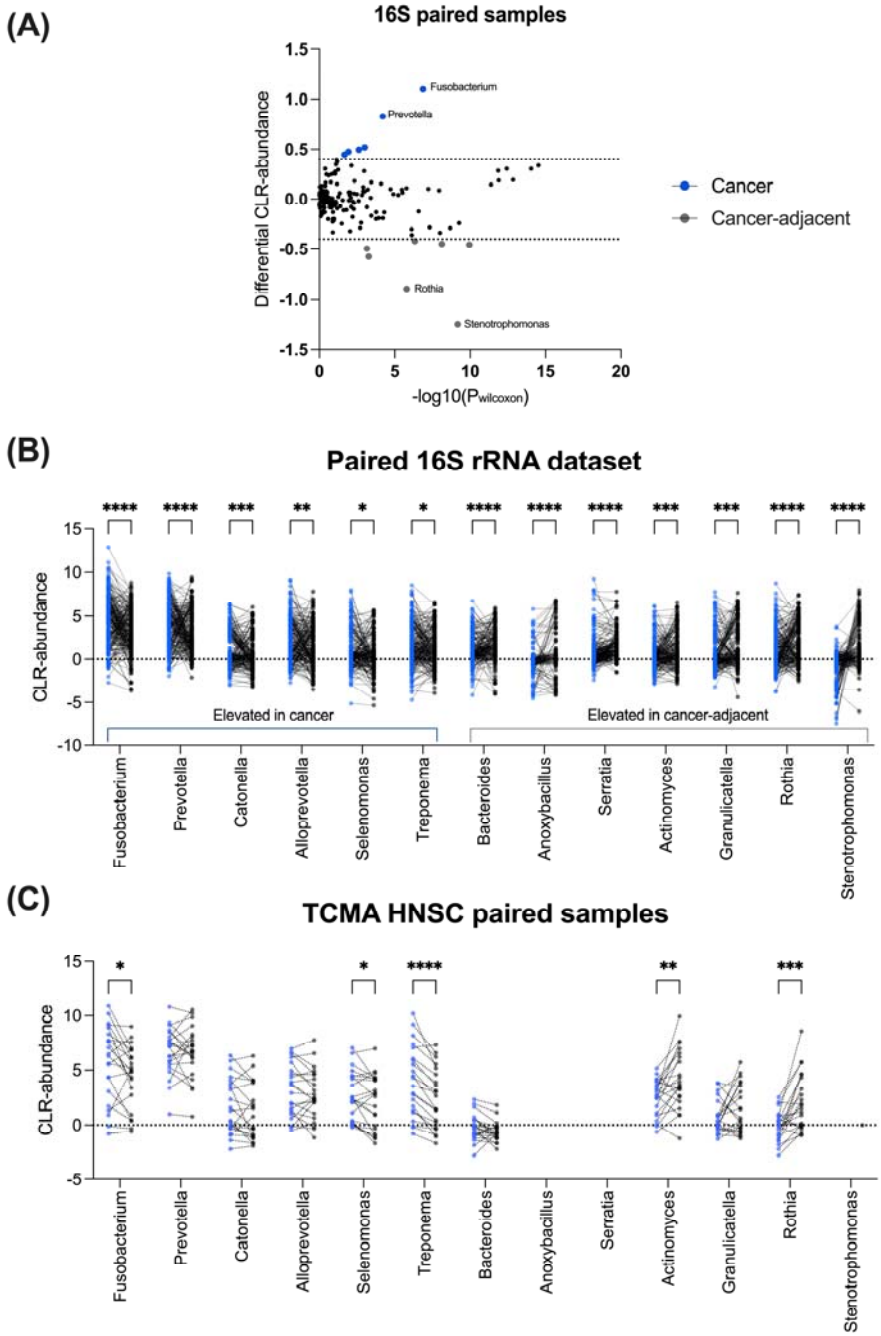


Figure 4: Comparison of tissue microbiota in paired cancer and cancer-adjacent tissue samples using different sequencing datasets. (A) Paired Wilcoxon matched-pairs signed rank test on paired 16S rRNA sequencing cancer and cancer-adjacent tissue samples. 76 bacteria were significantly different in sample groups ($p < 0.05$) using paired Wilcoxon matched-pairs signed rank test and 13 bacteria genera were identified as top bacteria with

differential CLR-abundance (Diff. CLR-abundance > 0.4 or < -0.4). Blue and red dot points represent bacteria that were higher in abundance in cancer and cancer-adjacent tissues respectively. CLR-abundance of paired cancer and cancer-adjacent samples from **(B)** 16s rRNA sequencing and **(C)** TCMA WGS sequencing datasets. Wilcoxon matched pairs signed rank test was performed for both 16s rRNA (n = 287) and TCMA (n = 22) datasets. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

3.4 Tissue microbiota diversity correlates with cancer functional gene expression signatures.

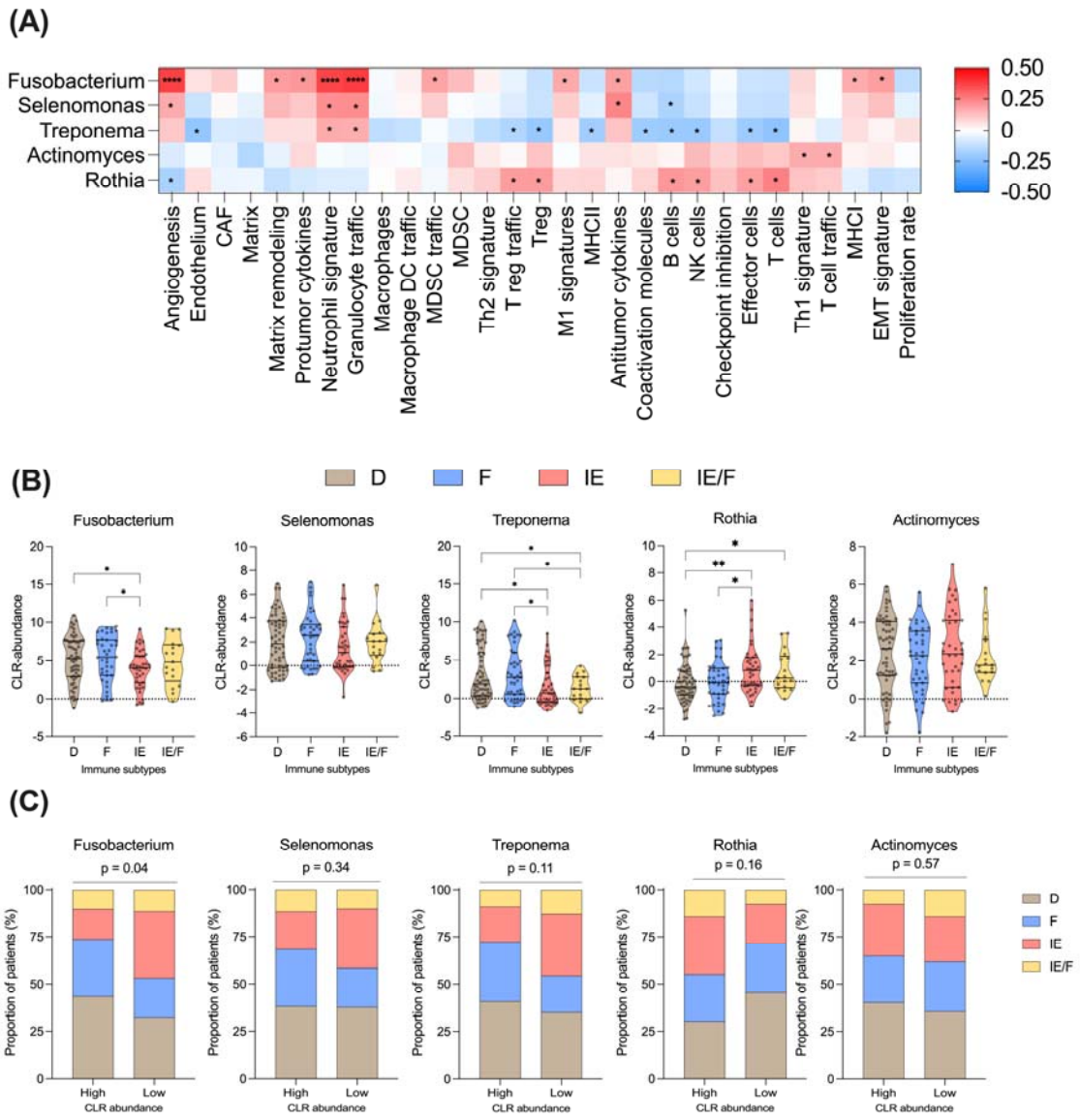
Since *Fusobacterium, Selenomonas, Treponema, Actinomyces,* and *Rothia* displayed significant differences between paired cancer and cancer-adjacent tissues, we performed correlation analyses to investigate the possible relationship between these genera and the tumour transcriptional profile and patient clinical features found in matched TCGA patients (n = 156).[105] Here, TCGA transcriptomic data were classified into 29 functional gene expression signatures (FGES), which represent major functional components and characteristics of cancer cell populations.[105] These 29 FGES can then be used to further classify cancers into four major immune subtypes (Desert, Fibrotic, Immune-enriched/non-fibrotic, and Immune-enriched/fibrotic).[105] We correlated the CLR-abundance of *Fusobacterium, Selenomonas, Treponema, Actinomyces,* and *Rothia* from TCGA-HNSC patients with their respective FGES scores and immune subtype.

