JOURNAL OF MEDICAL MICROBIOLOGY

RESEARCH ARTICLE

Yeo *et al.*, *Journal of Medical Microbiology* 2024;73:001799

DOI 10.1099/jmm.0.001799

MICROBIOLOGY SOCIETY

OPEN ACCESS

Identification of consensus head and neck cancer-associated microbiota signatures: a systematic review and 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,*]

Abstract

Introduction. Multiple reports have attempted to describe the tumour microbiota in head and neck cancer (HNSC).

Gap statement. However, these have failed to produce a consistent microbiota signature, which may undermine understanding the importance of bacterial-mediated effects in HNSC.

Aim. The aim of this study is to consolidate these datasets and identify a consensus microbiota signature in HNSC.

Methodology. We analysed 12 published HNSC 16S rRNA microbial datasets collected from cancer, cancer-adjacent and non-cancer tissues to generate a consensus microbiota signature. These signatures were then validated using The Cancer Microbiome Atlas (TCMA) database and correlated with the tumour microenvironment phenotypes and patient's clinical outcome.

Results. We identified a consensus microbial signature at the genus level to differentiate between HNSC sample types, with cancer and cancer-adjacent tissues sharing more similarity than non-cancer tissues. Univariate analysis on 16S rRNA datasets identified significant differences in the abundance of 34 bacterial genera among the tissue types. Paired cancer and cancer-adjacent tissue analyses in 16S rRNA and TCMA datasets identified increased abundance in *Fusobacterium* in cancer tissues and decreased abundance of *Atopobium*, *Rothia* and *Actinomyces* in cancer-adjacent tissues. Furthermore, these bacteria were associated with different tumour microenvironment phenotypes. Notably, high *Fusobacterium* signature was associated with high neutrophil (r=0.37, P<0.0001), angiogenesis (r=0.38, P<0.0001) and granulocyte signatures (r=0.38, P<0.0001) and better overall patient survival [continuous: HR 0.8482, 95% confidence interval (CI) 0.7758–0.9273, P=0.0003].

Conclusion. Our meta-analysis demonstrates a consensus microbiota signature for HNSC, highlighting its potential importance in this disease.

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 (HNSCs) containing one of the highest level of intratumoral microbial infiltrates, while glioblastomas having the least amount of microbes [1, 2]. This 'intratumoral

Received 30 July 2023; Accepted 05 January 2024; Published 01 February 2024

Author affiliations: [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, Woodville South SA, 5011, Australia; [3]Department of Haematology and Oncology, Basil Hetzel Institute for Translational Health Research and The Queen Elizabeth Hospital, Central Adelaide Local Health Network, Woodville South SA, 5011, Australia.
***Correspondence:** Kevin Fenix, kevin.fenix@adelaide.edu.au

Keywords: tumour microbiota; head and neck cancer; 16S rRNA sequencing; meta-analysis.

Abbreviations: ANOSIM, Analysis of similarities; CLR, Central log-ratio; FGES, functional gene expression signatures; HNSC, Head and neck cancer; PCA, Principal component analysis; PERMANOVA, Permutational multivariate analysis of variance; PICRUST2, Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2; rRNA, Ribosomal RNA; SCFA, Short chain fatty acids; sPLS-DA, Multivariate sparse partial least squares discriminant analysis; TCGA, The Cancer Genome Atlas; TCMA, The Cancer Microbiome Atlas; TME, Tumour microenvironment; WGS, Whole genome sequencing.

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

Eight supplementary figures and ten supplementary tables are available with the online version of this article.

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) [3, 4]. 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 [5, 6].

Direct interaction between specific bacterial species and the tumour and the TME can induce chemoresistance, promote tumour progression, enhance therapeutic responses and modulate anti-tumour immunity through various mechanisms [7–9]. Bacteria can metabolize an active drug into its inactive form or induce autophagy in cancer cells, which can promote chemoresistance [9]. Moreover, specific bacterial species can mount or suppress anti-tumour responses [10, 11]. Most notably, *Fusobacterium nucleatum* colocalizes 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 and carcinoembryonic antigen-related cell adhesion molecule 1 receptors or sugar groups (e.g. tumour-expressed galactose-N-acetylgalactosamine), which may then promote chemoresistance and suppress anti-tumour immunity [8, 12–15]. Alternatively, *Bifidobacterium* species enhance anti-tumour immunity and the efficacy of PD-1 immunotherapy responses [6, 16, 17].

The release of bacterial metabolites such as short-chain fatty acids (SCFAs), amino acids, vitamins and bile acids can indirectly affect the tumour and the TME [18, 19]. 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 [20, 21]. However, butyrate has also been shown to have pro-tumorigenic effects by inducing senescence-associated inflammatory phenotypes and inhibiting natural killer cell functions [22, 23]. Bacteria-derived indole and its derivatives (e.g. indole-3-acetic 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 [24, 25]. 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 [26–28]. Most of these studies compared the microbiota diversity and bacterial relative abundance between cancer and healthy samples using 16S ribosomal RNA (rRNA) sequencing [26–28], while two studies additionally correlated the impact of the microbiota with matched transcriptome analysis [29, 30]. 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 [26–28]. Most bacteria identified in HNSC are oral commensal bacteria from the genera *Streptococcus, Rothia, Fusobacterium, Haemophilus* and *Prevotella* [31–33]. 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 [34]. However, inconsistencies are observed for microbes such as *Streptococcus*, *Actinomyces* and *Prevotella*, meaning that it is necessary to identify a consensus microbiota signature for HNSC [27, 31, 32, 35, 36]. Previous systematic reviews have tried to identify and describe the dysbiosis in HNSC using processed [37–43] or re-analysed [44] data from studies containing heterogeneous HNSC sample types such as saliva, swab, oral rinse and tissues. Differences in sequencing techniques, data processing and sample types among the original studies were mostly unaccounted for in these reviews and may limit interpretation of the overall findings.

In this study, we systematically reviewed the literature and performed a meta-analysis to consolidate the currently heterogeneous HNSC-associated microbiota data. The selected 16S rRNA sequencing datasets were analysed consistently to minimize variability between different sample cohorts and adjusted for batch effects [45]. 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 TME and clinical outcomes.

METHODS

This study was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement. A PRISMA checklist for the abstract and review is available in the Supplementary Material.

Search and study selection

The following criteria were used to select datasets: (1) tissue samples, (2) presence of metadata to distinguish sample types and (3) Illumina short-read amplicon sequencing of 16S rRNA V3 to V5 primers (Fig. 1, Table 1). The exclusion criteria were (1) non-tissue samples (i.e. saliva, swabs and oral wash), (2) samples without metadata and sequencing data access and (3) sequencing data obtained from techniques other than Illumina short-read amplicon sequencing of 16S rRNA V3 to V5 primers. National Center for Biotechnology Information (NCBI) PubMed, Scopus and NCBI Sequence Read Archive (SRA) database search was performed on 16 August 2022 by two independent reviewers and datasets after this date were not included (Table S1, available in the online

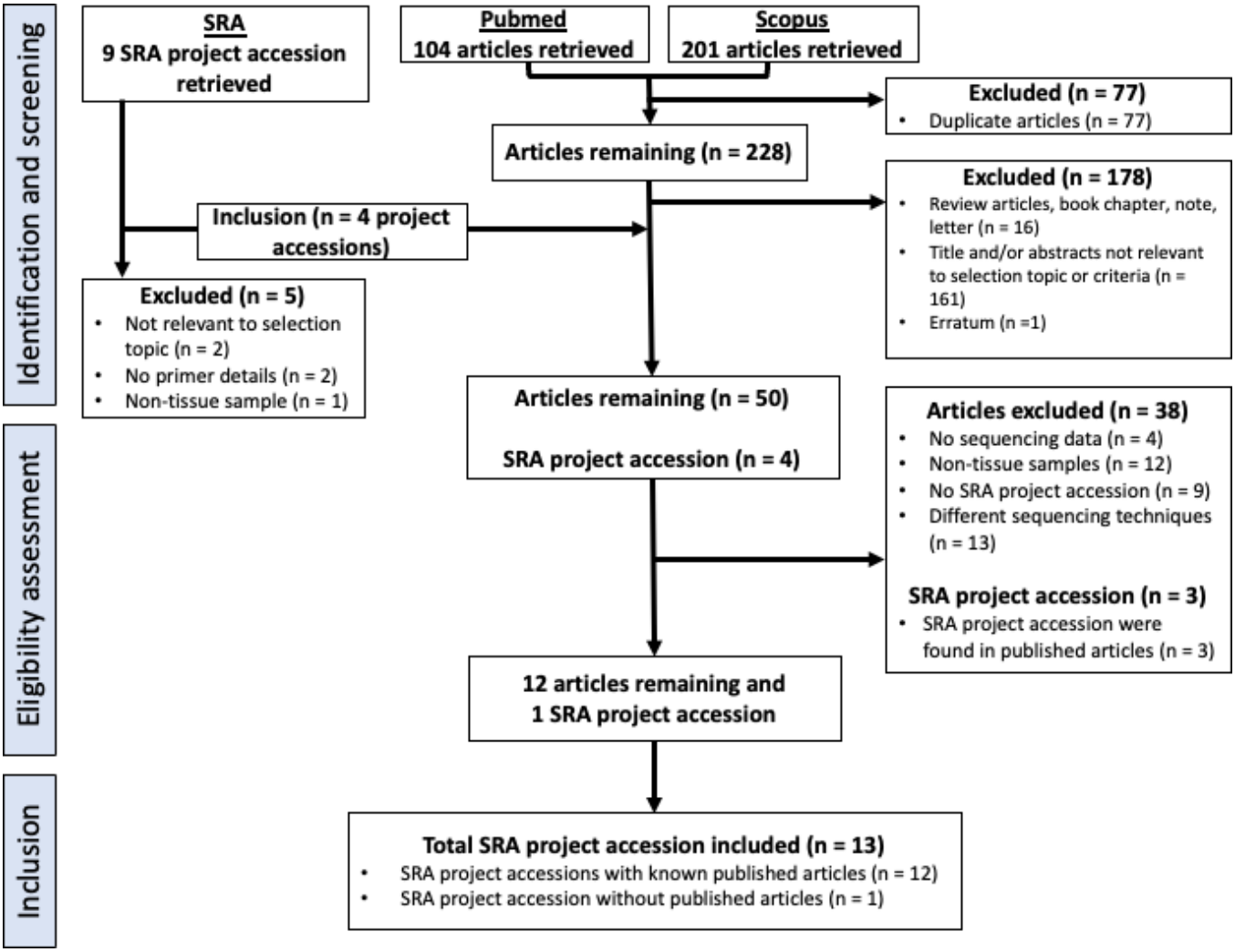

Fig. 1. Study selection flow chart.

version of this article). Risk of bias (RoB) assessment was performed using RoB2 assessment spreadsheet (ROB2_IRPG_beta_v9), and search terms are provided in Table S1 [46].

