Title

Oral fungal profiling and risk of nasopharyngeal carcinoma: a population-based case-control study.

Authors

Chen, Yufeng
Li, Wanxin
Chang, Ellen
et al.

Publication Date

2023-10-01

DOI

10.1016/j.ebiom.2023.104813

Peer reviewed

Oral fungal profiling and risk of nasopharyngeal carcinoma: a population-based case-control study

Yufeng Chen,[a,n] Wanxin Li,[b,n] Ellen T. Chang,[c] Justine W. Debelius,[d,e] Lokeshwaran Manoharan,[f] Yuming Zheng,[g,h] Yancheng Li,[g,i] Guangwu Huang,[j,k] Hans-Olov Adami,[a,l] Rob Knight,[m] Yonglin Cai,[g,h,o,**] Zhe Zhang,[j,k,o,***] and Weimin Ye[a,b,o,*]

[a]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, 17177, Sweden
[b]Department of Epidemiology and Health Statistics & Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Fujian Medical University, Fuzhou, 350000, China
[c]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, 94143, USA
[d]Center for Translational Microbiome Research, Department of Microbiology, Tumor and Cancer Biology, Karolinska Institutet, Stockholm, 17177, Sweden
[e]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, 21205, USA
[f]National Bioinformatics Infrastructure Sweden (NBIS), Lund University, Lund, 22100, Sweden
[g]Guangxi Health Commission Key Laboratory of Molecular Epidemiology of Nasopharyngeal Carcinoma, Wuzhou Red Cross Hospital, Wuzhou, 543002, China
[h]Department of Preventive Medicine, Wuzhou Cancer Center, Wuzhou, 543002, China
[i]Cangwu Institute for Nasopharyngeal Carcinoma Control and Prevention, Wuzhou, 543002, China
[j]Department of Otolaryngology-Head & Neck Surgery, First Affiliated Hospital of Guangxi Medical University, Nanning, 530021, China
[k]Key Laboratory of High-Incidence-Tumor Prevention & Treatment (Guangxi Medical University), Ministry of Education, Nanning, 530021, China
[l]Clinical Effectiveness Group, Institute of Health and Society, University of Oslo, Oslo, NO-0316, Norway
[m]Department of Pediatrics, University of California San Diego, CA, 92093, USA

eBioMedicine
2023;96: 104813
Published Online xxx
https://doi.org/10.1016/j.ebiom.2023.104813

Summary

Background Dysbiosis of the oral mycobiome has been linked to some diseases, including cancers. However, the role of oral fungal communities in nasopharyngeal carcinoma (NPC) carcinogenesis has not previously been investigated.

Methods We characterized the oral salivary fungal mycobiome in 476 untreated incident NPC patients and 537 population-based controls using fungal internal transcribed spacer (ITS)-2 sequencing. The relationship between oral fungal mycobiome and the risk of NPC was assessed through bioinformatic and biostatistical analyses.

Findings We found that lower fungal alpha diversity was associated with an increased odds of NPC [lower vs. higher: observed features (adjusted odds ratio [OR] = 5.81, 95% confidence interval [CI] = 3.60–9.38); Simpson diversity (1.53, 1.03–2.29); Shannon diversity (2.03, 1.35–3.04)]. We also observed a significant difference in global fungal community patterns between cases and controls based on Bray–Curtis dissimilarity ($P < 0.001$). Carriage of oral fungal species, specifically, *Saccharomyces cerevisiae*, *Candida tropicalis*, *Lodderomyces elongisporus*, *Candida albicans*, and *Fusarium poae*, was associated with significantly higher odds of NPC, with ORs ranging from 1.56 to 4.66. Individuals with both low fungal and low bacterial alpha diversity had a profoundly elevated risk of NPC.

Interpretation Our results suggest that dysbiosis in the oral mycobiome, characterized by a loss of fungal community diversity and overgrowth of several fungal organisms, is associated with a substantially increased risk of NPC.

Funding This work was funded by the US National Institutes of Health, the Swedish Research Council, the High-level Talents Research Start-up Project of Fujian Medical University, and the China Scholarship Council.

*Corresponding author. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, 171 77, Sweden.
**Corresponding author. Guangxi Health Commission Key Laboratory of Molecular Epidemiology of Nasopharyngeal Carcinoma, Wuzhou Red Cross Hospital, Wuzhou, 543002, China.
***Corresponding author. Department of Otolaryngology-Head & Neck Surgery, First Affiliated Hospital of Guangxi Medical University, Nanning, 530021, China.
E-mail addresses: weimin.ye@ki.se (W. Ye), cylzen@163.com (Y. Cai), zhangzhe1975@gmail.com (Z. Zhang).
[n]These authors contributed equally to this work.
[o]These authors contributed equally to this work.

Keywords: Fungi; ITS sequencing; Oral mycobiome; Microbiome; Nasopharyngeal carcinoma; Case–control study

Research in context

Evidence before this study

Dysbiosis of the oral mycobiome has been linked to head and neck and oral cancers. Prior evidence and our recent data, based on 16S rRNA gene amplicon sequencing, suggest that oral bacterial dysbiosis is involved in nasopharyngeal carcinoma (NPC) carcinogenesis. However, the role of the mycobiome, another core component of the oral microbiome, in NPC development has not yet been investigated in large-scale epidemiological studies.

Added value of this study

Advances in sequencing allow the evaluation of the role of the oral mycobiome in NPC development. We investigated the relationship between the oral fungal mycobiome and risk of NPC using fungal internal transcribed spacer (ITS)-2 sequencing in a population-based case–control study with 476 untreated incident NPC patients and 537 population controls. In addition, we analysed interactions between fungal components and bacterial components in relation to the risk of NPC.

Implications of all the available evidence

Our results suggest that dysbiosis in the oral mycobiome, characterized by a loss of fungal community diversity and overgrowth of several pathogenic or opportunistic pathogenic fungal organisms, is associated with a substantially increased risk of NPC. We observe strong additive interactions between the fungal and bacterial microbiomes on NPC risk. These findings, if confirmed by prospective studies, might offer insights into NPC etiology and pave a potential way for NPC prevention and control.

Introduction

The human oral cavity harbors our bodies' most diverse and complex microorganisms, which are referred to as the oral microbiome.[1] As in other body sites, bacterial and fungal communities are the primary members of the oral microbiome, and the bacterial component is predominant.[2,3] Although still markedly understudied compared with the oral bacteriome, the oral mycobiome has received increasing attention in recent years as high-throughput sequencing technologies are developing.[4] The complexity and biodiversity of the oral mycobiome and its role in health have started to be uncovered, and dysbiosis of the oral mycobiome has been linked to several diseases, including periodontitis, oral squamous cell carcinoma (OSCC), head and neck squamous cell carcinoma (HNSCC), and oral tongue cancer.[5–11]

Nasopharyngeal carcinoma (NPC) arises from the epithelium of the nasopharynx and is responsible for over 133,000 incident cases and 80,000 deaths annually worldwide as of 2020.[12] Although it is rare in most areas of the world, NPC is particularly common in some regions including southern China, Southeastern Asia, and northern and eastern Africa.[13] It is well accepted that genetic susceptibility, environmental factors, and infection with Epstein–Barr virus (EBV) are involved in the development of NPC.[13,14] EBV, as a known microbial factor, is the most established and extensively studied risk factor for NPC and plays a causal role in the etiology of probably all NPC in endemic areas.[15,16] However, apart from EBV, the role of other common types of microbial factors, including bacterial and fungal microbes, in NPC pathogenesis remains poorly understood.