We first correlated CLR-abundance with the FGE signatures. The CLR-abundance of *Fusobacterium* correlated (r > 0.3, p < 0.0001) with FGES related to angiogenesis, neutrophils and granulocyte traffic (Figure 5A). Other FGES such as matrix remodelling, protumour cytokines, MDSC traffic, M1 signature, antitumour cytokine, MHCI and EMT

signatures also positively correlated ($p < 0.05$) to CLR-abundance of *Fusobacterium* (Figure 5A). The CLR-abundance of *Selenomonas* showed a positive correlation ($p < 0.05$) to angiogenesis, neutrophil signature, granulocyte traffic and antitumour cytokines signatures, while negatively correlating ($p < 0.05$) to B cells (Figure 4A). Lastly, CLR-abundance of Treponema displayed a negative correlation ($p < 0.05$) to endothelium, T reg traffic, T reg, MHCII, Coactivation molecules, B cells, NK cells, Effector cells and T cells, while positively correlating to ($p < 0.05$) neutrophils and granulocyte traffic (Figure 5A).

Next, we investigated how CLR-abundance correlated to tissue immune subtyping. Cancer tissues classified as immune deserts (D) and fibrotic (F) which lack immune cell enrichment correlated with higher *Fusobacterium* and *Treponema* CLR-abundance. On the other hand, cancer tissues that are immune-enriched / non-fibrotic (IE) or immune-enriched / fibrotic (IE/F) correlated with greater *Rothia*. No significant correlation in immune subtypes were observed for *Selenomonas* and *Actinomyces* (Figure 5B). To identify the differences in immune subtypes between high and low CLR-abundance of each bacterial genera, we further segregated patients based on the upper and lower 35% CLR-abundance quartiles. As expected, patients with IE and IE/F tumour subtypes showed significant association with low CLR-abundance of *Fusobacterium* (chi-square test, $p = 0.04$). While not reaching statistical significance, more patients with IE and IE/F tumour subtypes have low CLR-abundance of *Selenomonas* (chi-square test, $p = 0.33$) and *Treponema* (chi-square test, $p = 0.11$), opposite to high CLR-abundance for *Rothia* (Figure 5C). Conversely, patients with D and F subtypes had higher CLR-abundance of *Fusobacterium*, *Selenomonas* or *Treponema* (Figure 5C). Lastly, the proportion of patients in each immune subtype were similar in high and low CLR-abundance *Actinomyces* groups. Taken together, these show that *Fusobacterium*,

451 *Selenomonas* or *Treponema* are associated with poor T-cell infiltration compared to *Rothia*
452 which may have implications in selecting patients suitable for immunotherapy.



453

454 **Figure 5: Correlation analysis of *Fusobacterium, Selenomonas, Treponema, Rothia***
455 **and *Actinomyces* to the tumour transcriptional profiles.**

456 **(A)** 29 functional gene expression (FGES) signature scores derived from Bagaev et al
457 (2021) were used to correlated with CLR-abundance of genera *Fusobacterium,*
458 *Selenomonas, Treponema, Actinomyces, and Rothia,* using Pearson's correlation
459 method. Asterisk (*) represents significant correlation ($p < 0.05$), and red and blue

scales represents positive and negative correlation respectively. **(B)** The CLR-abundance of each bacterial genera within each tumour microenvironment immune subtype (D – Desert, F – Fibrotic, IE – Immune-enriched/Non-fibrotic, IE/F – Immune-enriched/Fibrotic). Kruskal-Wallis test with uncorrected Dunn's test was performed to compare CLR-abundance in all immune groups. *p < 0.05, **p < 0.01. **(C)** The proportion of patients in each tumour immune subtype with high and low CLR-abundance in each bacterial genera. High and low bacteria CLR-abundance groups were determined by upper and lower 35% quartiles respectively. Chi-squared test was performed to determine association between high/low bacterial genera CLR-abundance and proportion of patients in each tumour subtype.

3.5 Evaluation of microbiota abundance with clinical features and survival

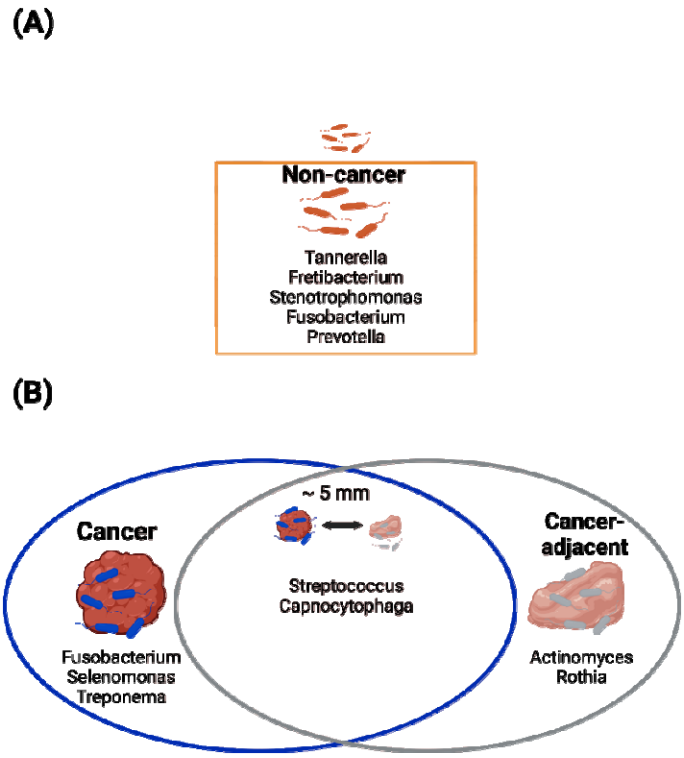
Univariate and multivariate Cox proportional hazard models were used to investigate the association between the intratumoral microbiota and clinical features. Univariate Cox proportional hazard model identified that current smokers (HR 2.235, 95% CI 1.146 – 4.359, p = 0.018), HPV-negative (HR 2.273, 95% CI 1.158 – 4.459, p = 0.017), and low CLR-abundance of *Fusobacterium* (HR 0.8883, 95% CI 0.8183 – 0.9642, p = 0.005) were risk factors for reduced overall survival (Table 2). Further multivariate Cox proportional hazard models identified that HPV-negative (HR 2.853, 95% CI 1.1991 – 6.7882, p = 0.0178) and low CLR-abundance of *Fusobacterium* (Continuous: HR 0.8482, 95% CI 0.7758– 0.9273, p = 0.0003; Low: HR 2.579, 95% CI 1.3687 – 4.860, p = 0.0034) were independent hazards for overall survival, but not current smokers (Table 2).

Table 2: Univariate and multivariable Cox proportional hazard models for overall survival

			Univariate		Multivariable	
		n	HR (95% CI)	p-value	HR (95% CI)	p-value
Age (years)	< 65	106				
	≥ 65	47	0.9832 (0.594 – 1.626)	0.947		
Sex	Female	41				
	Male	112	0.868 (0.525 – 1.436)	0.581		
Staging	I	4				
	II	30	2.331 (0.302 – 17.97)	0.417		
	III	31	2.275 (0.2950 – 17.54)	0.430		
	IV	87	3.275 (0.4492 – 23.87)	0.242		
HPV status	Positive	37				
	Negative	107	2.273 (1.158 – 4.459)	0.017*	2.853 (1.1991 – 6.7882)	0.0178*
Smoking	Non-smoker	37				
	Current	43	2.235 (1.146 – 4.359)	0.018*	1.3788 (0.5383 – 3.5317)	0.50329
	Previous	71	1.488 (0.7804 – 2.838)	0.227	0.7821 (0.3130 – 1.9545)	0.59894
Fusobacterium	Continuous	153	0.8883 (0.8183 - 0.9642)	0.005**	0.8482 (0.7758– 0.9273)	0.0003**
	High	53				
	Low	53	2.0592 (1.17 – 3.625)	0.0123*	2.579 (1.3687 – 4.860)	0.0034*
Selenomonas	Continuous		0.9712 (0.8714 –1.082)	0.597		
	High	53				
	Low	52	1.205 (0.7094 – 2.048)	0.49		
Treponema	Continuous	153	0.9467 (0.8768 – 0.719)	0.162		
	High	53				
	Low	53	1.432 (0.8092 – 2.535)	0.217		
Rothia	Continuous	153	1.029 (0.8936 – 1.184)	0.694		
	High	54				
	Low	54	0.6552 (0.3585 – 1.198)	0.17		
Actinomyces	Continuous	153	0.9652 (0.8512 – 1.094)	0.58		
	High	53				
	Low	52	1.006 (0.5679 – 1.783)	0.983		

4 Discussion:

Several studies have investigated the microbial signature in HNSC using different sequencing approaches and sample types, such as tissues, swabs, and oral fluids. However, these studies have reported inconsistent findings regarding the presence of specific bacterial genera. Consequently, a consensus microbial signature for head and neck tissues has yet to be established. In this study, we aimed to address this gap by conducting a meta-analysis of 11 studies and presenting a consensus tissue microbiota signature for head and neck tissues. We analyzed 16S rRNA sequencing datasets from 903 tissue samples, including 396 cancer tissues, 251 cancer-adjacent tissues, and 256 non-cancer tissues. Our analysis revealed significant differences in the abundance of 33 bacterial genera among the various tissue types. Specifically, we observed that cancer tissues and cancer-adjacent tissues exhibited greater similarity to each other compared to non-cancer tissues. These findings suggest distinct microbial profiles in cancer and cancer-adjacent tissues compared to non-cancer tissues. Non-cancer tissues exhibited the lowest differences in β-diversity and contained elevated levels of bacterial genera such as *Tannerella*, *Fretibacterium*, *Stenotrophomonas*, *Fusobacterium*, and *Prevotella* (Figure 6A). While cancer and cancer-adjacent tissues displayed similar microbiota based on β-diversity indexes, further analysis using paired and unpaired univariate methods enabled differentiation of these tissues at the genera level (Figure 6B). Importantly, these abundance signatures were validated using additional data from TCMA. Matching TCMA samples with transcriptomic data derived from TCGA) and clinical features provided insights into the contributions of individual genera in HNSC. Notably, we found that a high abundance of *Fusobacterium* was associated with better overall survival in HNSC patients Overall, our study contributes to the establishment of a consensus tissue microbiota signature for HNSC, shedding light on the distinct microbial profiles in different tissue types and their potential implications for clinical outcomes.