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

Previously published raw sequences were retrieved from the NCBI SRA using pysradb (Fig. S1) [47]. 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 >5 mm 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 [48]. 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 unit (OTU) classification using QIIME2 and SILVA reference database (version silva-138–99-nb-classifier) [49].

Raw microbial reads were filtered, central log ratio (CLR)-transformed and batch-adjusted using Phyloseq and mixOmics as described previously [50, 51]. Batch adjustments have been assessed and PLSDA-batch correction values were used for subsequent 16S rRNA data analysis (Figs S2–S5). As microbiome datasets are inherently compositional [52], CLR transformation addresses generate scale-invariant values, which allows datasets to remain unaffected by variations in library sizes among samples. Briefly, low abundances 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 [50]. The CLR abundance was used for subsequent statistical and discriminant analyses. A total of 1032 SRA samples from 13 SRA projects were downloaded. After filtering, 938 SRA samples remained from 12 SRA projects (Tables 1 and S2).

Table 1. Study accession and sample size post-filtering

Accession no.	Sample size			Primers
	Cancer	Cancer-adjacent	Non-cancer	
PRJNA412445	16	0	0	V4–V5
PRJNA555458	0	0	5	V3–V4
PRJNA596113	102	53	0	V3–V4
PRJNA597251	19	18	0	V3–V4
PRJNA666746	50	50	0	V3–V4
PRJNA666891	6	0	9	V4
PRJNA685226	13	13	0	V3–V4
PRJNA699728	37	0	200	V4
PRJNA736864	19	20	0	V3–V4
PRJNA803155	40	0	0	V4–V5
PRJNA822685	75	79	0	V3–V4
PRJNA866676	37	36	41	V3–V4

Discriminant analysis of 16S rRNA dataset

To discriminate the microbial signatures between sample types, we employed both multivariate and univariate discriminant analyses. For β-diversity analysis, the CLR abundance of all genera was ordinated using Euclidean distance and plotted on a principal component analysis (PCA) plot using the mixOmics R package. The β-diversity for each sample was calculated as distance to centroid for each tissue group 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 tests 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 means to distinguish between sample types.

Multivariate sparse partial least squares discriminant analysis (sPLS-DA) was applied on batch-adjusted datasets to identify discriminating genera within each sample type [50]. The area under curve (AUC) of the receiver operating characteristics (ROC) curve was calculated using mixOmics in R v4.1.3 [50, 53]. 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. A heatmap of all representative bacteria in each sPLS-DA was presented with sample type clustered according to Euclidean distance and Ward's linkage.

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

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 [54]. Functional abundance was processed and analysed similarly as described for raw microbial reads. The univariate Kruskal–Wallis test and the post-hoc Wilcoxon test were performed as previously described to compare differences between groups.

Reanalysis of tissue microbiome data from TCMA

Decontaminated microbial read counts 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 the 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 above. As samples were already pre-processed in the TCMA dataset, no further filtering or batch adjustment was required. Microbiome statistical analysis was performed similarly to that for 16S sequencing datasets. Metadata were obtained from cBioPortal for Cancer Genomics [55]. The β-diversity of tissue samples from different sites was compared to determine whether samples from different sites contained different bacterial communities (Fig. S7). Furthermore, β-diversity analysis was compared between cancer and cancer-adjacent tissues for paired tissues, similar to 16S rRNA analysis (Fig. S8).

Microbiome correlation analysis with TME and survival analysis

The TME immune subtype and 29 functional gene expression signature (FGES) scores were previously described by Bagaev *et al.* using transcriptomic datasets from TCGA [56]. The 29 FGESs represent the major functional components and immune, stromal and other cellular populations of the tumour [56]. Pearson's correlation test was applied to determine the correlation between FGES scores and selected bacterial genera. The four TME immune subtypes were – desert (D), fibrotic (F), immune-enriched (IE) and immune-enriched/fibrotic (IE/F) (described in Table S2) [56]. Specifically, tissues with IE and IE/F phenotypes contain high T-cell infiltration, while D and F phenotypes have low T-cell infiltration (Table S2) [56]. Using a cut-off of high (top 35th percentile) and low (bottom 35th percentile) CLR abundances, the proportion of each patient within each of the four TME subtypes was 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 analyses. The chi-squared (χ^2) test was performed in Prism 9 to determine the association between high/low bacterial genera CLR abundance and the proportion of patients within each tumour subtype.

Statistical analysis

The Kruskal–Wallis test with Bonferroni's multiple comparison test was used for comparisons made between all tissue groups unless stated otherwise. The post-hoc Wilcoxon matched pairs signed rank test with Bonferroni's multiple comparison test was used to compare differences between unpaired tissue samples. For all paired cancer and cancer-adjacent samples, the Wilcoxon matched pairs signed rank test was performed. The univariate and multivariate Cox proportional hazard model analysis was performed using survminer in R. Statistical analysis was performed using R and Prism 9.

RESULTS

Study characteristics and dataset

Using our search strategy, a total of 9 SRA project accessions, 104 PubMed and 201 Scopus articles were retrieved. After removing duplicates and following our exclusion criteria (see Methods), 12 published articles were identified with a SRA project accession. We also found and included one additional SRA project accession (unpublished but uploaded dataset) (Fig. 1). In total, 13 SRA project accessions containing a total of 1032 head and neck tissue samples were collected [31, 32, 35, 57–63]. After data pre-processing, a total of 938 samples remained, consisting of 414 cancer, 269 cancer-adjacent and 255 non-cancer tissues (Table S2). These tissue samples comprised 508 male, 205 females and 225 unknown donor tissue samples (Table S2). The RoB2 assessment on studies showed the overall low risk of bias (Table S1) [46].

Multivariate analysis identifies homogeneous microbial abundance and functions between cancer and cancer-adjacent samples in contrast to non-cancer samples

After sample processing and aggregation of 16S rRNA data at the genus level, a total of 179 distinct bacterial genera were identified (Fig. S1). Differences in the microbiota and β-diversity between tissue types were assessed using PCA and PERMANOVA test (Fig. 2a, b). The β-diversity index was calculated for cancer [mean=14.3, 95% confidence interval (CI)=13.7–15.0] and cancer-adjacent (mean=15.1, 95% CI=14.4–15.8) tissues, revealing similar levels of β-diversity. In contrast, non-cancer tissues (mean=9.3, 95% CI=8.63–9.97) exhibited lower β-diversity (PERMANOVA – overall R^2=0.050, P=0.0001) (Fig. 2b). A post-hoc pairwise test identified significant differences in β-diversity between cancer and non-cancer (R^2=0.048, P=0.001), cancer and cancer-adjacent (R^2=0.013, P=0.001), and non-cancer and cancer-adjacent samples (R^2=0.071, P=0.001) (Fig. 2b). These findings were consistent with additional multivariate and univariate statistical analyses, ANOSIM (R=0.016, P=0.036) and FFMANOVA (P<0.0001) (Table S5).

Multivariate sparse partial least squares discriminant analysis (sPLS-DA) identified 126 representative bacterial genera in sPLS-DA components 1 and 2, which were discriminant between tissue types (Fig. 2c–e). The AUC values were computed for different sample comparisons: cancer versus others (AUC=0.76, P<0.05), non-cancer versus others (AUC=0.93, P<0.05) and cancer-adjacent versus others (AUC=0.84, P<0.05). These results demonstrate that sPLS-DA components 1 and 2 (Fig. 2d) can effectively differentiate between tissue types. Lastly, the majority of cancer and cancer-adjacent samples clustered together and were distinct from non-cancer samples, as determined by the Euclidean distance metric (Fig. 2e).