Poor oral hygiene is associated with an increased risk of NPC,[17] suggesting that the oral microbiome may play a role in NPC development. Prior studies, including our recent analysis based on 16S rRNA gene amplicon sequencing, have shown significant differences in the oral bacterial microbiome between pre-treatment NPC patients and controls.[18,19] In addition, the oral bacterial microbiome has been linked to NPC prognosis; evidence from prospective cohort studies showed that lower alpha diversity was related to higher all-cause mortality[20] and high bacterial load was associated with poorer prognosis.[21] However, the existing microbiome studies of NPC exclusively focused on the bacterial component, while the role of the mycobiome, another core component of the oral microbiome, in NPC has not yet been investigated in large-scale epidemiological studies.

Hence, using fungal amplicon sequencing and bioinformatic/statistical analysis, we aimed to characterise oral fungal profiles in untreated incident NPC cases and population controls, and to evaluate the association between the oral mycobiome and NPC risk using a population-based case–control study in an NPC-endemic area of southern China.

Methods

Study population, participant data collection, and sample collection

Recruitment of study participants has been described in detail previously.[19,22] In brief, incident NPC cases were recruited between March 2010 and December 2013 through a rapid case ascertainment network at hospitals in Zhaoqing of Guangdong Province and Wuzhou, Guiping, and Pingnan of Guangxi Autonomous Region in southern China. Controls frequency-matched to cases on age and sex were randomly selected from the total population in each study area and recruited between

November 2010 and November 2014. All recruited participants were aged 20–74 years, residing in the study areas, and without a prior history of malignant diseases or congenital or acquired immunodeficiency.

A structured electronic questionnaire interview was administered to collect information covering demographics, residential and occupational history, history of chronic ear, nose, and throat (ENT) diseases, oral hygiene, family history of malignancies, cigarette smoking, alcohol and tea consumption, dietary habits, and use of Chinese herbal medicine. Saliva samples were collected at the time of interview from 92.0% of NPC cases and 94.6% of controls. Participants were required not to eat or chew gum for 30 min before sample collection. Two to four ml saliva from each subject was collected into a 50-ml falcon tube containing 2.5 ml lysis buffer (pH 8.0, 50 mM Tris, 50 mM EDTA, 50 mM sucrose, 100 mM NaCl, 10% SDS) and stored at –20 °C for less than 3 days before being long-term stored at –80 °C.

In this analysis, we included saliva samples collected from the Wuzhou site, as one subset of abovementioned large case–control study.[19,22] The Wuzhou region was selected for this analysis due to its relatively high incidence of NPC and low residential mobility. The four residential communities included in this study (Wuzhou, Cangwu, Tengxian, and Cenxi) covered 95% of the population in the Wuzhou region. In total 1081 samples (542 cases, 539 controls) were sent for sequencing. We excluded two samples that had ambiguous identifiers and could not be mapped to the participant data, 4 samples with fewer than 1000 reads after denoising, and 62 cases' samples that were collected after treatment, leaving 1013 saliva samples for the final analysis (476 cases, 537 controls, comprising 76% [476/627] of all pretreatment cases and 81% [537/665] of all controls enrolled from the Wuzhou study site) (Figure S1).

Ethics

The study was approved by the institutional review boards of Harvard School of Public Health (reference no. P16848-101), Sun Yat-sen University Cancer Center (reference no. YP2008075), the Institute for Viral Disease Control and Prevention of the Chinese Center for Disease Control and Prevention (reference no. CCDC-R01CA115873), First Affiliated Hospital of Guangxi Medical University (reference no. GXMU-R01CA115873), and the Stockholm ethical review board (reference no. 2009/1293-31/3). Written or oral informed consent was obtained from all study participants.

DNA extraction, PCR amplification, library preparation, and sequencing

We extracted DNA from 1 ml solution of saliva sample using a two-step protocol, including lysozyme lysis preprocessing (lysozyme from chicken egg white; Sigma–Aldrich) and mechanical bead beating, and followed by using the TIANamp blood DNA kit (TIANGEN Biotech Co., Ltd, Beijing, China) per the manufacturer's instructions. Amplicon libraries were constructed by targeting the fungal internal transcribed spacer (ITS)-2 region, using the following primers[23,24]: ITS3ngsmixes: CANCGATGAAGAACGYRG; and ITS4ngsUni: CCTSCSCTTANTDATATGC. We selected the primers based on previous published work by Tedersoo et al. illustrating that these primers are modified to reduce mismatches and recommended candidates for high-throughput sequencing analysis of the ITS-2 region in fungi.[23] Illumina's specific adapter sequences were linked to the ITS-2 primer pairs per Illumina's Fungal Metagenomic Sequencing Demonstrated Protocol.

First-stage PCR conditions and cycles were optimized as follows: we added 5 µl of salivary DNA to 12.5 µl KAPA HiFi HotStart ReadyMix (Kapa Biosystems, Salt River, Cape Town, South Africa), 2 µl (10 µM) of primers (1 µl forward and reverse each), and 5.5 µl PCR-grade water (total volume: 25 µl). Samples underwent initial denaturation at 95 °C for 5 min, followed by 30 cycles at 98 °C for 30 s, 57 °C for 30 s, and 72 °C for 30 s with a final extension step at 72 °C for 5 min. After amplification, the PCR products were visualized using a 1.5% agarose gel to ensure successful amplification and an appropriately sized band. PCR products were then purified using AMPure XP beads (Beckman Coulter Genomics) with a ratio of 1.8 for volume of AMPure XP beads to PCR products, and a final elution performed in 40 µl PCR-grade water.

Second-stage index PCR was performed to barcode the samples with indexes (IDT® Illumina Nextera DNA Unique Dual Indexes, Illumina, Inc.) and to add the flow-cell attachment regions. We performed a second cleanup with a beads ratio of 1.2:1. Final libraries were quantified using a Qubit fluorometer 4.0 double-stranded DNA (dsDNA) assay (Invitrogen, Carlsbad, CA, USA), and the purity of libraries was measured on an Agilent 2100 Bioanalyzer system. We pooled amplicon libraries at equimolar concentrations. Pooled libraries of 10 pM with 15% phiX were loaded and sequenced on the Illumina MiSeq platform using the MiSeq Reagent Kit (v.3, 600 cycles) following the 2 × 300-bp paired-end sequencing protocol. We divided the samples into five batches for sequencing, and the samples were randomly distributed in these batches. Extraction blanks and respective positive and negative controls for the various PCR steps were run in parallel in each batch. A single-organism i.e. *Candida albicans* (ordered from the Culture Collection University of Gothenburg, CCUG Number: 44135) was used as the positive control.