512

513 **Figure 6: Summary of bacteria genera within cancer, cancer-adjacent and non-cancer**
514 **tissue samples. (A)** Elevated microbiota within non-cancer tissues compared to cancer and
515 cancer-adjacent tissues. **(B)** Elevated bacteria genera between cancer and cancer-adjacent
516 tissues.

517

518 Both multivariate and univariate discriminant analyses was able to differentiate different
519 tissue sample types based on microbial abundance. As previously reported, cancer and
520 cancer-adjacent tissues were more similar in microbial diversity when compared to non-
521 cancer tissues.[36-38, 42, 43, 45, 49, 51, 52, 55, 57-59, 85] At the genus level, both paired and unpaired
522 abundance analysis of cancer and cancer-adjacent tissues showed consistent enrichment for
523 *Fusobacterium* and *Rothia* in cancer tissues.[36, 37, 45, 52, 55, 58, 59] In contrast, *Prevotella* was
524 enriched within cancer tissues compared to cancer-adjacent tissues, and no differences were
525 observed for *Streptococcus*.[36-38, 41, 43, 45, 51, 52, 55, 57-59, 85]

Previous studies have reported conflicting result where *Fusobacterium* was more in cancer tissues as compared to non-cancer and cancer-adjacent tissues.[38, 42, 52, 57, 58] However, we found that Fusobacterium was most abundant in non-cancer tissues. *Fusobacterium* is an abundant commensal bacteria found largely in the oral cavity (buccal, hard palate, gingiva, tonsils, tongue) and saliva of healthy individuals, suggesting a potential role within the healthy oral microbiota.[106-108] *In vitro* experiments in HNSC cell lines showed that *Fusobacterium nucleatum* infection promotes cancer cell invasion, proliferation, autophagy, and PD-L1 expression.[109-113] It is unknown whether there are strain and species level differences found in *Fusobacterium* isolated in cancer and non-cancer tissues to explain such seemingly contradictory findings. Additionally, non-cancer tissue from cancer patients may also have different tissue microbiota profiles from healthy donor tissues which is currently unavailable for this study. Also, most of the experiments showing an oncogenic role for *F. nucleatum* were carried out *in vitro* and thus did not consider a potential mitigating role of the immune system. Moreover, the abundance of *F. nucleatum* both in absolute terms and relative to other bacteria present in the tumour microbiota might influence the oncogenic potential of *F. nucleatum*. Further experiments are required to evaluate the role of *Fusobacterium* in HNSC.

We observed that *Streptococcus*, another highly abundant oral commensal genera[106-108], was increased specifically in cancer and cancer-adjacent tissue when compared to non-cancer tissues. However, there was no significant difference in *Streptococcus* abundance between cancer and cancer-adjacent tissues. Within the oral cavity, certain pathogenic *Streptococcus* species, like *S. mutans*, can contribute to periodontitis by acidifying the environment.[114] In oral cancer, *S. mutans* has been shown to promote tumour proliferation and invasion,

551 potentially through upregulation of IL-6 in infected cells[115]. On the other hand, *Streptococcus*
552 species from the *mitis* (*S. oralis*, *S parasanguinis*, *S.mitis*) and *sanguinis* (*S. sanguinis*, *S.*
553 *gordonii*) groups, can break down lactic acid or pyruvate into hydrogen peroxide, thereby
554 antagonising pathogenic species such as *S. mutans*.[114] In oral cancer, *S. mitis*, *S. salivarius*, *S.*
555 *anginosus* were found to display anti-tumour effects, including reducing cancer cell viability
556 and promoting $CD8^+$ cytotoxic T cell responses[116-120]. These findings indicate that the
557 abundance of specific *Streptococcus* species may contribute to pathogenesis, disease severity,
558 or exert anti-tumour effects. It is important to note that these studies underscore the
559 limitations of identifying microbiota at the genus level using short-read 16S rRNA
560 sequencing. To address these limitations, recent advances in sequencing technologies such as
561 long-read 16S rRNA amplicon sequencing (e.g., PacBio, Nanopore) or shotgun
562 metagenomics can be employed to reveal species- or strain-specific diversity within the
563 microbiota.[121-123] Such advancements can provide a more comprehensive understanding of
564 the specific species and strains that play a role in oral cancer pathogenesis and anti-tumour
565 effects.
566
567 To compare the metabolic potential of different head and neck tissue types, a functional
568 prediction analysis was performed using PICRUSt2 on the 16S rRNA sequencing data. The
569 analysis revealed an enrichment of several amino acids and metabolites, including L-
570 aspartate, L-asparagine, acetate, butanoate, and lactate, in cancer tissues compared to non-
571 cancer and cancer-adjacent tissues. L-aspartate and L-asparagine, as substrates for nucleotide
572 biosynthesis and regulators of amino acid homeostasis and anabolic metabolism, have been
573 reported to promote tumour proliferation.[124-126] Butanoate, acetate, and lactate can serve as
574 energy sources for cells by converting into acetyl-CoA, which can then be utilized in the
575 tricarboxylic acid (TCA) cycle to produce ATP.[127-130] The role of butanoate in tumorigenesis

576 depends on the specific tumour and the TME, as it can exhibit tumour-promoting or
577 suppressive properties. [31, 131-133] Lactate, a well-studied metabolite produced by both cancer
578 cells and bacteria, can modulate the TME by inactivating natural killer cells, promoting
579 polarisation of M2-like tumour-associated macrophages, and stimulating the growth of T-
580 regulatory cells.[134] Collectively, these findings suggest that bacteria infiltrating HNSC tissues
581 possess functional capacities that may promote cancer progression. Further validation studies
582 are warranted to better understand the role of these metabolic pathways in HNSC and the
583 contribution of bacteria in shaping the TME.
584
585 We further explored the relationship between the abundance of the five cancer-associated
586 bacterial genera, *Fusobacterium, Selenomonas, Treponema, Actinomyces,* and *Rothia,* and the
587 TME phenotype and clinical outcomes. *Fusobacterium* was associated with a lack of T-cell
588 immune infiltration in HNSC, similar to colorectal and oesophageal cancers.[21, 135-137]
589 Furthermore, *Fusobacterium* can chemoattract neutrophils via release of SCFA and can also
590 modulate neutrophils and endothelial cell functions *in vitro*.[138-142] Interestingly, we observed
591 that patients with low levels of *Fusobacterium* within the tumour tissue had shorter overall
592 survival, consistent to previous reports in HNSC.[45, 60, 143] In contrast, opposite findings have
593 been reported for colorectal, gastric and oesophageal cancers, suggesting that *Fusobacterium*
594 may have a different role in HNSC.[144-147] We also found that *Treponema* correlated with an
595 lack of immune infiltration in HNSC. Although the effect of *Treponema* infiltration in HNSC
596 is still unknown, these bacteria have been associated with an upregulation of immune
597 suppressive cells and can suppress innate immune responses.[148-150] In our analysis, *Rothia*
598 was found to correlate with an immune-enriched TME. Limited information is available
599 regarding the role of *Rothia* in cancer; however, *Rothia dentocariosa* has been shown to
600 induce Toll-like receptor 2 (TLR-2) mediated TNF-alpha inflammatory response in human

601 embryonic kidney cells and THP-1 monocytes.[151] *Selenomonas* and *Actinomyces* did not
602 significantly correlate with TME subtypes in our analysis. However, *Selenomonas sputigena*
603 infected gingival epithelial cells can promote neutrophil and monocyte recruitment.[152]
604 *Actinomyces* has been associated with young-onset colorectal cancers, showing a preferential
605 localisation with cancer-associated fibroblasts in the TME.[153] These findings underscore the
606 importance of validating and understanding the underlying mechanisms through which these
607 bacteria can modulate the tumour microenvironment in HNSC.
608
609 This study represents the first comprehensive comparison of 16S rRNA (V3-V5) microbial
610 sequencing across multiple studies to identify consensus HNSC-associated microbiota
611 signatures in cancer, cancer-adjacent and non-cancer tissues. To ensure consistency, a
612 uniform bioinformatics approach was employed. However, it is important to acknowledge the
613 inherent limitations of this study. Variations in sample collection, preparation, and
614 sequencing among different laboratories introduce batch effects that could contribute to the
615 inconsistencies observed across different reports. To mitigate these effects, we applied
616 PLSDA-batch adjustment to the pooled datasets.[91] Conventional short-read 16S rRNA
617 sequencing provides information only up to the genus level, which restricts the ability to
618 identify specific bacterial species or strains that may be relevant to disease outcomes.[121]
619 Overcoming this limitation would require advanced sequencing technologies such as long-
620 read 16S rRNA amplicon sequencing or shotgun metagenomics to reveal species- or strain-
621 level diversity within the microbiota. Furthermore, the availability of complete clinical
622 metadata in published datasets reporting 16S rRNA sequencing is limited, restricting our
623 ability to make comprehensive clinical associations. Therefore, our clinical associations were
624 primarily based on TCMA/TCGA datasets. Despite these limitations, this study confirms
625 distinct differences in the microbiota composition among cancer, cancer-adjacent and non-

cancer HNSC tissue samples. The strength of our study lies in the meta-analysis of a substantial number of samples, totalling 903. Additionally, our analysis indicates that a high load of *Fusobacterium* within HNSC tissues may be associated with a favourable survival outcome. The correlation analysis of the microbiota with functional predictions, functional gene enrichment signature, and immune subtyping of the tumour and TME provides novel avenues for further exploration.