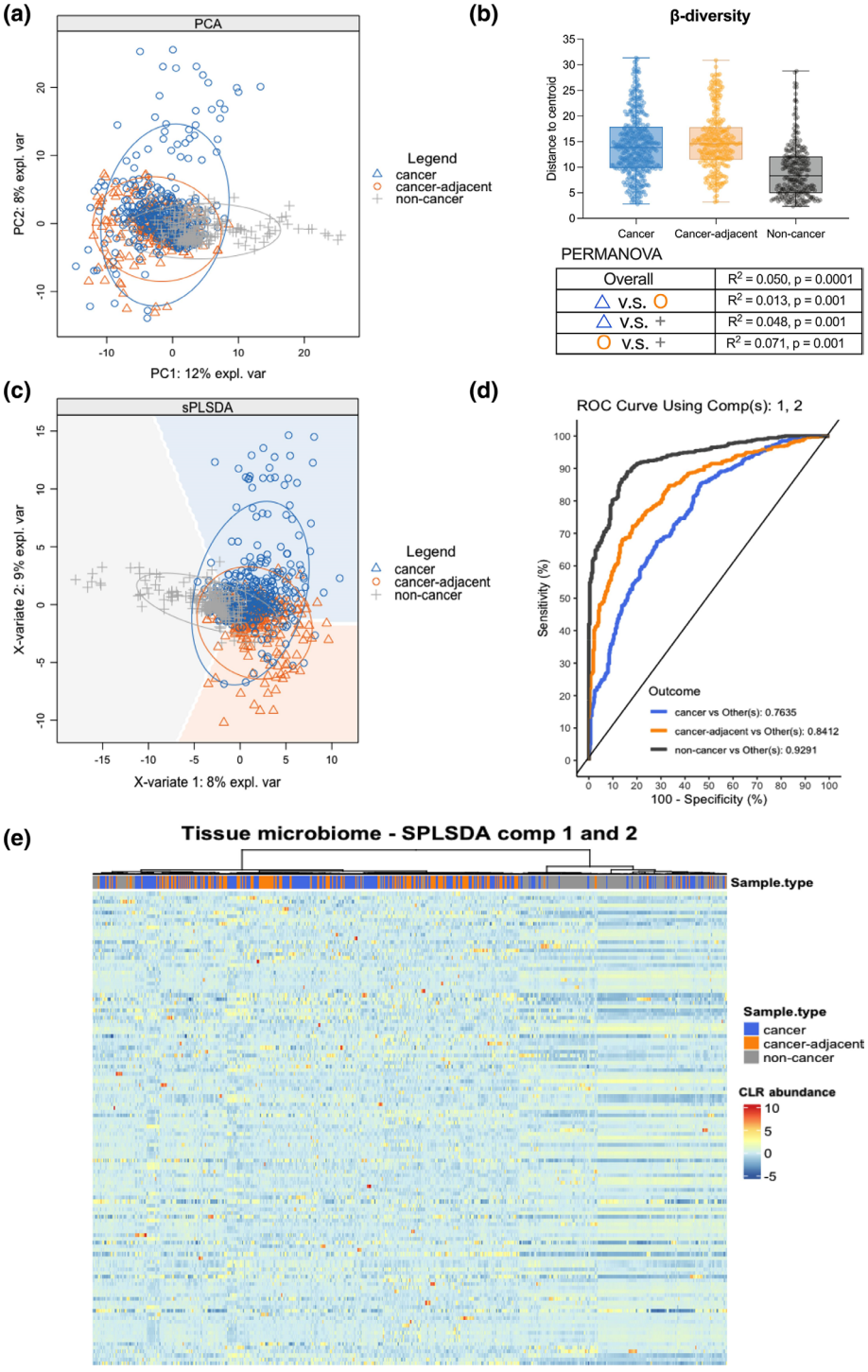

Fig. 2. Multivariate discriminant analysis (sPLS-DA and PERMANOVA) of tissue 16S rRNA microbiota to discriminate between cancer, cancer-adjacent and non-cancer tissues. (a) PCA plot of tissue CLR abundance microbiota based on Euclidean distance. (b) Dispersion of β-diversity (top-right panel) for each sample type, with error bar representing 95% CI. 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% CI 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 126 bacterial genera on components 1 and 2. (d) ROC curve and AUC values determined from sPLS-DA analysis were used to access the discriminatory potential of sPLS-DA components 1 and 2. (e) Heatmap representing 126 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.

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 179 bacterial genera, 34 were identified as significantly different among tissue types using the Kruskal–Wallis test (P_{adjust}<0.05) (Table S6). Notably, 31 of these were also identified as representative bacterial genera in sPLS-DA discriminant analysis (Table S6). These 34 genera are denoted as bacterial genera of interest (Table S6). The top 20 differentially abundant genera, based on the effect size (η^2), are presented in Fig. 3. The post-hoc unpaired Wilcoxon test with Bonferroni–Dunn's multiple comparison test was performed on these genera to determine the mean differences in the CLR abundance between tissue types (Fig. 3a, Table S7). Since most published studies compared cancer to non-cancer or cancer to cancer-adjacent tissues, we performed the post-hoc test for these comparisons (Fig. 3b). We identified 29 out of 34 genera as significantly different [P_{adjust} (#) <0.05] between cancer and non-cancer tissues (Fig. 3a, b, Table S7). Non-cancer tissues contained more *Fretibacterium* (CLR abundance diff.=1.41, SE=0.12), *Stenotrophomonas* (CLR abundance diff.=0.79, SE=0.12), *Fusobacterium* (CLR abundance diff.=0.87, SE=0.15) and *Tannerella* (CLR abundance diff.=0.72, SE=0.10), while cancer tissues had a greater CLR abundance of *Neisseria* (CLR abundance diff.=2.31, SE=0.16), *Capnocytophaga* (CLR abundance diff.=1.98, SE=0.15) and *Streptococcus* (CLR abundance diff.=1.86, SE=0.18) (Fig. 3a, b). *Capnocytophaga* abundance in cancer tissues was consistent with previous findings [36, 64, 65], while contradictory findings were identified for the abundance for *Streptococcus* [32, 36, 61, 64, 66] and *Fusobacterium* [32, 60, 61, 64, 66, 67].

For cancer and cancer-adjacent tissues, 11 out of 34 bacterial genera were significantly different [post-hoc unpaired Wilcoxon test P_{adjust} (*) <0.05] (Fig. 3a, c, Table S7). Similar to many studies, *Fusobacterium* (CLR abundance diff.=1.14, SE=0.20) displayed significantly higher CLR abundance in cancer tissues than cancer-adjacent tissues, while *Rothia* (CLR abundance diff.=0.98, SE=0.17), *Stenotrophomonas* (CLR abundance diff.=1.25, SE=0.14) and *Serratia* (CLR abundance diff.=0.62, SE=0.12) had higher CLR abundances in cancer-adjacent tissues than cancer tissues (Fig. 3a) [31, 33, 35, 59, 66–69]. Additionally, we found that *Prevotella* (CLR abundance diff.=0.86, SE=0.19) was elevated in cancer tissues as compared to cancer-adjacent tissues [35, 59, 66, 67, 69]. Unlike previous studies, we did not observe any significant differences in *Streptococcus* abundance between cancer and cancer-adjacent tissues [31, 33, 59, 66, 68–70].

Lastly, 29 of the 34 top bacterial genera were significantly different (post-hoc unpaired Wilcoxon test P_{adjust} <0.05) when comparing non-cancer to cancer-adjacent tissue samples (Fig. 3a, Table S7). Genera *Neisseria* (CLR abundance diff.=3.05, SE=0.19), *Rothia* (CLR abundance diff.=1.98, SE=0.15) and *Streptococcus* (CLR abundance diff.=1.71, SE=0.20) were higher in CLR abundance in cancer-adjacent tissues, while *Fusobacterium* (CLR abundance diff.=2.01, SE=0.16) and *Prevotella* (CLR abundance diff.=1.18, SE=0.20) were greater in CLR abundance in non-cancer tissues (Fig. 3a).

To provide functional insights into microbial abundance between cancer tissues and other tissue types, we applied PICRUSt2 to predict possible differences in MetaCyc pathway functional CLR abundance [54]. After filtering low-abundance functional pathways, we identified a total of 365 MetaCyc pathways. Using the Kruskal–Wallis test, 181 MetaCyc pathways were identified as significantly different among sample types (P_{adjust}<0.05) (Table S8). Post-hoc analysis identified 130/181 and 121/181 pathways that were significantly different between cancer–non-cancer and cancer–cancer-adjacent tissue comparisons, respectively (Table S9).

Cancer tissues, when compared to non-cancer tissues, were enriched in pathways involving the synthesis of L-methionine, inosine-5′-phosphate and cysteine, while non-cancer tissues were enriched in the degradation of L-lysine, L-glutamine, N-acetylglucosamine (GlcNac), N-acetylmannosamine (ManNac) and N-acetylneuraminate (Neu5Ac) (Fig. 3c). Cancer tissues were more similar to cancer-adjacent tissues, albeit enrichment was identified in pathways involving biosynthesis of adenosylcobalamin and cob(II)yrinate a,c-diamide, and enrichment in pathways involving degradation of L-glutamate, L-lysine, glucose and xylose when compared to cancer-adjacent tissues (Fig. 3c).

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

To understand microbial abundance differences between cancer tissues and cancer-adjacent tissues within the same patients, we performed the 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 were identified between the patient's paired cancer and cancer-adjacent tissues (Fig. S8).

However, 76 bacterial genera were significantly different between paired tissue samples (Fig. 4a, Table S10). Bacterial genera with the greatest differences in CLR abundance were then identified by using a cut-off of >0.4 and <−0.4 (Fig. 3a). Using this cut-off, we found that *Fusobacterium*, *Gemella*, *Prevotella*, *Bergeyella* and *Capnocytophaga* were elevated in cancer tissues vs cancer-adjacent tissues, while *Atopobium*, *Neisseria*, *Actinomyces*, *Granulicatella*, *Rothia* and *Stenotrophomonas* were greater in cancer-adjacent tissues compared to cancer tissues (Fig. 4a, b). Similarly, five of these bacteria were also found to be significantly different in unpaired tissue analysis (Tables S7 and S10). Contrary to published studies on unpaired samples, *Streptococcus,* an abundant oral commensal, was not significantly different in our paired sample analysis [31, 33, 59, 66, 68–70].