Sequence data processing

The Illumina sequencing quality report indicated that the quality of sequences was relatively high (overall run

metrics: ≥Q30 = 80.5%, clusters PF = 90.4%). We processed the Illumina-generated demultiplexed fungal ITS sequences using QIIME2[25] (v. 2020.8.0). Primers' and adaptors' sequences were trimmed using Cutadapt.[26] Sequences were quality-filtered and denoised using DADA2[27] with appropriate parameters (i.e., no bases were trimmed from the beginning of the sequences; trim lengths were 280 bases for forward and 210 bases for reverse sequences, respectively) based on the Interactive Quality Plot to filter out any phiX and chimeric sequences, and to construct an amplicon sequence variants (ASVs) table. After quality filtering, an extremely few reads were left for the blank and negative controls [median reads (interquartile range) = 60 (21–427)], and high consistency was achieved for the positive controls across batches (coefficient of variation = 0.95%) (Table S1). In total, there were 76,369,485 raw reads, and 32,665,463 high-quality reads were retained after quality-filtering procedures for downstream analysis. We performed taxonomic assignments with a naive Bayesian classifier trained against the latest UNITE reference database[28] (release for QIIME, v. 8.3). Low-prevalence ASVs present in less than 2 samples were removed from the ASV table (no samples were excluded in this step). Finally, we used the quality-filtered ASV and taxonomy tables for further analysis in R (version 4.1).[29] Participant data, ASV table, taxonomy table, and representative sequences were integrated into one object using the R package "Phyloseq" (v. 1.36).[30]

Diversity analyses

To avoid deviation caused by effects of different sequencing depths, we rarefied sequencing depth to 5000 reads per sample without replacement for diversity analysis; 3.9% (40/1013) samples with <5000 reads were removed, leaving samples of 450 cases and 523 controls. Alpha diversity refers to the variation of microbes within a sample or microbial ecosystem. We used observed ASVs, Simpson diversity, and Shannon diversity to indicate alpha diversity. Specifically, observed ASVs provide a measure of richness, i.e., the number of different sequence variants in the sample; Simpson diversity measures evenness, i.e., the relative abundance of different sequence variants in the sample; and Shannon diversity measures both richness and abundance. Each indicator of alpha diversity was expressed as a continuous variable or categorized into low, medium, or high diversity by tertiles of the distribution among controls.

Beta diversity describes the variation of microbial communities between samples or microbial ecosystems. To evaluate beta diversity, we computed Bray–Curtis dissimilarity, which considers relative abundance, using the R package "Vegan" (v. 2.5-7).[31,32] We performed permutational multivariate analysis of variance (PERMANOVA) (Adonis test in "Vegan") to calculate the variation explained by each of our collected host factors. In each Adonis test, we minimally adjusted for age, sex, and sequencing run, and *P* values were generated based on 9999 permutations. All *P* values were then adjusted for false discovery rate (FDR) using the Benjamini-Hochberg method.[33,34] We used principal coordinates analysis (PCoA) on Bray–Curtis dissimilarity to visualise beta diversity by disease status.

Differential abundance analysis

We performed a differential abundance analysis on unrarefied data using the R package Analysis of Compositions of Microbiomes with Bias Correction (ANCOMBC, v. 1.2.2)[35] to identify differentially abundant taxa between NPC cases and controls, adjusting for potential covariates that contribute to fungal composition. These covariates were selected based on a statistically significant influence on alpha and/or beta diversity (see Table S2 and Fig. 1b) and included age, sex, sequencing run, EBV infection status, residential community, educational level, current occupation, tooth loss, tooth brushing frequency, cigarette smoking, and tea drinking. For fungal species that were found to be differentially abundant by disease status, we classified the carriage of each species as no carriage (i.e., where relative abundance was equal to zero), low, or high carriage, with the cutoff between low and high based on the median abundance among controls.

Statistical analysis

We compared the distribution of demographics and other potential confounders between NPC cases and controls using the Chi-squared test for categorical covariates and Student's t test for continuous variables. For between-group comparisons on alpha diversity, we used Wilcoxon and Kruskal–Wallis sum rank tests, where appropriate. Potentially significant alpha diversity predictors were identified based on a P value < 0.05 for inclusion in a subsequent regression model.

Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated using multivariate logistic regression to estimate associations of categorical alpha diversities with risk of NPC, adjusted for age, sex, sequencing run, EBV infection status, and residential community. For species enriched in NPC cases and controls, we also used multivariate logistic regression to evaluate associations between carriage of these species and risk of NPC, adjusting for the potential confounders included in the differential abundance analysis described above. We conducted linear trend tests by treating categorical variables as ordinal. In addition, for the associations of carriage of fungal species with NPC risk, we performed subgroup analysis by cigarette smoking, alcohol drinking, consumption of herbal medicine or salted fish, body mass index (BMI), tooth loss, tooth brushing frequency, family history of NPC, and history of ENT disease. We used likelihood ratio tests to assess heterogeneity of estimates by introducing an interaction term to the logistic models.