In conclusion, our study establishes a consensus microbial signature for head and neck tissues, shedding light on the distinct microbial profiles present in head and neck cancer (HNSC). These findings have the potential to serve as targets for future treatment approaches in HNSC. Nevertheless, it is crucial to acknowledge the limitations identified in our study and recognize the need for further research to address these limitations. Additional investigations are required to gain a deeper understanding of the functional implications of the identified microbiota differences in HNSC. By addressing these gaps, we can advance our knowledge and pave the way for more effective therapeutic interventions in HNSC.

5. Conflict of Interest

Authors state no conflict of interest.

6. Author Contributions

Conceptualisation K.Y.,K.F.; methodology, investigation, and data analysis, K.Y.,R.L.,F.W., E.S., G.B., and L.M.; resources A.P.,P.W. and S.V.; writing - original draft preparation, K.Y., E.S., S.V., and K.F; writing-review and editing, K.Y., G.B., E.S., A.P., P.W., R.V., S.V., and K.F.; supervision R.V.,A.P., S.V., and K.F.; funding acquisition, A.P.,P.W., and S.V.; All authors have read and agreed to the published version of the manuscript.

7. Acknowledgments

This work is supported by an NHMRC investigator grant APP1196832 to P.W., a The Garnett Passe and Rodney Williams Senior Fellowship to S.V., and The University of Adelaide Postgraduate Research Scholarship to K.Y., R.L., F.W and L.M. Illustration in Figure 6 was generated using Biorender.

8. Data Availability Statement

The data used to support the findings of this study are included within the article and within supplementary material.

Reference:

1. Dohlman AB, Arguijo Mendoza D, Ding S, et al. The cancer microbiome atlas: a pan-cancer comparative analysis to distinguish tissue-resident microbiota from contaminants. *Cell Host Microbe*. Feb 10 2021;29(2):281-298.e5. doi:10.1016/j.chom.2020.12.001
2. Nejman D, Livyatan I, Fuks G, et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. *Science*. May 29 2020;368(6494):973-980. doi:10.1126/science.aay9189
3. Xuan C, Shamonki JM, Chung A, et al. Microbial dysbiosis is associated with human breast cancer. *PLoS One*. 2014;9(1):e83744. doi:10.1371/journal.pone.0083744
4. Jain T, Sharma P, Are AC, Vickers SM, Dudeja V. New Insights Into the Cancer-Microbiome-Immune Axis: Decrypting a Decade of Discoveries. *Front Immunol*. 2021;12:622064. doi:10.3389/fimmu.2021.622064
5. Xavier JB, Young VB, Skufca J, et al. The Cancer Microbiome: Distinguishing Direct and Indirect Effects Requires a Systemic View. *Trends Cancer*. Mar 2020;6(3):192-204. doi:10.1016/j.trecan.2020.01.004
6. Yang L, Li A, Wang Y, Zhang Y. Intratumoral microbiota: roles in cancer initiation, development and therapeutic efficacy. *Signal Transduct Target Ther*. Jan 16 2023;8(1):35. doi:10.1038/s41392-022-01304-4
7. Vétizou M, Pitt JM, Daillère R, et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science*. Nov 27 2015;350(6264):1079-84. doi:10.1126/science.aad1329
8. Sivan A, Corrales L, Hubert N, et al. Commensal Bifidobacterium promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science*. Nov 27 2015;350(6264):1084-9. doi:10.1126/science.aac4255
9. Iida N, Dzutsev A, Stewart CA, et al. Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science*. Nov 22 2013;342(6161):967-70. doi:10.1126/science.1240527
10. Viaud S, Saccheri F, Mignot G, et al. The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide. *Science*. Nov 22 2013;342(6161):971-6. doi:10.1126/science.1240537
11. Fu A, Yao B, Dong T, et al. Tumor-resident intracellular microbiota promotes metastatic colonization in breast cancer. *Cell*. Apr 14 2022;185(8):1356-1372.e26. doi:10.1016/j.cell.2022.02.027
12. Geller LT, Barzily-Rokni M, Danino T, et al. Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science*. Sep 15 2017;357(6356):1156-1160. doi:10.1126/science.aah5043
13. Yu T, Guo F, Yu Y, et al. Fusobacterium nucleatum Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. *Cell*. Jul 27 2017;170(3):548-563.e16. doi:10.1016/j.cell.2017.07.008
14. Spanogiannopoulos P, Kyaw TS, Guthrie BGH, et al. Host and gut bacteria share metabolic pathways for anti-cancer drug metabolism. *Nat Microbiol*. Oct 2022;7(10):1605-1620. doi:10.1038/s41564-022-01226-5
15. Gur C, Ibrahim Y, Isaacson B, et al. Binding of the Fap2 protein of Fusobacterium nucleatum to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity*. Feb 17 2015;42(2):344-355. doi:10.1016/j.immuni.2015.01.010
16. Kalaora S, Nagler A, Nejman D, et al. Identification of bacteria-derived HLA-bound peptides in melanoma. *Nature*. Apr 2021;592(7852):138-143. doi:10.1038/s41586-021-03368-8

17. Naghavian R, Faigle W, Oldrati P, et al. Microbial peptides activate tumour-infiltrating lymphocytes in glioblastoma. *Nature*. May 2023;617(7962):807-817. doi:10.1038/s41586-023-06081-w
18. Abed J, Emgård JE, Zamir G, et al. Fap2 Mediates Fusobacterium nucleatum Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc. *Cell Host Microbe*. Aug 10 2016;20(2):215-25. doi:10.1016/j.chom.2016.07.006
19. Parhi L, Alon-Maimon T, Sol A, et al. Breast cancer colonization by Fusobacterium nucleatum accelerates tumor growth and metastatic progression. *Nat Commun*. Jun 26 2020;11(1):3259. doi:10.1038/s41467-020-16967-2
20. Zhang S, Yang Y, Weng W, et al. Fusobacterium nucleatum promotes chemoresistance to 5-fluorouracil by upregulation of BIRC3 expression in colorectal cancer. *J Exp Clin Cancer Res*. Jan 10 2019;38(1):14. doi:10.1186/s13046-018-0985-y
21. Wu J, Li Q, Fu X. Fusobacterium nucleatum Contributes to the Carcinogenesis of Colorectal Cancer by Inducing Inflammation and Suppressing Host Immunity. *Transl Oncol*. Jun 2019;12(6):846-851. doi:10.1016/j.tranon.2019.03.003
22. Gur C, Maalouf N, Shhadeh A, et al. Fusobacterium nucleatum supresses anti-tumor immunity by activating CEACAM1. *Oncoimmunology*. 2019;8(6):e1581531. doi:10.1080/2162402x.2019.1581531
23. Yoon Y, Kim G, Jeon BN, Fang S, Park H. Bifidobacterium Strain-Specific Enhances the Efficacy of Cancer Therapeutics in Tumor-Bearing Mice. *Cancers (Basel)*. Feb 25 2021;13(5)doi:10.3390/cancers13050957
24. Asadollahi P, Ghanavati R, Rohani M, Razavi S, Esghaei M, Talebi M. Anti-cancer effects of Bifidobacterium species in colon cancer cells and a mouse model of carcinogenesis. *PLoS One*. 2020;15(5):e0232930. doi:10.1371/journal.pone.0232930
25. Rossi T, Vergara D, Fanini F, Maffia M, Bravaccini S, Pirini F. Microbiota-Derived Metabolites in Tumor Progression and Metastasis. *Int J Mol Sci*. Aug 12 2020;21(16)doi:10.3390/ijms21165786
26. Krautkramer KA, Fan J, Bäckhed F. Gut microbial metabolites as multi-kingdom intermediates. *Nat Rev Microbiol*. Feb 2021;19(2):77-94. doi:10.1038/s41579-020-0438-4
27. Bachem A, Makhlouf C, Binger KJ, et al. Microbiota-Derived Short-Chain Fatty Acids Promote the Memory Potential of Antigen-Activated CD8(+) T Cells. *Immunity*. Aug 20 2019;51(2):285-297.e5. doi:10.1016/j.immuni.2019.06.002
28. Wang W, Fang D, Zhang H, et al. Sodium Butyrate Selectively Kills Cancer Cells and Inhibits Migration in Colorectal Cancer by Targeting Thioredoxin-1. *Onco Targets Ther*. 2020;13:4691-4704. doi:10.2147/ott.S235575
29. Liang Y, Rao Z, Du D, Wang Y, Fang T. Butyrate prevents the migration and invasion, and aerobic glycolysis in gastric cancer via inhibiting Wnt/β-catenin/c-Myc signaling. *Drug Dev Res*. May 2023;84(3):532-541. doi:10.1002/ddr.22043
30. He Y, Fu L, Li Y, et al. Gut microbial metabolites facilitate anticancer therapy efficacy by modulating cytotoxic CD8(+) T cell immunity. *Cell Metab*. May 4 2021;33(5):988-1000.e7. doi:10.1016/j.cmet.2021.03.002
31. Okumura S, Konishi Y, Narukawa M, et al. Gut bacteria identified in colorectal cancer patients promote tumourigenesis via butyrate secretion. *Nat Commun*. Sep 28 2021;12(1):5674. doi:10.1038/s41467-021-25965-x
32. Zaiatz-Bittencourt V, Jones F, Tosetto M, et al. Butyrate limits human natural killer cell effector function. *Sci Rep*. Feb 15 2023;13(1):2715. doi:10.1038/s41598-023-29731-5
33. Bender MJ, McPherson AC, Phelps CM, et al. Dietary tryptophan metabolite released by intratumoral Lactobacillus reuteri facilitates immune checkpoint inhibitor treatment. *Cell*. Apr 27 2023;186(9):1846-1862.e26. doi:10.1016/j.cell.2023.03.011