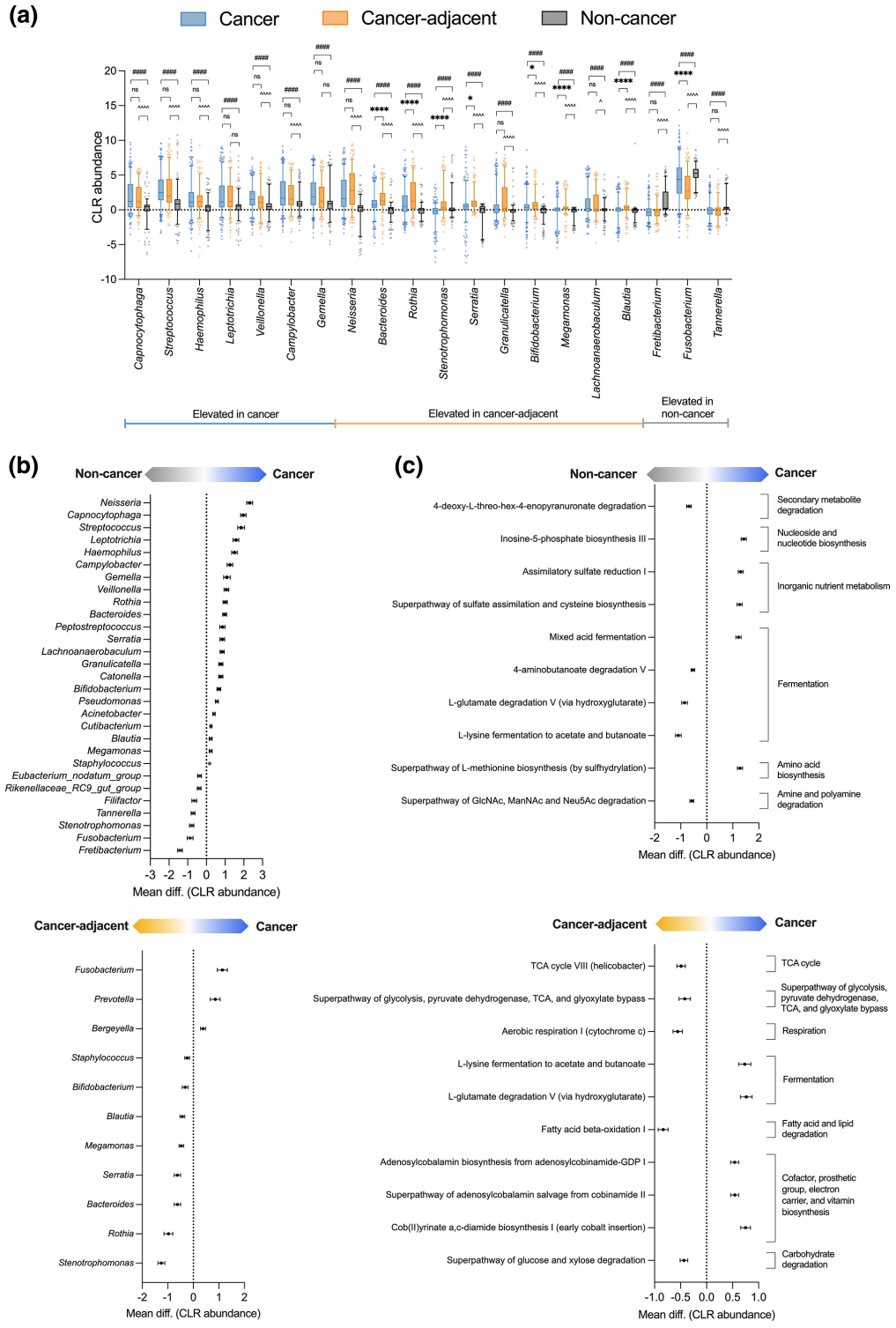

Fig. 3. Comparison of bacterial CLR abundance and functional prediction between sample types. (a) Top 20 bacterial genera (based on effect size) in CLR-normalized abundances between sample groups using the Kruskal–Wallis test with Bonferroni's multiple comparison test. Thirty-four out of 179 genera were identified as significantly different (P_{adjust}<0.05) using the Kruskal–Wallis test. The post-hoc Wilcoxon test with Bonferroni–Dunn's multiple comparison test was performed to identify groupwise differences between cancer–non-cancer (#), cancer–cancer-adjacent (*) and non-cancer–cancer-adjacent (^). Post-hoc unpaired Wilcoxon test with Bonferroni–Dunn's multiple comparison test for (b) bacterial genera and (c) functional CLR abundance for cancer–non-cancer (top panel) and cancer–cancer-adjacent (bottom panel). *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

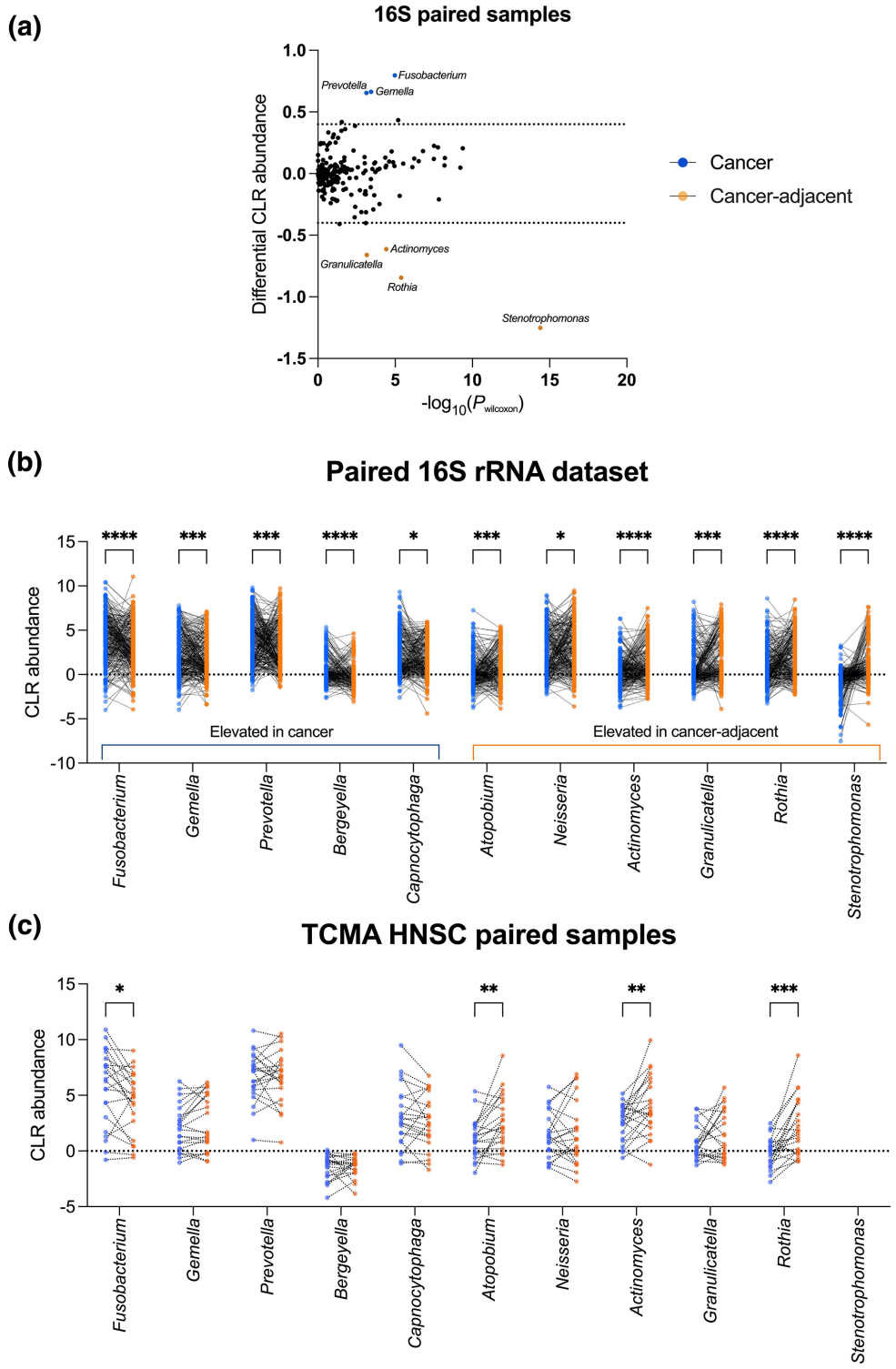

Fig. 4. Comparison of tissue microbiota in paired cancer and cancer-adjacent tissue samples using different sequencing datasets. (a) Paired Wilcoxon matched pairs signed rank test on paired 16S rRNA sequencing cancer and cancer-adjacent tissue samples. Seventy-six bacteria were significantly different in sample groups ($P<0.05$) using paired Wilcoxon matched pairs signed rank test and 13 bacterial 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. *$P<0.05$, **$P<0.01$, ***$P<0.001$, ****$P<0.0001$.

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*, while *Rothia, Atopobium* and *Actinomyces* were elevated (P<0.05) in cancer-adjacent tissues (Fig. 4d). In the TCMA dataset, *Stenotrophomonas* was not present due to pre-analysis filtering, while no significant differences in CLR abundance were observed for *Gemella*, *Prevotella*, *Bergeyella*, *Capnocytophaga*, *Neisseria* and *Granulicatella* (Fig. 4c). Notably, a similar trend in CLR abundance between cancer and cancer-adjacent samples was still observed for *Prevotella*, *Capnocytophaga*, *Neisseria* and *Granulicatella* in the TCMA dataset. Overall, the 16S rRNA and TCMA WGS datasets showed similar trends for most bacterial genera, regardless of sequencing techniques.