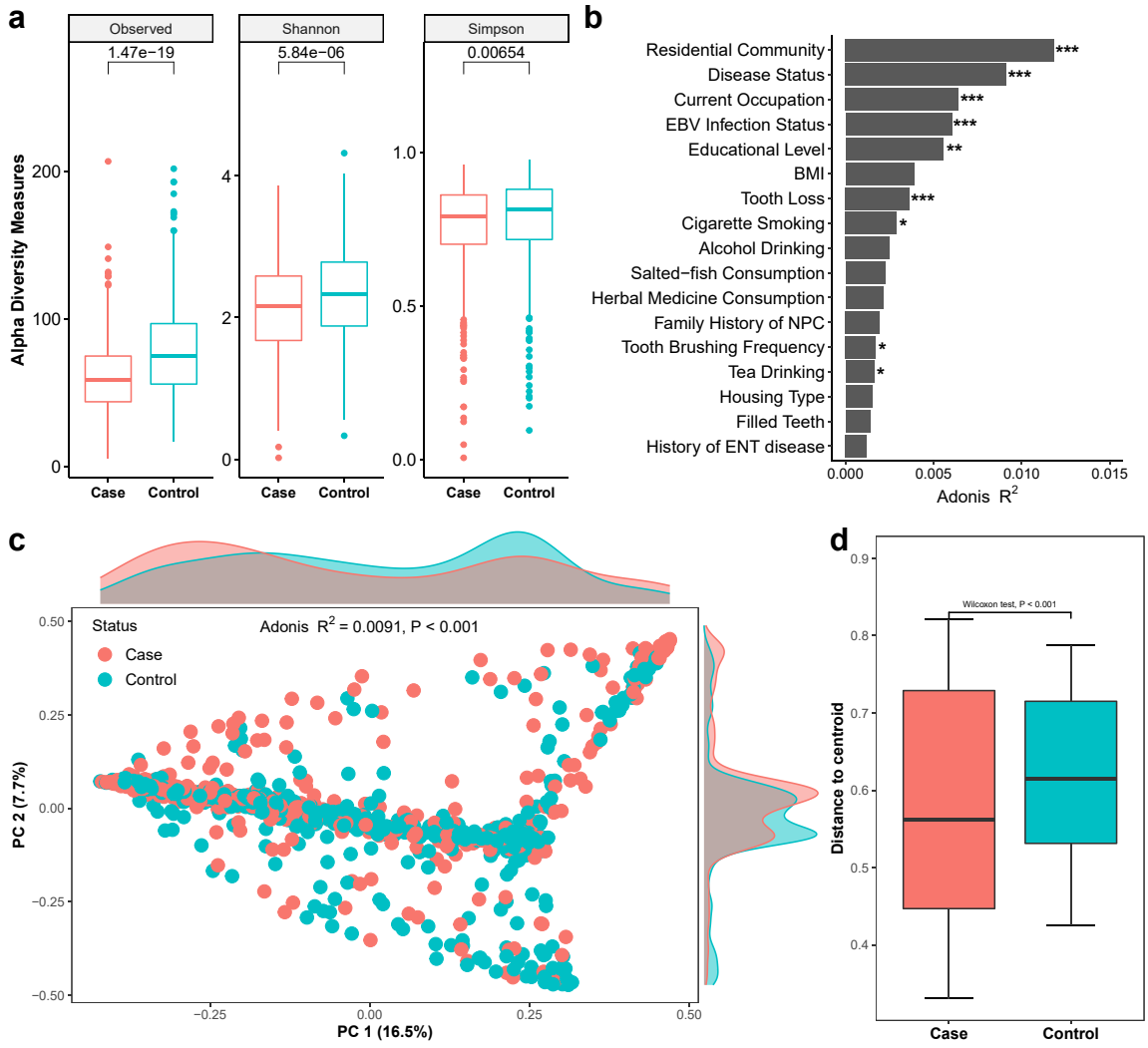

Fig. 1: **Comparison of the oral fungal microbiome between nasopharyngeal carcinoma cases and controls.** **(a)** Cases harboured significantly lower fungal richness (observed amplicon sequence variants [ASVs]) and diversity (Shannon and Simpson indexes) than controls (Wilcoxon tests, $P < 0.01$). **(b)** Adonis test (based on Bray–Curtis dissimilarity) adjusted for age, sex, and sequencing run, showed that there was a significant difference in fungal community structure by residential community, disease status, current occupation, Epstein–Barr virus infection status, educational level, tooth loss, cigarette smoking, tooth brushing frequency, and tea drinking. For 9999 permutations, ***, FDR-adjusted $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$. **(c)** Bray–Curtis dissimilarity-based principal coordinates analysis (PCoA) plot revealed a separation in fungal community structure between NPC cases and controls along PC1 and PC2. The axes are labelled with the variation explained, i.e., PC1 explained 16.5% and PC2 explained 7.7% of the variation. **(d)** The average Bray–Curtis dissimilarity in NPC cases was lower than that in controls (Wilcoxon test, $P < 0.001$).

In our previous study[19] using the same study samples, we noted that bacterial alpha diversity and three bacterial ASVs [i.e., Gran-7770 and Lact-ecat9 (more enriched in cases), and Gran-5a37 (more prevalent in controls)] were strongly associated with NPC risk. Therefore, using our previous 16S rRNA gene sequencing data,[19] we examined possible interactions between the fungal microbiome and these significant bacterial signals on the risk of NPC. Interactions were estimated on both the multiplicative and additive scales, with the former reflecting etiological relevance and the latter reflecting public health relevance.[36] We defined $OR_{11}/(OR_{10}{*}OR_{01}) > 1$ as positive, <1 as negative, and = 1 as no multiplicative interaction, respectively; and we defined relative excess risk due to interaction (RERI) > 0 as positive, <0 as negative, and = 0 as no additive interaction, respectively.[36]

We conducted two sensitivity analyses to assess the robustness of our primary findings: a) we excluded study subjects (66 cases, 18 controls) who had taken

antibiotics within one week before sample collection or had a chronic infectious disease; and b) we restricted to early-stage/localized NPC cases (i.e., AJCC 7th edition[37] stages I, II, and III; 232 cases). In addition, we explored the oral fungal microbiome by cancer stage among NPC cases.

A diagrammatic workflow for sequence data processing and bioinformatic and statistical analyses in this study is shown in Figure S2.

All statistical tests were two-sided, and P values < 0.05 were considered statistically significant. All analyses were conducted using R (version 4.1, R Foundation for Statistical Computing, Vienna, Austria).[29]

Role of the funding source

The funding sources had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. This work did not receive any support by a pharmaceutical company or other agency.

Results

Demographics of study population and sequencing summary

After denoising and quality filtering, 1013 subjects (476 pre-treatment cases, 537 controls) had both sufficiently high-quality sequences and participant data to be retained for analysis. The distributions of demographic characteristics and other covariates among the NPC cases and controls are summarized in Table 1. The majority (62%) of NPC cases' saliva samples were collected within one day after disease diagnosis, and over 95% were collected within 30 days after diagnosis; all controls' samples were collected on the date of interview. Nearly half of NPC cases (49%) were early-stage or localized at the time of diagnosis. Cases were more likely than controls to be less educated, live in a cottage (clay brick structure) or boat, have blue-collar jobs, consume herbal medicine, have a family history of NPC, have a lower frequency of brushing teeth, and have positive IgA antibodies against the EBV viral capsid antigen.

ITS-2 amplicon sequencing yielded an average of 32,246 reads per sample (cases vs. controls: 29,506 vs. 28,722 reads; range: 1122–137,861; median: 29,053) (Figure S3). In total, 9058 ASVs were identified from the 1013 participants.

Alpha diversity and beta diversity by disease status

Fig. 1a shows the average alpha diversity in NPC cases and controls. NPC cases exhibited significantly fewer observed ASVs, lower Simpson diversity, and lower Shannon diversity than controls (Wilcoxon tests, P values < 0.01). Compared with those harboring higher

Characteristics	Cases (n = 476)	Controls (n = 537)	P value[a]
	N (%)	N (%)	
Saliva collection after diagnosis/interview (days)			
Median (5%–95%)	1 (0–29)	0 (0–0)	
Age (years)			0.02
Mean (SD)	48.5 (10.7)	50.1 (10.7)	
Sex			0.69
Male	344 (72.3)	394 (73.4)	
Female	132 (27.7)	143 (26.6)	
Residential community			<0.001
Cangwu	133 (27.9)	97 (18.1)	
Cenxi	161 (33.8)	177 (33.0)	
Tengxian	86 (18.1)	158 (29.4)	
Wuzhou	96 (20.2)	105 (19.6)	
Educational level (years)			<0.001
<7	206 (43.3)	142 (26.4)	
7–9	164 (34.5)	217 (40.4)	
10–12	88 (18.5)	128 (23.8)	
>12	18 (3.8)	50 (9.3)	
BMI (10 years ago)			0.04
Underweight	54 (11.3)	57 (10.6)	
Normal	320 (67.2)	323 (60.1)	
Overweight	91 (19.1)	137 (25.5)	
Obese	11 (2.3)	20 (3.7)	
Housing type			<0.001
Building (concrete structure)	306 (64.3)	416 (77.5)	
Cottage (clay brick structure) or boat	170 (35.7)	121 (22.5)	
Job category			<0.001
Farmer	142 (29.8)	204 (38.0)	
Blue collar	197 (41.4)	144 (26.8)	
White collar	73 (15.3)	116 (21.6)	
Unemployed/others	64 (13.4)	73 (13.6)	
Cigarette smoking			0.18
Never	215 (45.2)	274 (51.0)	
Former	27 (5.7)	28 (5.2)	
Current	234 (49.2)	235 (43.8)	
Alcohol drinking			0.22
Never	321 (67.4)	389 (72.4)	
Former	21 (4.4)	20 (3.7)	
Current	134 (28.2)	128 (23.8)	
Tea drinking			0.42
Never	339 (71.2)	370 (68.9)	
Ever	137 (28.8)	167 (31.1)	
Salted fish consumption in adulthood			0.21
Never	385 (80.9)	436 (81.2)	
Yearly	64 (13.4)	82 (15.3)	
Monthly or more	27 (5.7)	19 (3.5)	
Herbal medicine consumption			0.04
Never	403 (84.7)	480 (89.4)	
Ever	61 (12.8)	52 (9.7)	
Unknown	12 (2.5)	5 (0.9)	