34. Tintelnot J, Xu Y, Lesker TR, et al. Microbiota-derived 3-IAA influences chemotherapy efficacy in pancreatic cancer. *Nature*. Mar 2023;615(7950):168-174. doi:10.1038/s41586-023-05728-y

35. Hezaveh K, Shinde RS, Klötgen A, et al. Tryptophan-derived microbial metabolites activate the aryl hydrocarbon receptor in tumor-associated macrophages to suppress anti-tumor immunity. *Immunity*. Feb 8 2022;55(2):324-340.e8. doi:10.1016/j.immuni.2022.01.006

36. Chang C, Geng F, Shi X, et al. The prevalence rate of periodontal pathogens and its association with oral squamous cell carcinoma. *Appl Microbiol Biotechnol*. Feb 2019;103(3):1393-1404. doi:10.1007/s00253-018-9475-6

37. Zhou J, Wang L, Yuan R, et al. Signatures of Mucosal Microbiome in Oral Squamous Cell Carcinoma Identified Using a Random Forest Model. *Cancer Manag Res*. 2020;12:5353-5363. doi:10.2147/cmar.S251021

38. Torralba MG, Aleti G, Li W, et al. Oral Microbial Species and Virulence Factors Associated with Oral Squamous Cell Carcinoma. *Microb Ecol*. Nov 2021;82(4):1030-1046. doi:10.1007/s00248-020-01596-5

39. Henrich B, Rumming M, Sczyrba A, et al. Mycoplasma salivarium as a dominant coloniser of Fanconi anaemia associated oral carcinoma. *PLoS One*. 2014;9(3):e92297. doi:10.1371/journal.pone.0092297

40. Chan JYK, Ng CWK, Lan L, et al. Restoration of the Oral Microbiota After Surgery for Head and Neck Squamous Cell Carcinoma Is Associated With Patient Outcomes. *Front Oncol*. 2021;11:737843. doi:10.3389/fonc.2021.737843

41. De Martin A, Lütge M, Stanossek Y, et al. Distinct microbial communities colonize tonsillar squamous cell carcinoma. *Oncoimmunology*. 2021;10(1):1945202. doi:10.1080/2162402x.2021.1945202

42. Zakrzewski M, Gannon OM, Panizza BJ, Saunders NA, Antonsson A. Human papillomavirus infection and tumor microenvironment are associated with the microbiota in patients with oropharyngeal cancers-pilot study. *Head Neck*. Nov 2021;43(11):3324-3330. doi:10.1002/hed.26821

43. Sarkar P, Malik S, Laha S, et al. Dysbiosis of Oral Microbiota During Oral Squamous Cell Carcinoma Development. *Front Oncol*. 2021;11:614448. doi:10.3389/fonc.2021.614448

44. Zhou X, Hao Y, Peng X, et al. The Clinical Potential of Oral Microbiota as a Screening Tool for Oral Squamous Cell Carcinomas. *Front Cell Infect Microbiol*. 2021;11:728933. doi:10.3389/fcimb.2021.728933

45. Chen Z, Wong PY, Ng CWK, et al. The Intersection between Oral Microbiota, Host Gene Methylation and Patient Outcomes in Head and Neck Squamous Cell Carcinoma. *Cancers (Basel)*. Nov 18 2020;12(11)doi:10.3390/cancers12113425

46. Perera M, Al-Hebshi NN, Perera I, et al. Inflammatory Bacteriome and Oral Squamous Cell Carcinoma. *J Dent Res*. Jun 2018;97(6):725-732. doi:10.1177/0022034518767118

47. Al-Hebshi NN, Nasher AT, Maryoud MY, et al. Inflammatory bacteriome featuring Fusobacterium nucleatum and Pseudomonas aeruginosa identified in association with oral squamous cell carcinoma. *Sci Rep*. May 12 2017;7(1):1834. doi:10.1038/s41598-017-02079-3

48. Schmidt BL, Kuczynski J, Bhattacharya A, et al. Changes in abundance of oral microbiota associated with oral cancer. *PLoS One*. 2014;9(6):e98741. doi:10.1371/journal.pone.0098741

49. Wang H, Funchain P, Bebek G, et al. Microbiomic differences in tumor and paired-normal tissue in head and neck squamous cell carcinomas. *Genome Med*. Feb 7 2017;9(1):14. doi:10.1186/s13073-017-0405-5

50. Guerrero-Preston R, Godoy-Vitorino F, Jedlicka A, et al. 16S rRNA amplicon sequencing identifies microbiota associated with oral cancer, human papilloma virus infection and surgical treatment. *Oncotarget*. Aug 9 2016;7(32):51320-51334. doi:10.18632/oncotarget.9710

51. Gong H, Shi Y, Zhou X, et al. Microbiota in the Throat and Risk Factors for Laryngeal Carcinoma. *Appl Environ Microbiol*. Dec 2014;80(23):7356-63. doi:10.1128/aem.02329-14

52. Gong HL, Shi Y, Zhou L, et al. The Composition of Microbiome in Larynx and the Throat Biodiversity between Laryngeal Squamous Cell Carcinoma Patients and Control Population. *PLoS One*. 2013;8(6):e66476. doi:10.1371/journal.pone.0066476

53. Pushalkar S, Ji X, Li Y, et al. Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. *BMC Microbiol*. Jul 20 2012;12:144. doi:10.1186/1471-2180-12-144

54. Burcher KM, Burcher JT, Inscore L, Bloomer CH, Furdui CM, Porosnicu M. A Review of the Role of Oral Microbiome in the Development, Detection, and Management of Head and Neck Squamous Cell Cancers. *Cancers (Basel).* Aug 25 2022;14(17)doi:10.3390/cancers14174116

55. Yang K, Wang Y, Zhang S, et al. Oral Microbiota Analysis of Tissue Pairs and Saliva Samples From Patients With Oral Squamous Cell Carcinoma - A Pilot Study. *Front Microbiol*. 2021;12:719601. doi:10.3389/fmicb.2021.719601

56. Gopinath D, Kunnath Menon R, Chun Wie C, et al. Salivary bacterial shifts in oral leukoplakia resemble the dysbiotic oral cancer bacteriome. *J Oral Microbiol*. Dec 9 2020;13(1):1857998. doi:10.1080/20002297.2020.1857998

57. Gong H, Shi Y, Xiao X, et al. Alterations of microbiota structure in the larynx relevant to laryngeal carcinoma. *Sci Rep*. Jul 14 2017;7(1):5507. doi:10.1038/s41598-017-05576-7

58. Dong Z, Zhang C, Zhao Q, et al. Alterations of bacterial communities of vocal cord mucous membrane increases the risk for glottic laryngeal squamous cell carcinoma. *J Cancer*. 2021;12(13):4049-4063. doi:10.7150/jca.54221

59. 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*. Aug 30 2017;7(1):9934. doi:10.1038/s41598-017-09786-x

60. Chan JYK, Cheung MK, Lan L, et al. Characterization of oral microbiota in HPV and non-HPV head and neck squamous cell carcinoma and its association with patient outcomes. *Oral Oncol*. Dec 2022;135:106245. doi:10.1016/j.oraloncology.2022.106245

61. Choi YS, Kim Y, Yoon HJ, et al. The presence of bacteria within tissue provides insights into the pathogenesis of oral lichen planus. *Sci Rep*. Jul 7 2016;6:29186. doi:10.1038/srep29186

62. Baek K, Lee J, Lee A, et al. Characterization of intratissue bacterial communities and isolation of Escherichia coli from oral lichen planus lesions. *Sci Rep*. Feb 26 2020;10(1):3495. doi:10.1038/s41598-020-60449-w