Tissue microbiota diversity correlates with cancer functional gene expression signatures

Since *Fusobacterium*, *Atopobium*, *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) [56]. Here, TCGA transcriptomic data were classified into 29 FGESs, which represent major functional components and characteristics of cancer cell populations [56]. These 29 FGESs can then be used to further classify cancers into four major immune subtypes (desert, fibrotic, immune-enriched/non-fibrotic and immune-enriched/fibrotic) [56]. We correlated the CLR abundance of *Fusobacterium*, *Atopobium*, *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 (Fig. 5a). Other FGESs such as matrix remodelling, protumour cytokines, MDSC traffic, M1 signature, antitumour cytokine, major histocompatibility complex class I (MHCI) and epithelial-mesenchymal transition (EMT) signatures were also positively correlated (P<0.05) with CLR abundance of *Fusobacterium* (Fig. 5a). The CLR abundance of *Atopobium* showed a negative correlation (P<0.05) with endothelium and matrix signatures (Fig. 4a). Lastly, the CLR abundance of *Rothia* displayed a positive correlation (P<0.05) with T reg traffic, T reg, B cell, NK cell, effector cell and T cell signatures, while having a negative correlation with (P<0.05) angiogenesis signature (Fig. 5a).

Next, we investigated how CLR abundance was correlated with tissue immune subtyping. Cancer tissues classified as immune deserts (D) and fibrotic (F), which lack immune cell enrichment, were correlated with higher *Fusobacterium* CLR abundance. On the other hand, cancer tissues that are immune-enriched/non-fibrotic (IE) or immune-enriched/fibrotic (IE/F) were correlated with greater *Rothia*. No significant correlation with immune subtypes was observed for *Actinomyces*, while high CLR abundance of *Atopobium* was associated with the D phenotype (Fig. 5b). To identify the differences in immune subtypes between high and low CLR abundances of each bacterial genera, we further segregated patients based on the upper and lower 35% CLR abundance quartiles. While not reaching statistical significance, more patients with IE and IE/F tumour subtypes have low CLR abundance of *Fusobacterium* (chi-squared test, P=0.19), opposite to the CLR abundance for *Rothia* (Fig. 5c). Conversely, patients with D and F subtypes had higher CLR abundance of *Fusobacterium* (Fig. 5c). Similar to *Rothia*, (chi-squared test, P=0.2), low CLR abundance of *Atopobium* is associated with reduced IE/F subtype (Fig. 5c). Lastly, the proportion of patients in each immune subtype was similar in high and low CLR abundance *Actinomyces* groups. Taken together, these show that *Fusobacterium* is associated with poor T-cell infiltration compared to *Rothia*, which may have implications for selecting patients suitable for immunotherapy.

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. The 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).

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 12 studies and presenting a consensus tissue microbiota signature for head and neck tissues. We analysed 16S rRNA sequencing datasets from 938 tissue samples, including 414 cancer tissues, 269 cancer-adjacent tissues and 255 non-cancer tissues. Our analysis revealed significant differences in the abundance of 34 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

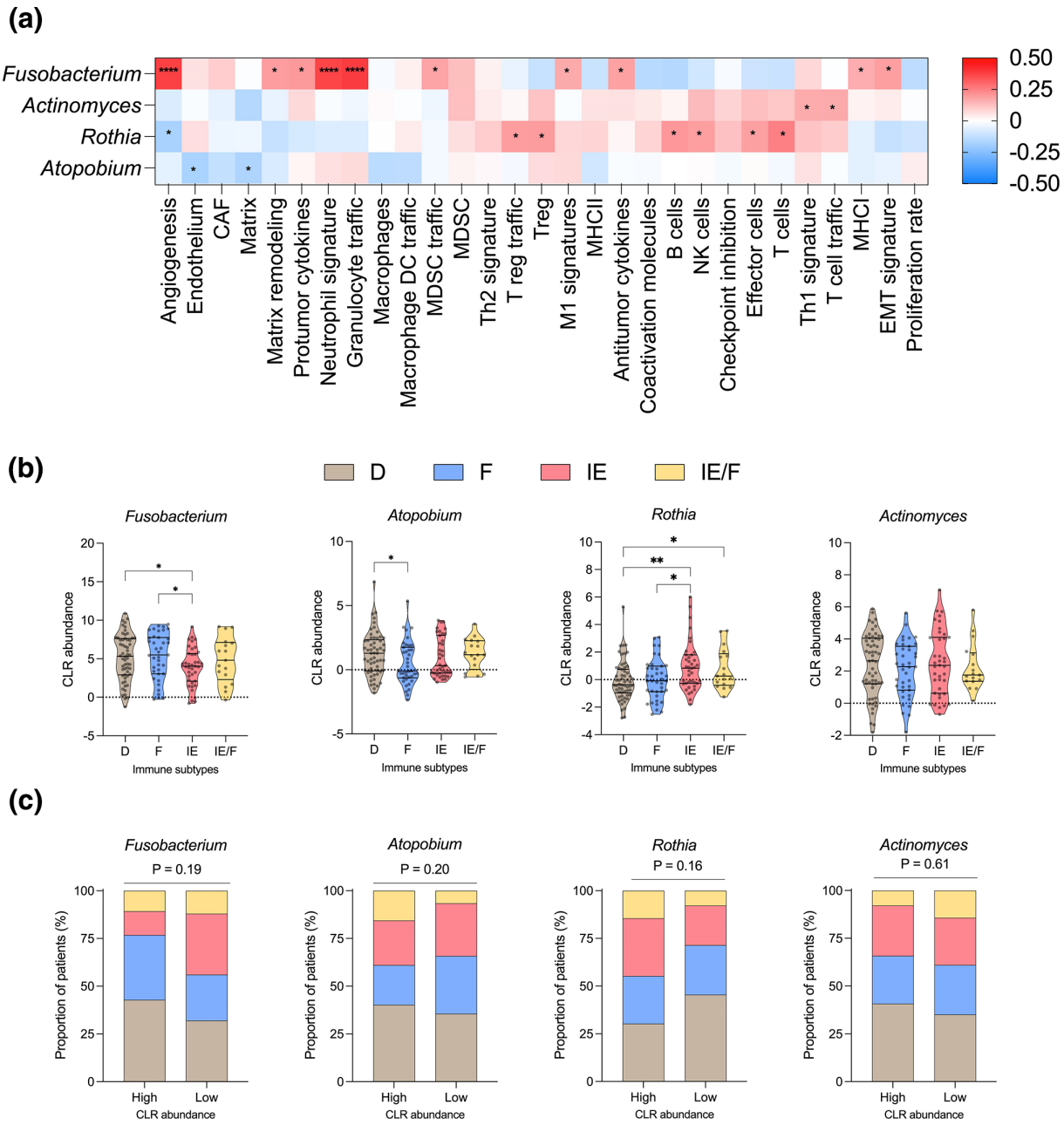

Fig. 5. Correlation analysis of *Fusobacterium*, *Atopobium*, *Rothia* and *Actinomyces* to the tumour transcriptional profiles. (a) Twenty-nine FGES scores derived from Bagaev *et al.* [56] were used for correlation with the CLR abundance of genera *Fusobacterium*, *Atopobium*, *Actinomyces* and *Rothia*, using Pearson's correlation method. Asterisks (*) represent significant correlation ($P<0.05$), and red and blue scales represent positive and negative correlations, respectively. (b) The CLR abundance of each bacterial genera within each TME 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 abundances in each bacterial genera. High and low bacterial CLR abundance groups were determined by the upper and lower 35% quartiles, respectively. Chi-squared test was performed to determine he association between high/low bacterial genera CLR abundance and the proportion of patients in each tumour subtype.

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* (Fig. 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 (Fig. 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

Table 2. Univariate and multivariate 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*
Atopobium	Continuous	153	1.0359 (0.8889–1.208)	0.651		
	High	53				
	Low	53	1.2781 (0.7385–2.212)	0.3805		
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		

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.

Both multivariate and univariate discriminant analyses were able to differentiate different tissue sample types based on microbial abundance. As previously reported, cancer and cancer-adjacent tissues were more similar in microbial diversity when compared to non-cancer tissues [31–33, 35, 36, 59, 60, 64, 66–71]. At the genus level, both paired and unpaired abundance analyses of cancer and cancer-adjacent tissues showed consistent enrichment for *Fusobacterium* and *Rothia* in cancer tissues [31, 33, 59, 66–69]. In contrast to previous studies, *Prevotella* and *Atopobium* were enriched within cancer tissues compared to cancer-adjacent tissues, and no differences were observed for *Streptococcus* [31–33, 35, 36, 59, 61, 64–70, 72].

Previous studies have reported conflicting results where *Fusobacterium* was increased in cancer tissues as compared to non-cancer and cancer-adjacent tissues [32, 60, 64, 66, 67]. 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 and tongue) and saliva of healthy individuals, suggesting a potential role within the healthy oral microbiota [73, 74]. *In vitro* experiments in HNSC cell lines showed that *Fusobacterium nucleatum* infection promotes cancer cell invasion, proliferation, autophagy and PD-L1 expression [75–77]. It is unknown whether there are strain- and species-level differences found in *Fusobacterium* isolated