(Table 1 continues on next page)

Characteristics	Cases (n = 476)	Controls (n = 537)	P value[a]
	N (%)	N (%)	
(Continued from previous page)			
Tooth loss after age 20 years			0.46
No	237 (49.8)	255 (47.5)	
Yes	239 (50.2)	282 (52.5)	
Filled teeth after age 20 years			0.05
No	375 (78.8)	449 (83.6)	
Yes	101 (21.2)	88 (16.4)	
Frequency of brushing teeth, times/d			<0.001
≤1	316 (66.4)	242 (45.1)	
≥2	160 (33.6)	295 (54.9)	
Family history of NPC			<0.001
No	412 (86.6)	515 (95.9)	
Yes	52 (10.9)	20 (3.7)	
Unknown	12 (2.5)	2 (0.4)	
History of ENT disease			0.36
Never	432 (90.8)	496 (92.4)	
Ever	44 (9.2)	41 (7.6)	
Serum EBV status[b]			<0.001
Negative	14 (2.9)	309 (57.5)	
Positive	388 (81.5)	187 (34.8)	
Unknown	74 (15.5)	41 (7.6)	
Cancer stage[c]			
I	10 (2.1)		
II	28 (5.9)		
III	194 (40.8)		
IV	230 (48.3)		
Unknown	14 (2.9)		

Abbreviations: SD, standard deviation; BMI, body mass index; NPC, nasopharyngeal carcinoma; ENT, ear, nose, and throat; EBV, Epstein–Barr virus. [a]P value for age comparison was determined by a 2-sided Student's t test. Other P values were determined by Chi-squared test. [b]EBV status was determined by IgA antibodies against the EBV viral capsid antigen. [c]Cancer stage was classified according to the 7th AJCC version.

Table 1: **Demographic characteristics of study population.**

alpha diversity, the odds of NPC was significantly greater for those with low observed ASVs (adjusted OR = 5.81, 95% CI = 3.60–9.38), low Simpson diversity (OR = 1.53, 95% CI = 1.03–2.29), or low Shannon diversity (OR = 2.03, 95% CI = 1.35–3.04). Odds of NPC also increased with reduced alpha diversity ($P_{trend} < 0.05$ for all three alpha diversity indexes) (Table 2).

Adonis models (9999 permutations, based on Bray–Curtis dissimilarity) minimally adjusted for sex, age, and sequencing run revealed a significant difference in fungal community pattern (beta diversity) by residential community, disease status, EBV infection status, educational level, BMI, tooth loss, housing type, history of rhinitis, and tea drinking (FDR-adjusted $P < 0.05$), with residential community and disease status accounting for the top two strongest explanatory factors for Bray–Curtis dissimilarity (Fig. 1b). Although there was some overlap, the PCoA plot illustrated that controls' fungal communities clustered separately from NPC cases' fungal communities (Adonis test, FDR-adjusted $P < 0.001$, Fig. 1c). Accordingly, the average Bray–Curtis dissimilarity in cases was significantly lower than that in controls (Wilcoxon test, $P = 0.001$, Fig. 1d). Similar results were obtained based on Jaccard dissimilarity (data not shown).

Relative abundance of fungal communities by disease status

We further filtered out ASVs with ambiguous (missing or unclassified) taxa annotations at each taxon level before taxonomic assignment. Subsequently, we classified the filtered ASVs into 6 phyla, 25 classes, 78 orders, 195 families, 377 genera, and 555 species. The distributions of relative abundance of the 10 most predominant fungal genera and the top 5 abundant fungal organisms on other taxonomic levels by disease status are presented in Fig. 2 and Figure S4. The top 2 most abundant phyla accounting for 99.6% of the total abundance in both NPC patients and controls were *Ascomycota* (mean relative abundance, 86.5%) and *Basidiomycota* (13.1%). The 5 most abundant genera were *Malassezia* (21.2%), *Candida* (17.6%), *Blumeria* (10.5%), *Cladosporium* (10.2%), and *Aspergillus* (6.3%).

To determine which fungal organisms distinguished between NPC cases and controls, we used ANCOMBC analyses adjusting for multiple confounders. In total, we identified 9 classes, 13 orders, 13 families, 14 genera, and 11 species that differed significantly between cases and controls (FDR-adjusted $P < 0.05$) (Fig. 3). For instance, the fungal species *Saccharomyces cerevisiae*,

Alpha diversity[a]	Cases (n = 450)	Controls (n = 523)	OR (95% CI)	
	N (%)	N (%)	Crude	Adjusted[b]
Observed ASVs				
High	50 (11.1)	171 (32.7)	1.00 (Ref.)	1.00 (Ref.)
Medium	143 (31.8)	176 (33.7)	2.78 (1.89–4.08)	2.71 (1.68–4.36)
Low	257 (57.1)	176 (33.7)	4.99 (3.45–7.22)	5.81 (3.60–9.38)
P_{trend}			<0.001	<0.001
Simpson diversity				
High	120 (26.7)	175 (33.5)	1.00 (Ref.)	1.00 (Ref.)
Medium	156 (34.7)	173 (33.1)	1.32 (0.96–1.81)	1.22 (0.83–1.81)
Low	174 (38.7)	175 (33.5)	1.45 (1.06–1.98)	1.53 (1.03–2.29)
P_{trend}			0.02	0.04
Shannon diversity				
High	99 (22.0)	175 (33.5)	1.00 (Ref.)	1.00 (Ref.)
Medium	154 (34.2)	174 (33.3)	1.56 (1.13–2.17)	1.55 (1.03–2.32)
Low	197 (43.8)	174 (33.3)	2.00 (1.45–2.76)	2.03 (1.35–3.04)
P_{trend}			<0.001	<0.001

Abbreviation: OR, odds ratio; CI, confidence interval; ASVs, amplicon sequence variants. [a]Samples (26 cases and 14 controls) with fewer than 5000 reads were not included in the diversity analysis. Alpha diversity was categorized by tertiles of the distribution among controls. Linear trend tests were conducted by treating the categorical variable as an ordinal variable. [b]Estimates were calculated using logistic regression models adjusted for age, sex, sequencing run, Epstein–Barr virus infection status, and residential community.