63. Bao K, Li X, Poveda L, et al. Proteome and Microbiome Mapping of Human Gingival Tissue in Health and Disease. *Front Cell Infect Microbiol*. 2020;10:588155. doi:10.3389/fcimb.2020.588155

64. Sawant S, Dugad J, Parikh D, Srinivasan S, Singh H. Identification & correlation of bacterial diversity in oral cancer and long-term tobacco chewers- A case-control pilot study. *J Med Microbiol*. Sep 2021;70(9)doi:10.1099/jmm.0.001417

65. Frank DN, Qiu Y, Cao Y, et al. A dysbiotic microbiome promotes head and neck squamous cell carcinoma. *Oncogene*. Feb 2022;41(9):1269-1280. doi:10.1038/s41388-021-02137-1
66. Sharma AK, DeBusk WT, Stepanov I, Gomez A, Khariwala SS. Oral Microbiome Profiling in Smokers with and without Head and Neck Cancer Reveals Variations Between Health and Disease. *Cancer Prev Res (Phila)*. May 2020;13(5):463-474. doi:10.1158/1940-6207.Capr-19-0459
67. Lau HC, Hsueh CY, Gong H, et al. Oropharynx microbiota transitions in hypopharyngeal carcinoma treatment of induced chemotherapy followed by surgery. *BMC Microbiol*. Nov 9 2021;21(1):310. doi:10.1186/s12866-021-02362-4
68. Hsueh CY, Gong H, Cong N, et al. Throat Microbial Community Structure and Functional Changes in Postsurgery Laryngeal Carcinoma Patients. *Appl Environ Microbiol*. Nov 24 2020;86(24)doi:10.1128/aem.01849-20
69. Panda M, Rai AK, Rahman T, et al. Alterations of salivary microbial community associated with oropharyngeal and hypopharyngeal squamous cell carcinoma patients. *Arch Microbiol*. May 2020;202(4):785-805. doi:10.1007/s00203-019-01790-1
70. Vesty A, Gear K, Biswas K, Radcliff FJ, Taylor MW, Douglas RG. Microbial and inflammatory-based salivary biomarkers of head and neck squamous cell carcinoma. *Clin Exp Dent Res*. Dec 2018;4(6):255-262. doi:10.1002/cre2.139
71. Lee WH, Chen HM, Yang SF, et al. Bacterial alterations in salivary microbiota and their association in oral cancer. *Sci Rep*. Nov 28 2017;7(1):16540. doi:10.1038/s41598-017-16418-x
72. Amer A, Galvin S, Healy CM, Moran GP. The Microbiome of Potentially Malignant Oral Leukoplakia Exhibits Enrichment for Fusobacterium, Leptotrichia, Campylobacter, and Rothia Species. *Front Microbiol*. 2017;8:2391. doi:10.3389/fmicb.2017.02391
73. Chen MY, Chen JW, Wu LW, et al. Carcinogenesis of Male Oral Submucous Fibrosis Alters Salivary Microbiomes. *J Dent Res*. Apr 2021;100(4):397-405. doi:10.1177/0022034520968750
74. Debelius JW, Huang T, Cai Y, et al. Subspecies Niche Specialization in the Oral Microbiome Is Associated with Nasopharyngeal Carcinoma Risk. *mSystems*. Jul 7 2020;5(4)doi:10.1128/mSystems.00065-20
75. Kumpitsch C, Moissl-Eichinger C, Pock J, Thurnher D, Wolf A. Preliminary insights into the impact of primary radiochemotherapy on the salivary microbiome in head and neck squamous cell carcinoma. *Sci Rep*. Oct 6 2020;10(1):16582. doi:10.1038/s41598-020-73515-0
76. Wolf A, Moissl-Eichinger C, Perras A, Koskinen K, Tomazic PV, Thurnher D. The salivary microbiome as an indicator of carcinogenesis in patients with oropharyngeal squamous cell carcinoma: A pilot study. *Sci Rep*. Jul 19 2017;7(1):5867. doi:10.1038/s41598-017-06361-2
77. Zhu XX, Yang XJ, Chao YL, et al. The Potential Effect of Oral Microbiota in the Prediction of Mucositis During Radiotherapy for Nasopharyngeal Carcinoma. *EBioMedicine*. Apr 2017;18:23-31. doi:10.1016/j.ebiom.2017.02.002
78. Furquim CP, Soares GM, Ribeiro LL, et al. The Salivary Microbiome and Oral Cancer Risk: a Pilot Study in Fanconi Anemia. *J Dent Res*. Mar 2017;96(3):292-299. doi:10.1177/0022034516678169
79. Zhang J, Liu H, Liang X, et al. Investigation of salivary function and oral microbiota of radiation caries-free people with nasopharyngeal carcinoma. *PLoS One*. 2015;10(4):e0123137. doi:10.1371/journal.pone.0123137

80. Hu YJ, Wang Q, Jiang YT, et al. Characterization of oral bacterial diversity of irradiated patients by high-throughput sequencing. *Int J Oral Sci*. Mar 2013;5(1):21-5. doi:10.1038/ijos.2013.15
81. Pushalkar S, Mane SP, Ji X, et al. Microbial diversity in saliva of oral squamous cell carcinoma. *FEMS Immunol Med Microbiol*. Apr 2011;61(3):269-77. doi:10.1111/j.1574-695X.2010.00773.x
82. Hu YJ, Shao ZY, Wang Q, et al. Exploring the dynamic core microbiome of plaque microbiota during head-and-neck radiotherapy using pyrosequencing. *PLoS One*. 2013;8(2):e56343. doi:10.1371/journal.pone.0056343
83. Minarovits J. Anaerobic bacterial communities associated with oral carcinoma: Intratumoral, surface-biofilm and salivary microbiota. *Anaerobe*. Apr 2021;68:102300. doi:10.1016/j.anaerobe.2020.102300
84. Orlandi E, Iacovelli NA, Tombolini V, et al. Potential role of microbiome in oncogenesis, outcome prediction and therapeutic targeting for head and neck cancer. *Oral Oncol*. Dec 2019;99:104453. doi:10.1016/j.oraloncology.2019.104453
85. Gopinath D, Menon RK, Wie CC, et al. Differences in the bacteriome of swab, saliva, and tissue biopsies in oral cancer. *Sci Rep*. Jan 13 2021;11(1):1181. doi:10.1038/s41598-020-80859-0
86. Jain V, Baraniya D, El-Hadedy DE, et al. Integrative Metatranscriptomic Analysis Reveals Disease-specific Microbiome–host Interactions in Oral Squamous Cell Carcinoma. *Cancer Res Commun*. 2023;3(5):807-820. doi:10.1158/2767-9764.CRC-22-0349
87. Wang X, Zhao Z, Tang N, et al. Microbial Community Analysis of Saliva and Biopsies in Patients With Oral Lichen Planus. *Front Microbiol*. 2020;11:629. doi:10.3389/fmicb.2020.00629
88. Zhang Z, Feng Q, Li M, et al. Age-Related Cancer-Associated Microbiota Potentially Promotes Oral Squamous Cell Cancer Tumorigenesis by Distinct Mechanisms. *Front Microbiol*. 2022;13:852566. doi:10.3389/fmicb.2022.852566
89. Nie F, Wang L, Huang Y, et al. Characteristics of Microbial Distribution in Different Oral Niches of Oral Squamous Cell Carcinoma. *Front Cell Infect Microbiol*. 2022;12:905653. doi:10.3389/fcimb.2022.905653
90. Qiao H, Li H, Wen X, Tan X, Yang C, Liu N. Multi-Omics Integration Reveals the Crucial Role of Fusobacterium in the Inflammatory Immune Microenvironment in Head and Neck Squamous Cell Carcinoma. *Microbiol Spectr*. Aug 31 2022;10(4):e0106822. doi:10.1128/spectrum.01068-22
91. Wang Y, KA LC. PLSDA-batch: a multivariate framework to correct for batch effects in microbiome data. *Brief Bioinform*. Mar 19 2023;24(2)doi:10.1093/bib/bbac622
92. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Rev Esp Cardiol (Engl Ed)*. Sep 2021;74(9):790-799. Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. doi:10.1016/j.rec.2021.07.010
93. Choudhary S. pysradb: A Python package to query next-generation sequencing metadata and data from NCBI Sequence Read Archive. *F1000Res*. 2019;8:532. doi:10.12688/f1000research.18676.1
94. Leinonen R, Sugawara H, Shumway M. The sequence read archive. *Nucleic Acids Res*. Jan 2011;39(Database issue):D19-21. doi:10.1093/nar/gkq1019
95. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. Aug 2019;37(8):852-857. doi:10.1038/s41587-019-0209-9