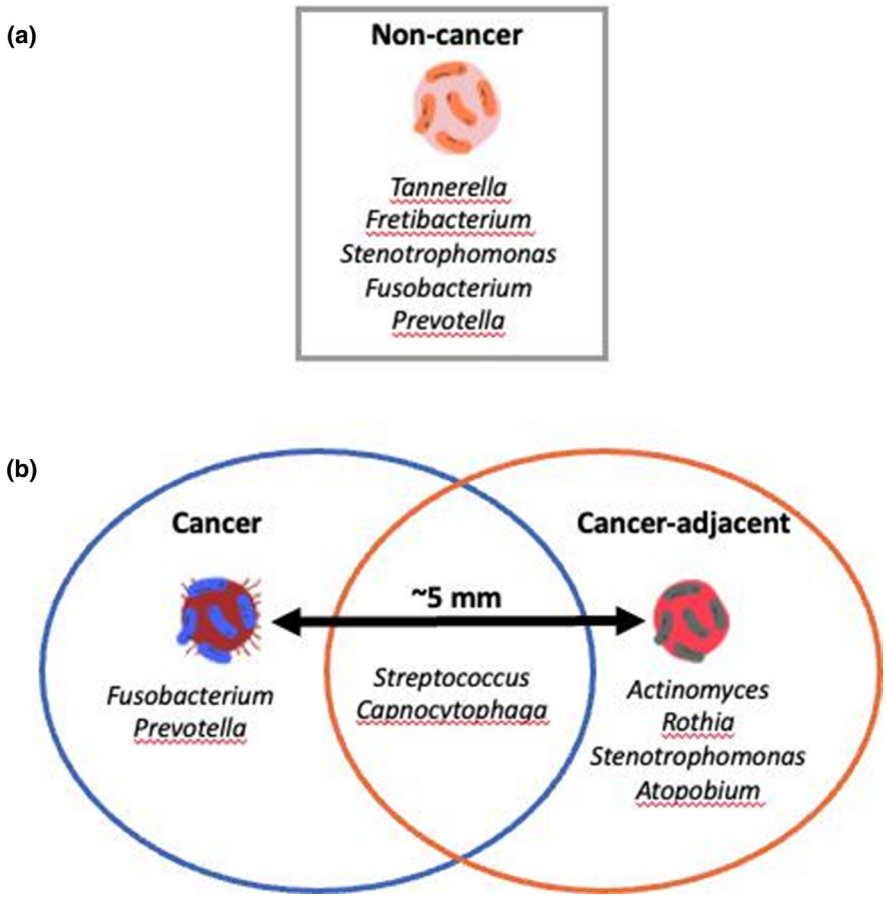

Fig. 6. Summary of bacterial genera within cancer, cancer-adjacent and non-cancer tissue samples. (a) Elevated microbiota within non-cancer tissues compared to cancer and cancer-adjacent tissues. (b) Elevated bacterial genera between cancer and cancer-adjacent tissues.

in cancer and non-cancer tissues to explain such seemingly contradictory findings. Additionally, non-cancer tissues from cancer patients may also have different tissue microbiota profiles from healthy donor tissues, for which data are currently unavailable in this study. Furthermore, 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 [73, 74], was increased specifically in cancer and cancer-adjacent tissues 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 *Streptococcus* species, such as *S. mutans*, can contribute to dental caries by acidifying the environment [78]. In oral cancer, *S. mutans* has been shown to promote tumour proliferation and invasion, potentially through upregulation of IL-6 in infected cells [79]. On the other hand, *Streptococcus* species from the *mitis* (*S. oralis*, *S parasanguinis* and *S.mitis*) and *sanguinis* (*S. sanguinis* and *S. gordonii*) groups can break down lactic acid or pyruvate into hydrogen peroxide, thereby antagonizing pathogenic species such as *S. mutans* [78]. In oral cancer, *S. mitis*, *S. salivarius* and *S. anginosus* were found to display anti-tumour effects, including reducing cancer cell viability and promoting $CD8^+$ cytotoxic T cell responses [80–83]. These findings indicate that the abundance of specific *Streptococcus* species may contribute to pathogenesis and disease severity or exert anti-tumour effects. It is important to note that these studies underscore the limitations of identifying microbiota at the genus level using short-read 16S rRNA sequencing. To address these limitations, recent advances in sequencing technologies such as long-read 16S rRNA amplicon sequencing (e.g. PacBio and Nanopore) or shotgun metagenomics can be employed to reveal species- or strain-specific diversity within the microbiota [84, 85]. Such advancements can provide a more comprehensive understanding of the specific species and strains that play a role in oral cancer pathogenesis and anti-tumour effects.

To compare the metabolic potential of different head and neck tissue types, a functional prediction analysis was performed using PICRUSt2 on the 16S rRNA sequencing data. Functional analysis predicted an enrichment of pathways that involve the production of several amino acids and metabolites, including L-methionine, acetate, butanoate and lactate, in cancer tissues compared to non-cancer and cancer-adjacent tissues. Methionine deprivation has been shown to sensitize numerous cancers such as breast and colorectal cancer cells [86, 87]. Moreover, competition between cancer and T cells for methionine can result in T cell impairment [87–89]. Butanoate and acetate can serve as energy sources for cells by converting into acetyl-CoA, which can then be utilized in the tricarboxylic acid (TCA) cycle to produce ATP [90, 91]. The role of butanoate in tumorigenesis depends on the specific tumour and the TME as it can exhibit tumour-promoting or suppressive properties [22, 92, 93]. Collectively, these findings suggest that bacteria infiltrating HNSC tissues possess functional capacities that may promote cancer progression. Further validation studies are warranted to better understand the role of these metabolic pathways in HNSC and the contribution of bacteria in shaping the TME.

We further explored the relationship between the abundance of the four bacterial genera associated with cancer and cancer-adjacent tissues, *Fusobacterium*, *Atopobium*, *Actinomyces* and *Rothia*, and the TME phenotype and clinical outcomes. *Fusobacterium* was associated with a lack of T-cell immune infiltration in HNSC, similar to colorectal and oesophageal cancers [94, 95]. Furthermore, *Fusobacterium* can chemoattract neutrophils via release of SCFA and can also modulate neutrophils and endothelial cell functions *in vitro* [96–98]. Interestingly, we observed that patients with low levels of *Fusobacterium* within the tumour tissue had shorter overall survival, consistent with previous reports for HNSC [59, 99, 100]. In contrast, opposite findings have been reported for colorectal, gastric and oesophageal cancers, suggesting that *Fusobacterium* may have a different role in HNSC [101–104]. In our analysis, we also found that *Rothia* was correlated with an immune-enriched TME. Limited information is available regarding the role of *Rothia* in cancer; however, *Rothia dentocariosa* has been shown to induce Toll-like receptor 2 (TLR-2)-mediated TNF-alpha inflammatory response in human embryonic kidney cells and THP-1 monocytes [105]. *Atopobium,* a oral and vagina commensal, was also found to be higher in tissues with an T-cell immune-deprived phenotype. While not studied extensively, specific *Atopobium* species such as *Atopobium minutum* cause apoptosis in colon cancer cell lines, while *Atopobium parvula* induces a pro-inflammatory response in cervical cells [106, 107]. Lastly, *Actinomyces* did not significantly correlate with TME subtypes in our analysis. However, *Actinomyces* has been associated with young-onset colorectal cancers, showing a preferential localization with cancer-associated fibroblasts in the TME [108]. These findings underscore the importance of validating and understanding the underlying mechanisms through which these bacteria can modulate the TME in HNSC.

This study represents the first comprehensive comparison of 16S rRNA (V3–V5) microbial sequencing across multiple studies to identify consensus HNSC-associated microbiota signatures in cancer, cancer-adjacent and non-cancer tissues. To ensure consistency, a uniform bioinformatics approach was employed. However, it is important to acknowledge the inherent limitations of this study. Variations in sample collection, preparation and sequencing among different laboratories introduce batch effects that could contribute to the inconsistencies observed across different reports. To mitigate these effects, we applied PLSDA batch adjustment to the pooled datasets [45]. Another limitation of this study is the lack of controls to eliminate background or contaminants sequenced as most studies do not have negative controls or sequencing files uploaded to the SRA database. Despite its flaws, we removed low-abundance bacteria to try to ameliorate this issue. Additionally, cohort-specific taxa that are most likely to reflect background or contaminants will likely be accounted for in the batch correction steps of our analysis.

Conventional short-read 16S rRNA sequencing only provides information up to the genus level, which restricts our ability to identify specific bacterial species or strains that may be relevant to disease outcomes [84]. Overcoming this limitation would require advanced sequencing technologies such as long-read 16S rRNA amplicon sequencing or shotgun metagenomics to reveal species- or strain-level diversity within the microbiota. Additionally, 16S rRNA sequencing can only identify the presence of bacterial DNA and does not identify the presence of viable bacteria within the tumour. Hence, other methods, such as transmission electron microscopy, culturomics, single-cell RNA sequencing or spatial transcriptomics, may better show viable and functional bacteria within tissue samples [2, 109]. Another important consideration is that the clinical metadata associated with the published 16S rRNA sequencing datasets are limited, restricting our ability to make comprehensive clinical associations. Thus, our clinical associations were primarily based on TCMA/TCGA datasets. Since TCMA/TCGA datasets are cross-sectional, a future longitudinal study would be more informative to understand the relationship between the microbiome, patient prognosis and treatment outcomes.

Despite these limitations, this study confirms 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, totalling 938 samples. 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. 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.

Funding information
The authors received no specific grant from any funding agency.

Acknowledgements
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. The illustration in Fig. 6 was generated using Canva.