Table 2: **ORs and 95% CIs for nasopharyngeal carcinoma in relation to oral fungal alpha diversity.**

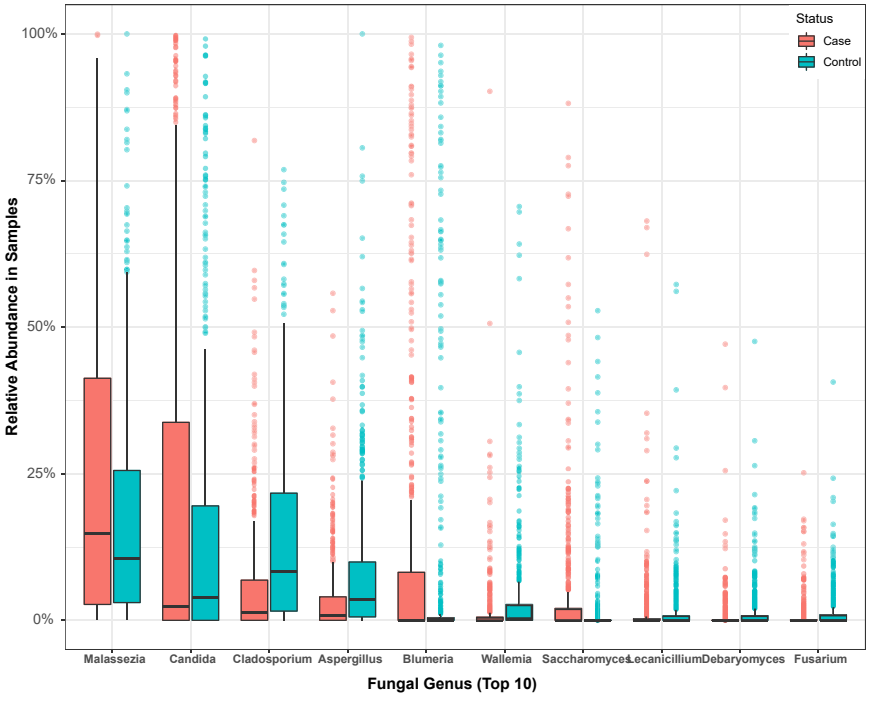

Fig. 2: **Distribution of relative abundance of predominant oral fungal genera in nasopharyngeal carcinoma cases and controls.**

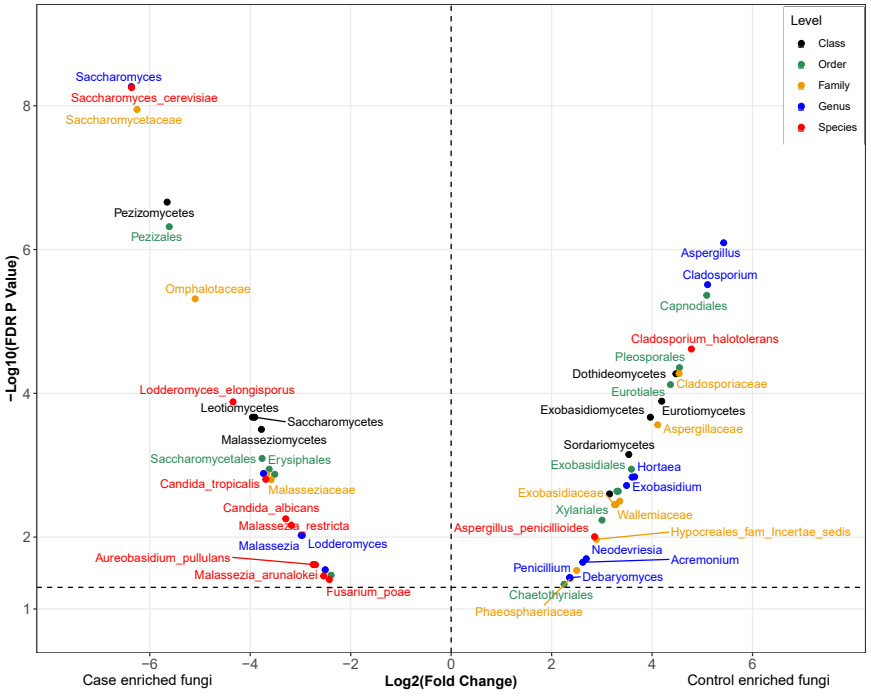

Fig. 3: **Differentially abundant oral fungal organisms on corresponding taxonomic levels in nasopharyngeal carcinoma cases and controls**. R package ANCOMBC (v. 1.2.2) was used for the differential abundance analysis, adjusted for age, sex, sequencing run, residential community, Epstein–Barr virus infection status, current occupation, educational level, tooth loss, cigarette smoking, tooth brushing frequency, and tea drinking. Horizontal dashed line indicates significance at FDR-adjusted $P < 0.05$. Only fungal organisms with FDR-adjusted $P < 0.05$ are shown. Fungal organisms enriched in NPC cases are shown on the left side of the vertical dashed line; those enriched among controls are shown on the right side of the vertical dashed line.

Candida tropicalis, *Lodderomyces elongisporus*, *Candida albicans*, *Fusarium poae*, *Malassezia arunalokei*, *Malassezia restricta*, *Malassezia globosa*, and *Aureobasidium pullulans* were significantly enriched in NPC patients, whereas *Cladosporium halotolerans* and *Aspergillus penicillioides* were significantly more abundant in controls.

For the differentially enriched fungal species, we further assessed associations between carriage of these species and risk of NPC. Among the nine species enriched in NPC cases, except for the three *Malassezia* species and *Aureobasidium pullulans*, carriage of the other five species was associated with a 1.6- to 4.7-fold elevated odds of NPC, compared with no carriage (Table 3). For example, the fully adjusted OR for low carriage of *Saccharomyces cerevisae* vs. none was 2.25 (95% CI = 1.29–3.91), and that for high carriage was 4.66 (95% CI = 2.78–7.81) (P_{trend} < 0.001). In contrast, compared with no carriage, carriage of *Cladosporium halotolerans* and *Aspergillus penicillioides* was associated with a 81% and 64% lower odds of NPC, respectively. The OR estimates increased profoundly with carriage of more types of high-risk species, and the estimates tended to be lower with carriage of more types of low-risk species (Table S3). We found little evidence of heterogeneity in the associations with carriage of these fungal species by cigarette smoking, alcohol drinking, consumption of herbal medicine or salted fish, BMI, tooth loss, tooth brushing frequency, family history of NPC, or history of ENT disease (data not shown). However, the associations with some species appeared to be modified by current occupation and tea drinking (Table S4). For example, the associations with high-risk species were stronger among non-farmers and non-tea drinkers, compared with farmers and tea drinkers, respectively.

Interaction of mycobiome and bacteriome

We observed significant positive additive interactions of fungal alpha diversity and bacterial alpha diversity on NPC risk (Table 4). For instance, individuals with both low fungal and low bacterial alpha diversities had a markedly elevated odds of NPC compared with those with both high alpha diversities (measures by observed ASVs: OR = 12.9, 95% CI = 7.41–22.5; RERI = 7.34, 95% CI = 2.23–12.5). Positive multiplicative interaction was detected between fungal Shannon diversity and bacterial observed ASVs or Faith's phylogenetic diversity.

In addition, we noted significant positive additive and multiplicative interactions between *Cladosporium halotolerans* and Gran-5a37. We found negative multiplicative interactions between Gran-7770 and *Saccharomyces cerevisiae* or *Lodderomyces elongisporus*, but little evidence of additive or multiplicative interactions between other examined fungal species and bacterial ASVs (Table 5).