96. Rohart F, Gautier B, Singh A, KA LC. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol*. Nov 2017;13(11):e1005752. doi:10.1371/journal.pcbi.1005752
97. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 2013;8(4):e61217. doi:10.1371/journal.pone.0061217
98. Moentadj R, Wang Y, Bowerman K, et al. Streptococcus species enriched in the oral cavity of patients with RA are a source of peptidoglycan-polysaccharide polymers that can induce arthritis in mice. *Ann Rheum Dis*. May 2021;80(5):573-581. doi:10.1136/annrheumdis-2020-219009
99. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front Microbiol*. 2017;8:2224. doi:10.3389/fmicb.2017.02224
100. Khomich M, Måge I, Rud I, Berget I. Analysing microbiome intervention design studies: Comparison of alternative multivariate statistical methods. *PLoS One*. 2021;16(11):e0259973. doi:10.1371/journal.pone.0259973
101. Cao Y, Dong Q, Wang D, Zhang P, Liu Y, Niu C. microbiomeMarker: an R/Bioconductor package for microbiome marker identification and visualization. *Bioinformatics*. Aug 10 2022;38(16):4027-4029. doi:10.1093/bioinformatics/btac438
102. Douglas GM, Maffei VJ, Zaneveld JR, et al. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol*. Jun 2020;38(6):685-688. doi:10.1038/s41587-020-0548-6
103. Caspi R, Billington R, Keseler IM, et al. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res*. Jan 8 2020;48(D1):D445-d453. doi:10.1093/nar/gkz862
104. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. Apr 2 2013;6(269):pl1. doi:10.1126/scisignal.2004088
105. Bagaev A, Kotlov N, Nomie K, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell*. Jun 14 2021;39(6):845-865.e7. doi:10.1016/j.ccell.2021.04.014
106. Eren AM, Borisy GG, Huse SM, Mark Welch JL. Oligotyping analysis of the human oral microbiome. *Proc Natl Acad Sci U S A*. Jul 15 2014;111(28):E2875-84. doi:10.1073/pnas.1409644111
107. Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. Biogeography of a human oral microbiome at the micron scale. *Proc Natl Acad Sci U S A*. Feb 9 2016;113(6):E791-800. doi:10.1073/pnas.1522149113
108. Wilbert SA, Mark Welch JL, Borisy GG. Spatial Ecology of the Human Tongue Dorsum Microbiome. *Cell Rep*. Mar 24 2020;30(12):4003-4015.e3. doi:10.1016/j.celrep.2020.02.097
109. Shao W, Fujiwara N, Mouri Y, et al. Conversion from epithelial to partial-EMT phenotype by Fusobacterium nucleatum infection promotes invasion of oral cancer cells. *Sci Rep*. Jul 22 2021;11(1):14943. doi:10.1038/s41598-021-94384-1
110. Chen G, Gao C, Jiang S, et al. Fusobacterium nucleatum outer membrane vesicles activate autophagy to promote oral cancer metastasis. *J Adv Res*. Apr 13 2023;doi:10.1016/j.jare.2023.04.002
111. Geng F, Zhang Y, Lu Z, Zhang S, Pan Y. Fusobacterium nucleatum Caused DNA Damage and Promoted Cell Proliferation by the Ku70/p53 Pathway in Oral Cancer Cells. *DNA Cell Biol*. Jan 2020;39(1):144-151. doi:10.1089/dna.2019.5064
112. Zhang S, Li C, Liu J, et al. Fusobacterium nucleatum promotes epithelial-mesenchymal transiton through regulation of the lncRNA MIR4435-2HG/miR-296-

5p/Akt2/SNAI1 signaling pathway. *Febs j*. Sep 2020;287(18):4032-4047. doi:10.1111/febs.15233

113. Binder Gallimidi A, Fischman S, Revach B, et al. Periodontal pathogens Porphyromonas gingivalis and Fusobacterium nucleatum promote tumor progression in an oral-specific chemical carcinogenesis model. *Oncotarget*. Sep 8 2015;6(26):22613-23. doi:10.18632/oncotarget.4209

114. Baty JJ, Stoner SN, Scoffield JA. Oral Commensal Streptococci: Gatekeepers of the Oral Cavity. *J Bacteriol*. Nov 15 2022;204(11):e0025722. doi:10.1128/jb.00257-22

115. Tsai MS, Chen YY, Chen WC, Chen MF. Streptococcus mutans promotes tumor progression in oral squamous cell carcinoma. *J Cancer*. 2022;13(12):3358-3367. doi:10.7150/jca.73310

116. Baraniya D, Jain V, Lucarelli R, et al. Screening of Health-Associated Oral Bacteria for Anticancer Properties in vitro. *Front Cell Infect Microbiol*. 2020;10:575656. doi:10.3389/fcimb.2020.575656

117. Xu Y, Jia Y, Chen L, Gao J, Yang D. Effect of Streptococcus anginosus on biological response of tongue squamous cell carcinoma cells. *BMC Oral Health*. Mar 20 2021;21(1):141. doi:10.1186/s12903-021-01505-3

118. Baraniya D, Chitrala KN, Al-Hebshi NN. Global transcriptional response of oral squamous cell carcinoma cell lines to health-associated oral bacteria - an in vitro study. *J Oral Microbiol*. 2022;14(1):2073866. doi:10.1080/20002297.2022.2073866

119. Wang J, Sun F, Lin X, Li Z, Mao X, Jiang C. Cytotoxic T cell responses to Streptococcus are associated with improved prognosis of oral squamous cell carcinoma. *Exp Cell Res*. Jan 1 2018;362(1):203-208. doi:10.1016/j.yexcr.2017.11.018

120. Wang J, Yang L, Mao X, Li Z, Lin X, Jiang C. Streptococcus salivarius-mediated CD8(+) T cell stimulation required antigen presentation by macrophages in oral squamous cell carcinoma. *Exp Cell Res*. May 15 2018;366(2):121-126. doi:10.1016/j.yexcr.2018.03.007

121. Curry KD, Wang Q, Nute MG, et al. Emu: species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. *Nat Methods*. Jul 2022;19(7):845-853. doi:10.1038/s41592-022-01520-4

122. Johnson JS, Spakowicz DJ, Hong BY, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*. Nov 6 2019;10(1):5029. doi:10.1038/s41467-019-13036-1

123. Gehrig JL, Portik DM, Driscoll MD, et al. Finding the right fit: evaluation of short-read and long-read sequencing approaches to maximize the utility of clinical microbiome data. *Microb Genom*. Mar 2022;8(3)doi:10.1099/mgen.0.000794

124. Garcia-Bermudez J, Baudrier L, La K, et al. Aspartate is a limiting metabolite for cancer cell proliferation under hypoxia and in tumours. *Nat Cell Biol*. Jul 2018;20(7):775-781. doi:10.1038/s41556-018-0118-z

125. Krall AS, Xu S, Graeber TG, Braas D, Christofk HR. Asparagine promotes cancer cell proliferation through use as an amino acid exchange factor. *Nat Commun*. Apr 29 2016;7:11457. doi:10.1038/ncomms11457

126. Halbrook CJ, Thurston G, Boyer S, et al. Differential integrated stress response and asparagine production drive symbiosis and therapy resistance of pancreatic adenocarcinoma cells. *Nat Cancer*. Nov 2022;3(11):1386-1403. doi:10.1038/s43018-022-00463-1

127. Donohoe DR, Garge N, Zhang X, et al. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab*. May 4 2011;13(5):517-26. doi:10.1016/j.cmet.2011.02.018

128. Mashimo T, Pichumani K, Vemireddy V, et al. Acetate is a bioenergetic substrate for human glioblastoma and brain metastases. *Cell*. Dec 18 2014;159(7):1603-14. doi:10.1016/j.cell.2014.11.025

129. Lyssiotis CA, Cantley LC. Acetate fuels the cancer engine. *Cell*. Dec 18 2014;159(7):1492-4. doi:10.1016/j.cell.2014.12.009
130. Comerford SA, Huang Z, Du X, et al. Acetate dependence of tumors. *Cell*. Dec 18 2014;159(7):1591-602. doi:10.1016/j.cell.2014.11.020
131. Donohoe DR, Collins LB, Wali A, Bigler R, Sun W, Bultman SJ. The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Mol Cell*. Nov 30 2012;48(4):612-26. doi:10.1016/j.molcel.2012.08.033
132. Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell*. Jun 2 2016;165(6):1332-1345. doi:10.1016/j.cell.2016.05.041
133. van der Hee B, Wells JM. Microbial Regulation of Host Physiology by Short-chain Fatty Acids. *Trends Microbiol*. Aug 2021;29(8):700-712. doi:10.1016/j.tim.2021.02.001
134. Li Z, Wang Q, Huang X, et al. Lactate in the tumor microenvironment: A rising star for targeted tumor therapy. *Front Nutr*. 2023;10:1113739. doi:10.3389/fnut.2023.1113739
135. Mima K, Sukawa Y, Nishihara R, et al. Fusobacterium nucleatum and T Cells in Colorectal Carcinoma. *JAMA Oncol*. Aug 2015;1(5):653-61. doi:10.1001/jamaoncol.2015.1377
136. Kim HS, Kim CG, Kim WK, et al. Fusobacterium nucleatum induces a tumor microenvironment with diminished adaptive immunity against colorectal cancers. *Front Cell Infect Microbiol*. 2023;13:1101291. doi:10.3389/fcimb.2023.1101291
137. Kosumi K, Baba Y, Yamamura K, et al. Intratumour Fusobacterium nucleatum and immune response to oesophageal cancer. *Br J Cancer*. Apr 2023;128(6):1155-1165. doi:10.1038/s41416-022-02112-x
138. Dahlstrand Rudin A, Khamzeh A, Venkatakrishnan V, Basic A, Christenson K, Bylund J. Short chain fatty acids released by Fusobacterium nucleatum are neutrophil chemoattractants acting via free fatty acid receptor 2 (FFAR2). *Cell Microbiol*. Aug 2021;23(8):e13348. doi:10.1111/cmi.13348
139. Mendes RT, Nguyen D, Stephens D, et al. Endothelial Cell Response to Fusobacterium nucleatum. *Infect Immun*. Jul 2016;84(7):2141-2148. doi:10.1128/iai.01305-15
140. Wang Q, Zhao L, Xu C, Zhou J, Wu Y. Fusobacterium nucleatum stimulates monocyte adhesion to and transmigration through endothelial cells. *Arch Oral Biol*. Apr 2019;100:86-92. doi:10.1016/j.archoralbio.2019.02.013
141. Wright HJ, Chapple IL, Matthews JB, Cooper PR. Fusobacterium nucleatum regulation of neutrophil transcription. *J Periodontal Res*. Feb 2011;46(1):1-12. doi:10.1111/j.1600-0765.2010.01299.x
142. Zhou T, Meng X, Wang D, Fu W, Li X. Neutrophil Transcriptional Deregulation by the Periodontal Pathogen Fusobacterium nucleatum in Gastric Cancer: A Bioinformatic Study. *Dis Markers*. 2022;2022:9584507. doi:10.1155/2022/9584507
143. Neuzillet C, Marchais M, Vacher S, et al. Prognostic value of intratumoral Fusobacterium nucleatum and association with immune-related gene expression in oral squamous cell carcinoma patients. *Sci Rep*. Apr 12 2021;11(1):7870. doi:10.1038/s41598-021-86816-9
144. Mima K, Nishihara R, Qian ZR, et al. Fusobacterium nucleatum in colorectal carcinoma tissue and patient prognosis. *Gut*. Dec 2016;65(12):1973-1980. doi:10.1136/gutjnl-2015-310101
145. Lehr K, Nikitina D, Vilchez-Vargas R, et al. Microbial composition of tumorous and adjacent gastric tissue is associated with prognosis of gastric cancer. *Sci Rep*. Mar 21 2023;13(1):4640. doi:10.1038/s41598-023-31740-3