Author contributions
Conceptualization, K.Y. and 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.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Dohlman AB, Arguijo Mendoza D, Ding S, Gao M, Dressman H, *et al*. The cancer microbiome atlas: a pan-cancer comparative analysis to distinguish tissue-resident microbiota from contaminants. *Cell Host Microbe* 2021;29:281–298.
2. Nejman D, Livyatan I, Fuks G, Gavert N, Zwang Y, *et al*. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. *Science* 2020;368:973–980.
3. Xavier JB, Young VB, Skufca J, Ginty F, Testerman T, *et al*. The cancer microbiome: distinguishing direct and indirect effects requires a systemic view. *Trends Cancer* 2020;6:192–204.
4. Yang L, Li A, Wang Y, Zhang Y. Intratumoral microbiota: roles in cancer initiation, development and therapeutic efficacy. *Signal Transduct Target Ther* 2023;8:35.
5. Vétizou M, Pitt JM, Daillère R, Lepage P, Waldschmitt N, *et al*. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* 2015;350:1079–1084.
6. Sivan A, Corrales L, Hubert N, Williams JB, Aquino-Michaels K, *et al*. Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science* 2015;350:1084–1089.
7. Fu A, Yao B, Dong T, Chen Y, Yao J, *et al*. Tumor-resident intracellular microbiota promotes metastatic colonization in breast cancer. *Cell* 2022;185:1356–1372.
8. Yu T, Guo F, Yu Y, Sun T, Ma D, *et al*. *Fusobacterium nucleatum* promotes chemoresistance to colorectal cancer by modulating autophagy. *Cell* 2017;170:548–563.
9. Spanogiannopoulos P, Kyaw TS, Guthrie BGH, Bradley PH, Lee JV, *et al*. Host and gut bacteria share metabolic pathways for anti-cancer drug metabolism. *Nat Microbiol* 2022;7:1605–1620.
10. Kalaora S, Nagler A, Nejman D, Alon M, Barbolin C, *et al*. Identification of bacteria-derived HLA-bound peptides in melanoma. *Nature* 2021;592:138–143.
11. Naghavian R, Faigle W, Oldrati P, Wang J, Toussaint NC, *et al*. Microbial peptides activate tumour-infiltrating lymphocytes in glioblastoma. *Nature* 2023;617:807–817.
12. Abed J, Emgård JEM, Zamir G, Faroja M, Almogy G, *et al*. Fap2 mediates *Fusobacterium nucleatum* colorectal adenocarcinoma enrichment by binding to tumor-expressed gal-GalNAc. *Cell Host Microbe* 2016;20:215–225.
13. Parhi L, Alon-Maimon T, Sol A, Nejman D, Shhadeh A, *et al*. Breast cancer colonization by *Fusobacterium nucleatum* accelerates tumor growth and metastatic progression. *Nat Commun* 2020;11:3259.
14. Gur C, Maalouf N, Shhadeh A, Berhani O, Singer BB, *et al*. *Fusobacterium nucleatum* supresses anti-tumor immunity by activating CEACAM1. *Oncoimmunology* 2019;8:e1581531.
15. Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, *et al*. Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity* 2015;42:344–355.
16. Yoon Y, Kim G, Jeon BN, Fang S, Park H. *Bifidobacterium* strain-specific rnhances the efficacy of cancer therapeutics in tumor-bearing mice. *Cancers* 2021;13:957.
17. Asadollahi P, Ghanavati R, Rohani M, Razavi S, Esghaei M, *et al*. Anti-cancer effects of *Bifidobacterium* species in colon cancer cells and a mouse model of carcinogenesis. *PLoS One* 2020;15:e0232930.
18. Rossi T, Vergara D, Fanini F, Maffia M, Bravaccini S, *et al*. Microbiota-derived metabolites in tumor progression and metastasis. *Int J Mol Sci* 2020;21:16.
19. Krautkramer KA, Fan J, Bäckhed F. Gut microbial metabolites as multi-kingdom intermediates. *Nat Rev Microbiol* 2021;19:77–94.
20. Bachem A, Makhlouf C, Binger KJ, de Souza DP, Tull D, *et al*. Microbiota-derived short-chain fatty acids promote the memory potential of antigen-activated $CD8^+$ T cells. *Immunity* 2019;51:285–297.
21. 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* 2023;84:532–541.
22. Okumura S, Konishi Y, Narukawa M, Sugiura Y, Yoshimoto S, *et al*. Gut bacteria identified in colorectal cancer patients promote tumourigenesis via butyrate secretion. *Nat Commun* 2021;12:5674.
23. Zaiatz-Bittencourt V, Jones F, Tosetto M, Scaife C, Cagney G, *et al*. Butyrate limits human natural killer cell effector function. *Sci Rep* 2023;13:2715.
24. Bender MJ, McPherson AC, Phelps CM, Pandey SP, Laughlin CR, *et al*. Dietary tryptophan metabolite released by intratumoral *Lactobacillus reuteri* facilitates immune checkpoint inhibitor treatment. *Cell* 2023;186:1846–1862.
25. Tintelnot J, Xu Y, Lesker TR, Schönlein M, Konczalla L, *et al*. Microbiota-derived 3-IAA influences chemotherapy efficacy in pancreatic cancer. *Nature* 2023;615:168–174.
26. Orlandi E, Iacovelli NA, Tombolini V, Rancati T, Polimeni A, *et al*. Potential role of microbiome in oncogenesis, outcome prediction and therapeutic targeting for head and neck cancer. *Oral Oncol* 2019;99:104453.

27. Burcher KM, Burcher JT, Inscore L, Bloomer CH, Furdui CM, *et al*. A review of the role of oral microbiome in the development, detection, and management of head and neck squamous cell cancers. *Cancers* 2022;14:17.
28. Orlandi E, Iacovelli NA, Tombolini V, Rancati T, Polimeni A, *et al*. Potential role of microbiome in oncogenesis, outcome prediction and therapeutic targeting for head and neck cancer. *Oral Oncol* 2019;99:104453.
29. Jain V, Baraniya D, El-Hadedy DE, Chen T, Slifker M, *et al*. Integrative metatranscriptomic analysis reveals disease-specific microbiome-host interactions in oral squamous cell carcinoma. *Cancer Res Commun* 2023;3:807–820.
30. Qiao H, Li H, Wen X, Tan X, Yang C, *et al*. Multi-omics integration reveals the crucial role of *Fusobacterium* in the inflammatory immune microenvironment in head and neck squamous cell carcinoma. *Microbiol Spectr* 2022;10:e0106822.
31. Zhou J, Wang L, Yuan R, Yu X, Chen Z, *et al*. Signatures of mucosal microbiome in oral squamous cell carcinoma identified using a random forest model. *Cancer Manag Res* 2020;12:5353–5363.
32. Torralba MG, Aleti G, Li W, Moncera KJ, Lin Y-H, *et al*. Oral microbial species and virulence factors associated with oral squamous cell carcinoma. *Microb Ecol* 2021;82:1030–1046.
33. Chang C, Geng F, Shi X, Li Y, Zhang X, *et al*. The prevalence rate of periodontal pathogens and its association with oral squamous cell carcinoma. *Appl Microbiol Biotechnol* 2019;103:1393–1404.
34. Al-Hebshi NN, Nasher AT, Maryoud MY, Homeida HE, Chen T, *et al*. Inflammatory bacteriome featuring *Fusobacterium nucleatum* and *Pseudomonas aeruginosa* identified in association with oral squamous cell carcinoma. *Sci Rep* 2017;7:1834.
35. Sarkar P, Malik S, Laha S, Das S, Bunk S, *et al*. Dysbiosis of oral microbiota during oral squamous cell carcinoma development. *Front Oncol* 2021;11:614448.
36. Gopinath D, Menon RK, Wie CC, Banerjee M, Panda S, *et al*. Differences in the bacteriome of swab, saliva, and tissue biopsies in oral cancer. *Sci Rep* 2021;11:1181.
37. Yu X, Shi Y, Yuan R, Chen Z, Dong Q, *et al*. Microbial dysbiosis in oral squamous cell carcinoma: a systematic review and meta-analysis. *Heliyon* 2023;9:e13198.
38. Gopinath D, Menon RK, Banerjee M, Su Yuxiong R, Botelho MG, *et al*. Culture-independent studies on bacterial dysbiosis in oral and oropharyngeal squamous cell carcinoma: a systematic review. *Crit Rev Oncol Hematol* 2019;139:31–40.
39. Peter TK, Withanage MHH, Comnick CL, Pendleton C, Dabdoub S, *et al*. Systematic review and meta-analysis of oral squamous cell carcinoma associated oral microbiome. *Front Microbiol* 2022;13:968304.
40. Ting HSL, Chen Z, Chan JYK. Systematic review on oral microbial dysbiosis and its clinical associations with head and neck squamous cell carcinoma. *Head & Neck* 2023;45:2120–2135.
41. Metsäniitty M, Hasnat S, Salo T, Salem A. Oral microbiota-a new frontier in the pathogenesis and management of head and neck cancers. *Cancers* 2021;14:46.
42. Su Mun L, Wye Lum S, Kong Yuiin Sze G, Hock Yoong C, Ching Yung K, *et al*. Association of microbiome with oral squamous cell carcinoma: a systematic review of the metagenomic studies. *Int J Environ Res Public Health* 2021;18:14.
43. Ramos RT, Sodré CS, de Sousa Rodrigues PMGR, da Silva AMP, Fuly MS, *et al*. High-throughput nucleotide sequencing for bacteriome studies in oral squamous cell carcinoma: a systematic review. *Oral Maxillofac Surg* 2020;24:387–401.
44. Delaney C, Veena CLR, Butcher MC, McLean W, Shaban SMA, *et al*. Limitations of using 16S rRNA microbiome sequencing to predict oral squamous cell carcinoma. *APMIS* 2023;131:262–276.
45. Wang Y, Lê Cao K-A. PLSDA-batch: a multivariate framework to correct for batch effects in microbiome data. *Brief Bioinform* 2023;24:bbac622.
46. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, *et al*. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019:l4898.
47. Choudhary S. pysradb: a python package to query next-generation sequencing metadata and data from NCBI Sequence Read Archive. *F1000Res* 2019;8:532.
48. Leinonen R, Sugawara H, Shumway M. The sequence read archive. *Nucleic Acids Res* 2011;39:D19–21.
49. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, *et al*. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852–857.
50. Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: an R package for omics feature selection and multiple data integration. *PLoS Comput Biol* 2017;13:e1005752.
51. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;8:e61217.
52. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome datasets are compositional: and this is not optional. *Front Microbiol* 2017;8:2224.
53. R Core Team R. R: A language and environment for statistical computing; 2013
54. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, *et al*. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* 2020;38:685–688.
55. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, *et al*. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013;6:pl1.
56. Bagaev A, Kotlov N, Nomie K, Svekolkin V, Gafurov A, *et al*. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* 2021;39:845–865.
57. Zhou X, Hao Y, Peng X, Li B, Han Q, *et al*. The clinical potential of oral microbiota as a screening tool for oral squamous cell carcinomas. *Front Cell Infect Microbiol* 2021;11:728933.
58. Wang X, Zhao Z, Tang N, Zhao Y, Xu J, *et al*. Microbial community analysis of saliva and biopsies in patients with oral lichen planus. *Front Microbiol* 2020;11:629.
59. Chen Z, Wong PY, Ng CWK, Lan L, Fung S, *et al*. The intersection between oral microbiota, host gene methylation and patient outcomes in head and neck squamous cell carcinoma. *Cancers* 2020;12:3425.
60. 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* 2021;43:3324–3330.
61. De Martin A, Lütge M, Stanossek Y, Engetschwiler C, Cupovic J, *et al*. Distinct microbial communities colonize tonsillar squamous cell carcinoma. *Oncoimmunology* 2021;10:1945202.
62. Zhang Z, Feng Q, Li M, Li Z, Xu Q, *et al*. Age-related cancer-associated microbiota potentially promotes oral squamous cell cancer tumorigenesis by distinct mechanisms. *Front Microbiol* 2022;13.
63. Nie F, Wang L, Huang Y, Yang P, Gong P, *et al*. Characteristics of microbial distribution in different oral niches of oral squamous cell carcinoma. *Front Cell Infect Microbiol* 2022;12:905653.
64. Gong H, Shi Y, Xiao X, Cao P, Wu C, *et al*. Alterations of microbiota structure in the larynx relevant to laryngeal carcinoma. *Sci Rep* 2017;7:5507.
65. Perera M, Al-Hebshi NN, Perera I, Ipe D, Ulett GC, *et al*. Inflammatory bacteriome and oral squamous cell carcinoma. *J Dent Res* 2018;97:725–732.
66. Gong H-L, Shi Y, Zhou L, Wu C-P, Cao P-Y, *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:e66476.
67. Dong Z, Zhang C, Zhao Q, Huangfu H, Xue X, *et al*. Alterations of bacterial communities of vocal cord mucous membrane