Fungal species[a]	Cases (n = 476)	Controls (n = 537)	OR (95% CI)	
	N (%)	N (%)	Crude	Adjusted[b]
Saccharomyces cerevisiae				
No carriage	307 (64.5)	433 (80.6)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	56 (11.8)	52 (9.7)	1.52 (1.01–2.28)	2.25 (1.29–3.91)
High carriage	113 (23.7)	52 (9.7)	3.06 (2.14–4.39)	4.66 (2.78–7.81)
P_{trend}			<0.001	<0.001
Candida tropicalis				
No carriage	393 (82.6)	493 (91.8)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	56 (11.8)	22 (4.1)	3.19 (1.92–5.32)	3.62 (1.76–7.46)
High carriage	27 (5.7)	22 (4.1)	1.54 (0.86–2.75)	1.50 (0.68–3.29)
P_{trend}			0.001	0.009
Lodderomyces elongisporus				
No carriage	417 (87.6)	492 (91.6)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	34 (7.1)	22 (4.1)	1.82 (1.05–3.17)	2.64 (1.21–5.77)
High carriage	25 (5.3)	23 (4.3)	1.28 (0.72–2.29)	1.40 (0.65–3.01)
P_{trend}			0.09	0.07
Candida albicans				
No carriage	261 (54.8)	297 (50.4)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	81 (17.0)	120 (22.3)	0.76 (0.54–1.02)	0.76 (0.48–1.19)
High carriage	134 (28.2)	120 (22.3)	1.36 (1.00–.84)	1.56 (1.02–2.40)
P_{trend}			0.15	0.08
Fusarium poae				
No carriage	414 (87.0)	469 (87.3)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	39 (8.2)	34 (6.3)	1.30 (0.81–2.10)	3.03 (1.50–6.14)
High carriage	23 (4.8)	34 (6.3)	0.77 (0.44–1.32)	0.67 (0.32–1.41)
P_{trend}			0.73	0.60
Malassezia arunalokei				
No carriage	145 (30.5)	187 (34.8)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	157 (33.0)	175 (32.6)	1.16 (0.85–1.57)	1.35 (0.90–2.04)
High carriage	174 (36.6)	175 (32.6)	1.28 (0.95–1.73)	1.48 (0.98–2.22)
P_{trend}			0.11	0.06
Malassezia restricta				
No carriage	80 (16.8)	102 (19.0)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	164 (34.5)	217 (40.4)	0.96 (0.68–1.38)	0.87 (0.53–1.43)
High carriage	232 (48.7)	218 (40.6)	1.36 (0.96–1.92)	1.38 (0.84–2.27)
P_{trend}			0.03	0.06
Malassezia globosa				
No carriage	199 (41.8)	246 (45.8)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	144 (30.3)	146 (27.2)	1.22 (0.91–1.64)	1.18 (0.79–1.78)
High carriage	133 (27.9)	145 (27.0)	1.13 (0.84–1.53)	1.30 (0.86–1.96)
P_{trend}			0.34	0.21
Aureobasidium pullulans				
No carriage	384 (80.7)	433 (80.6)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	50 (10.5)	52 (9.7)	1.08 (0.72–1.64)	1.29 (0.74–2.27)
High carriage	42 (8.8)	52 (9.7)	0.91 (0.59–1.40)	1.06 (0.61–1.84)
P_{trend}			0.82	0.61
Cladosporium halotolerans				
No carriage	285 (59.9)	197 (36.7)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	126 (26.5)	170 (31.7)	0.51 (0.38–0.69)	0.39 (0.26–0.59)
High carriage	65 (13.7)	170 (31.7)	0.26 (0.19–0.37)	0.19 (0.12–0.30)
P_{trend}			<0.001	<0.001

(Table 3 continues on next page)

Fungal species[a]	Cases (n = 476)	Controls (n = 537)	OR (95% CI)	
	N (%)	N (%)	Crude	Adjusted[b]
(Continued from previous page)				
Aspergillus penicillioides				
No carriage	325 (68.3)	288 (53.6)	1.00 (Ref.)	1.00 (Ref.)
Low carriage	107 (22.5)	124 (23.1)	0.77 (0.57–1.04)	0.73 (0.49–1.08)
High carriage	44 (9.2)	125 (23.3)	0.31 (0.21–0.46)	0.36 (0.22–0.58)
P_{trend}			<0.001	<0.001

Abbreviation: OR, odds ratio; CI, confidence interval. [a]Selected oral fungal species were significantly enriched in nasopharyngeal carcinoma cases (first 9 species) or controls (latter 2 species). Abundance of selected species was classified as no carriage (relative abundance was zero), low carriage, or high carriage (categorized by median abundance of the distribution among controls who carried the species). Linear trend tests were conducted by treating the categorical variable as an ordinal variable. [b]Estimates were calculated using logistic regression models adjusted for age, sex, sequencing run, EBV infection status, residential community, current occupation, educational level, tooth loss, tooth brushing frequency, cigarette smoking, and tea drinking.

Table 3: **ORs and 95% CIs for nasopharyngeal carcinoma in relation to carriage of selected oral fungal species.**

Sensitivity analyses

The results were almost unchanged after we excluded those who used antibiotics within one week before saliva collection or had a chronic infectious disease (Tables S5 and S6). Restricting the analysis to early-stage or localized NPC cases yielded similar but slightly attenuated associations compared with the primary analysis (Tables S7 and S8). Among the NPC cases, we found that patients with lower alpha diversity were more likely to be diagnosed at an advanced cancer stage (Table S9); no significant associations with stage at diagnosis were observed for beta diversity or differential abundance (data not shown).

Discussion

In this large, population-based study, NPC patients harboured fewer and less diverse oral fungal communities than controls, and lower alpha diversity was associated with increased risk of NPC in a dose–response manner. Beta diversity—that is, fungal community patterns measured by Bray–Curtis dissimilarity—also differed by disease status. Furthermore, we identified a number of differentially abundant oral fungal organisms between NPC cases and controls, including five species that were associated with higher risk of NPC, and two species that were associated with lower NPC risk. In addition, we noted an additive interaction between fungal alpha diversity and bacterial alpha diversity, which contributed to profoundly increased risk of NPC. These findings suggest a potential role of the oral fungal community, as a distinct component of the oral microbiome, in NPC aetiology.