146. Hsieh YY, Kuo WL, Hsu WT, Tung SY, Li C. Fusobacterium Nucleatum-Induced Tumor Mutation Burden Predicts Poor Survival of Gastric Cancer Patients. *Cancers (Basel)*. Dec 30 2022;15(1)doi:10.3390/cancers15010269

147. Zhang N, Liu Y, Yang H, et al. Clinical Significance of Fusobacterium nucleatum Infection and Regulatory T Cell Enrichment in Esophageal Squamous Cell Carcinoma. *Pathol Oncol Res*. 2021;27:1609846. doi:10.3389/pore.2021.1609846

148. Jo AR, Baek KJ, Shin JE, Choi Y. Mechanisms of IL-8 suppression by Treponema denticola in gingival epithelial cells. *Immunol Cell Biol*. Feb 2014;92(2):139-47. doi:10.1038/icb.2013.80

149. Babolin C, Amedei A, Ozolins D, Zilevica A, D'Elios MM, de Bernard M. TpF1 from Treponema pallidum activates inflammasome and promotes the development of regulatory T cells. *J Immunol*. Aug 1 2011;187(3):1377-84. doi:10.4049/jimmunol.1100615

150. Hashimoto M, Asai Y, Ogawa T. Treponemal phospholipids inhibit innate immune responses induced by pathogen-associated molecular patterns. *J Biol Chem*. Nov 7 2003;278(45):44205-13. doi:10.1074/jbc.M306735200

151. Kataoka H, Taniguchi M, Fukamachi H, Arimoto T, Morisaki H, Kuwata H. Rothia dentocariosa induces TNF-alpha production in a TLR2-dependent manner. *Pathog Dis*. Jun 2014;71(1):65-8. doi:10.1111/2049-632x.12115

152. Hawkes CG, Hinson AN, Vashishta A, et al. Selenomonas sputigena Interactions with Gingival Epithelial Cells That Promote Inflammation. *Infect Immun*. Feb 16 2023;91(2):e0031922. doi:10.1128/iai.00319-22

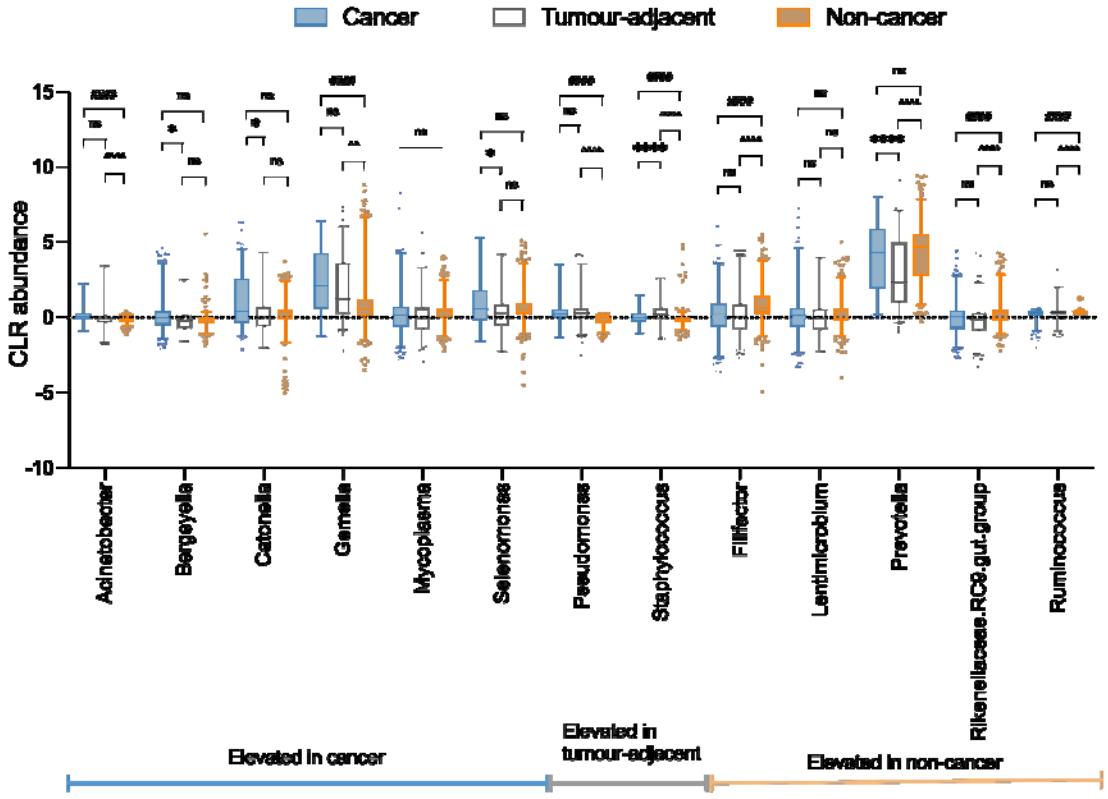
153. Xu Z, Lv Z, Chen F, et al. Dysbiosis of human tumor microbiome and aberrant residence of Actinomyces in tumor-associated fibroblasts in young-onset colorectal cancer. *Front Immunol*. 2022;13:1008975. doi:10.3389/fimmu.2022.1008975

1124 **Supplementary Figures**

1125

1126

1127



1128

1129 Figure S1: CLR-normalized abundances for remaining 13 bacteria between sample groups using unpaired
1130 Kruskal-Wallis test with Bonferroni's multiple comparison. Post-hoc Wilcoxon test with Bonferroni-Dunn's
1131 multiple comparison was performed to identify group-wise differences between Cancer – Non-cancer (#),
1132 Cancer – Cancer-adjacent (*), Non-cancer – Cancer-adjacent (^).

1133

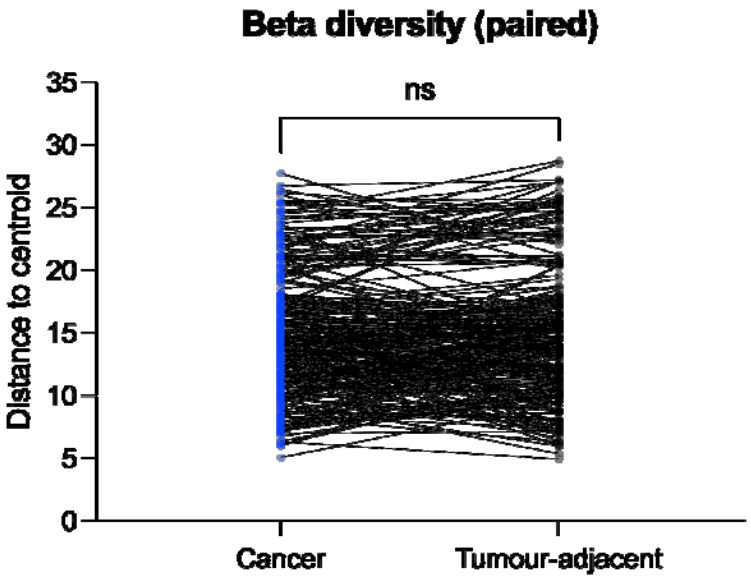


Figure S2: Beta-diversity between paired cancer and cancer-adjacent tissue samples. Paired Wilcoxon test was performed on Euclidean distance between each samples. No significant differences between cancer and cancer-adjacent tissue samples.