increases the risk for glottic laryngeal squamous cell carcinoma. *J Cancer* 2021;12:4049–4063.

68. Shin JM, Luo T, Kamarajan P, Fenno JC, Rickard AH, *et al*. Microbial communities associated with primary and metastatic head and neck squamous cell carcinoma - a high fusobacterial and low streptococcal signature. *Sci Rep* 2017;7:9934.
69. Yang K, Wang Y, Zhang S, Zhang D, Hu L, *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.
70. Gong H, Shi Y, Zhou X, Wu C, Cao P, *et al*. Microbiota in the throat and risk factors for laryngeal carcinoma. *Appl Environ Microbiol* 2014;80:7356–7363.
71. Wang H, Funchain P, Bebek G, Altemus J, Zhang H, *et al*. Microbiomic differences in tumor and paired-normal tissue in head and neck squamous cell carcinomas. *Genome Med* 2017;9:14.
72. Pushalkar S, Ji X, Li Y, Estilo C, Yegnanarayana R, *et al*. Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. *BMC Microbiol* 2012;12:144.
73. Eren AM, Borisy GG, Huse SM, Mark Welch JL. Oligotyping analysis of the human oral microbiome. *Proc Natl Acad Sci U S A* 2014;111:E2875–84.
74. 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* 2016;113:E791–800.
75. Shao W, Fujiwara N, Mouri Y, Kisoda S, Yoshida K, *et al*. Conversion from epithelial to partial-EMT phenotype by *Fusobacterium nucleatum* infection promotes invasion of oral cancer cells. *Sci Rep* 2021;11:14943.
76. Chen G, Gao C, Jiang S, Cai Q, Li R, *et al*. *Fusobacterium nucleatum* outer membrane vesicles activate autophagy to promote oral cancer metastasis. *J Adv Res* 2023:S2090-1232(23)00094-2.
77. 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* 2020;39:144–151.
78. Baty JJ, Stoner SN, Scoffield JA. Oral commensal streptococci: gatekeepers of the oral cavity. *J Bacteriol* 2022;204:e0025722.
79. Tsai MS, Chen YY, Chen WC, Chen MF. *Streptococcus mutans* promotes tumor progression in oral squamous cell carcinoma. *J Cancer* 2022;13:3358–3367.
80. Baraniya D, Jain V, Lucarelli R, Tam V, Vanderveer L, *et al*. Screening of health-associated oral bacteria for anticancer properties *in vitro Front Cell Infect Microbiol* 2020;10:575656.
81. 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* 2021;21:141.
82. 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;14:2073866.
83. Wang J, Sun F, Lin X, Li Z, Mao X, *et al*. Cytotoxic T cell responses to Streptococcus are associated with improved prognosis of oral squamous cell carcinoma. *Exp Cell Res* 2018;362:203–208.
84. Curry KD, Wang Q, Nute MG, Tyshaieva A, Reeves E, *et al*. Emu: species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. *Nat Methods* 2022;19:845–853.
85. Johnson JS, Spakowicz DJ, Hong B-Y, Petersen LM, Demkowicz P, *et al*. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* 2019;10:5029.
86. Gao X, Sanderson SM, Dai Z, Reid MA, Cooper DE, *et al*. Dietary methionine influences therapy in mouse cancer models and alters human metabolism. *Nature* 2019;572:397–401.
87. Sedillo JC, Cryns VL. Targeting the methionine addiction of cancer. *Am J Cancer Res* 2022;12:2249–2276.
88. Pandit M, Kil Y-S, Ahn J-H, Pokhrel RH, Gu Y, *et al*. Methionine consumption by cancer cells drives a progressive upregulation of PD-1 expression in CD4 T cells. *Nat Commun* 2023;14:2593.
89. Bian Y, Li W, Kremer DM, Sajjakulnukit P, Li S, *et al*. Cancer SLC43A2 alters T cell methionine metabolism and histone methylation. *Nature* 2020;585:277–282.
90. Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, *et al*. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* 2011;13:517–526.
91. Comerford SA, Huang Z, Du X, Wang Y, Cai L, *et al*. Acetate dependence of tumors. *Cell* 2014;159:1591–1602.
92. 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* 2016;165:1332–1345.
93. van der Hee B, Wells JM. Microbial regulation of host physiology by short-chain fatty acids. *Trends Microbiol* 2021;29:700–712.
94. Mima K, Sukawa Y, Nishihara R, Qian ZR, Yamauchi M, *et al*. *Fusobacterium nucleatum* and T cells in colorectal carcinoma. *JAMA Oncol* 2015;1:653–661.
95. Kosumi K, Baba Y, Yamamura K, Nomoto D, Okadome K, *et al*. Intratumour *Fusobacterium nucleatum* and immune response to oesophageal cancer. *Br J Cancer* 2023;128:1155–1165.
96. Dahlstrand Rudin A, Khamzeh A, Venkatakrishnan V, Basic A, Christenson K, *et al*. Short chain fatty acids released by *Fusobacterium nucleatum* are neutrophil chemoattractants acting via free fatty acid receptor 2 (FFAR2). *Cell Microbiol* 2021;23:e13348.
97. Mendes RT, Nguyen D, Stephens D, Pamuk F, Fernandes D, *et al*. Endothelial cell response to *Fusobacterium nucleatum*. *Infect Immun* 2016;84:2141–2148.
98. Wright HJ, Chapple ILC, Matthews JB, Cooper PR. *Fusobacterium nucleatum* regulation of neutrophil transcription. *J Periodontal Res* 2011;46:1–12.
99. Neuzillet C, Marchais M, Vacher S, Hilmi M, Schnitzler A, *et al*. Prognostic value of intratumoral *Fusobacterium nucleatum* and association with immune-related gene expression in oral squamous cell carcinoma patients. *Sci Rep* 2021;11:7870.
100. Chan JYK, Cheung MK, Lan L, Ng C, Lau EHL, *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* 2022;135:106245.
101. Mima K, Nishihara R, Qian ZR, Cao Y, Sukawa Y, *et al*. *Fusobacterium nucleatum* in colorectal carcinoma tissue and patient prognosis. *Gut* 2016;65:1973–1980.
102. Lehr K, Nikitina D, Vilchez-Vargas R, Steponaitiene R, Thon C, *et al*. Microbial composition of tumorous and adjacent gastric tissue is associated with prognosis of gastric cancer. *Sci Rep* 2023;13:4640.
103. Hsieh YY, Kuo WL, Hsu WT, Tung SY, Li C. *Fusobacterium nucleatum*-induced tumor mutation burden predicts poor survival of gastric cancer patients. *Cancers* 2022;15:269.
104. Zhang N, Liu Y, Yang H, Liang M, Wang X, *et al*. Clinical significance of *Fusobacterium nucleatum* infection and regulatory T cell enrichment in esophageal squamous cell carcinoma. *Pathol Oncol Res* 2021;27:1609846.
105. Kataoka H, Taniguchi M, Fukamachi H, Arimoto T, Morisaki H, *et al*. Rothia dentocariosa induces TNF-alpha production in a TLR2-dependent manner. *Pathog Dis* 2014;71:65–68.
106. Maarsingh JD, Łaniewski P, Herbst-Kralovetz MM. Immunometabolic and potential tumor-promoting changes in 3D cervical cell models infected with bacterial vaginosis-associated bacteria. *Commun Biol* 2022;5:725.
107. Altonsy MO, Andrews SC, Tuohy KM. Differential induction of apoptosis in human colonic carcinoma cells (Caco-2) by Atopobium, and commensal, probiotic and enteropathogenic bacteria: mediation by the mitochondrial pathway. *Int J Food Microbiol* 2010;137:190–203.

108. Xu Z, Lv Z, Chen F, Zhang Y, Xu Z, *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.

109. Galeano Niño JL, Wu H, LaCourse KD, Kempchinsky AG, Baryiames A, *et al*. Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. *Nature* 2022;611:810–817.