Our results indicate that both the richness and evenness of the oral fungal community were significantly lower in NPC patients; individuals with low alpha diversity had up to a 5.8-fold higher risk of NPC after adjustment for confounders. These findings suggest a significant reduction in fungal microbiome diversity that might have a link with NPC development. Similar observations were noted in studies that investigated the oral mycobiome in patients with OSCC, oral tongue cancer, and HNSCC, which also reported reduced alpha diversity in cancer patients compared with controls.[6–8,38] In addition, our multivariate Adonis models and PCoA plot showed a significant difference in fungal

Bacterial alpha diversity	Fungal alpha diversity			
	Observed ASVs		Shannon diversity	
	High	Low	High	Low
Observed ASVs				
High	1.00 (Ref.)	2.68 (1.47–4.89)	1.00 (Ref.)	0.87 (0.49–1.54)
Low	3.88 (2.16–6.98)	12.9 (7.41–22.5)	2.86 (1.69–4.84)	6.55 (3.85–11.1)
RERI for additive interaction	7.34 (2.23–12.5)		3.82 (1.30–6.34)	
P for multiplicative interaction	0.57		0.007	
Shannon diversity				
High	1.00 (Ref.)	2.50 (1.46–4.28)	1.00 (Ref.)	1.37 (0.82–2.28)
Low	1.75 (1.01–3.03)	7.05 (4.23–11.8)	2.17 (1.32–3.57)	3.86 (2.37–6.27)
RERI for additive interaction	3.81 (1.21–6.40)		1.31 (−0.10 to 2.72)	
P for multiplicative interaction	0.18		0.45	
Faith's phylogenetic diversity				
High	1.00 (Ref.)	2.64 (1.50–4.65)	1.00 (Ref.)	0.89 (0.52–1.53)
Low	2.47 (1.41–4.34)	8.57 (5.09–14.4)	1.95 (1.18–3.23)	4.45 (2.69–7.37)
RERI for additive interaction	4.46 (1.32–7.59)		2.61 (0.98–4.24)	
P for multiplicative interaction	0.46		0.007	

Abbreviation: ASVs, amplicon sequence variants; RERI, relative excess risk due to interaction. [a]Fungal or bacterial alpha diversity was categorized by median of the distribution among controls. Estimates were calculated using logistic regression models adjusted for age, sex, sequencing run, Epstein–Barr virus infection status, and residential community.

Table 4: **Interaction between oral fungal alpha diversity and bacterial alpha diversity on nasopharyngeal carcinoma[a].**

Case enriched fungal species and bacterial ASVs[b]				
Fungal species	**Bacterial ASVs**			
	Gran-7770		**Lact-ecat9**	
	No carriage	Carriage	No carriage	Carriage
Saccharomyces cerevisiae				
No carriage	1.00 (Ref.)	2.22 (1.45–3.40)	1.00 (Ref.)	1.63 (0.92–2.87)
Carriage	7.79 (3.68–16.47)	5.09 (2.97–8.74)	3.38 (2.18–5.25)	5.70 (2.22–14.7)
RERI for Additive interaction	−3.92 (−9.70 to 1.87)		1.69 (−3.72 to 7.11)	
P for multiplicative interaction	0.006		0.95	
Candida tropicalis				
No carriage	1.00 (Ref.)	1.80 (1.23–2.64)	1.00 (Ref.)	1.77 (1.04–3.01)
Carriage	4.40 (1.39–13.92)	3.49 (1.82–6.71)	2.81 (1.53–5.17)	2.39 (0.81–7.03)
RERI for additive interaction	−1.71 (−7.07 to 3.65)		−1.19 (−4.34 to 1.96)	
P for multiplicative interaction	0.22		0.28	
Lodderomyces elongisporus				
No carriage	1.00 (Ref.)	1.94 (1.33–2.83)	1.00 (Ref.)	1.37 (0.83–2.28)
Carriage	5.65 (1.66–19.24)	2.46 (1.24–4.88)	1.51 (0.83–2.74)	14.3 (2.46–83.6)
RERI for additive interaction	−4.12 (−11.2 to 2.93)		12.5 (−12.8 to 37.7)	
P for multiplicative interaction	0.04		0.05	
Candida albicans				
No carriage	1.00 (Ref.)	1.71 (1.08–2.72)	1.00 (Ref.)	2.65 (1.09–6.44)
Carriage	1.05 (0.58–1.92)	1.88 (1.15–3.07)	1.15 (0.78–1.7)	1.48 (0.83–2.65)
RERI for additive interaction	0.11 (−0.83 to 1.06)		−1.32 (−3.8 to 1.15)	
P for multiplicative interaction	0.91		0.18	
Fusarium poae				
No carriage	1.00 (Ref.)	1.71 (1.16–2.51)	1.00 (Ref.)	1.51 (0.90–2.53)
Carriage	1.33 (0.54–3.27)	2.92 (1.47–5.81)	1.39 (0.79–2.46)	3.76 (1.10–12.9)
RERI for additive interaction	0.88 (−1.22 to 2.99)		1.86 (−2.81 to 6.54)	
P for multiplicative interaction	0.65		0.42	
Control enriched fungal species and bacterial ASVs[c]				
Fungal species	**Bacterial ASV Gran-5a37**			
	Carriage		No carriage	
Cladosporium halotolerans				
Carriage	1.00 (Ref.)		1.57 (0.98–2.52)	
No carriage	2.87 (1.88–4.37)		12.1 (6.08–24.2)	
RERI for additive interaction	8.68 (0.76–16.6)			
P for multiplicative interaction	0.02			
Aspergillus penicillioides				
Carriage	1.00 (Ref.)		2.01 (1.16–3.48)	
No carriage	1.90 (1.25–2.89)		4.34 (2.46–7.63)	
RERI for additive interaction	1.43 (−0.72 to 3.59)			
P for multiplicative interaction	0.73			

Abbreviation: ASVs, amplicon sequence variants; RERI, relative excess risk due to interaction. [a]Estimates were calculated using logistic regression models adjusted for age, sex, sequencing run, EBV infection status, residential community, current occupation, educational level, tooth loss, tooth brushing frequency, cigarette smoking, and tea drinking. [b]Carriage of neither fungal species nor bacterial ASV was classified as the reference category. [c]Carriage of both control enriched fungal species and bacterial ASV as the reference category.

Table 5: **Interaction between oral fungal species and bacterial ASVs on nasopharyngeal carcinoma[a].**

composition and structure between cases and controls. Because NPC cases were newly diagnosed and samples were collected prior to radiotherapy or chemotherapy treatment, it is unlikely that the observed reduction in fungal microbiome diversity and difference in composition are treatment-related.[39,40] However, our retrospective study design was unable to determine whether these changes preceded NPC diagnosis; therefore, how these changes in the oral mycobiome might contribute to the aetiology of NPC needs to be further studied. Established evidence shows that balanced microbial community structure can prime the immune

system and provide colonization resistance against exogenous pathogens and indigenous pathobionts, whereas perturbation of stable states of commensal microbial communities, termed dysbiosis, could lead to immune dysregulation.[41–43] Therefore, a causal role of oral mycobiome dysbiosis in NPC onset is biologically plausible.

We identified several oral fungi that were differentially present between NPC patients and controls, characterized by an increased abundance in pathogenic or opportunistic pathogenic fungi and a decrease in symbiotic fungi in NPC patients. Carriage of certain fungal species was associated with profoundly higher risk of NPC. One possible mechanism for the increased abundances in these species could be that the patients might be already in an immune compromised or immunosuppressed state before and/or at disease onset, contributing to the overgrowth of these pathogenic or opportunistic pathogenic species.

While *Candida* was the most common fungal genus among both cases and controls, it was also one of the most differentially abundant genera between NPC patients vs. controls, specifically including *Candida albicans* and *Candida tropicalis* species. Overgrowth of *Candida albicans* was previously reported in the biopsies of OSCC and in the saliva of HNSCC patients.[6,7] *Candida* species infection has been implicated in carcinogenesis through multiple pathways.[44–46] *Candida* species have the capacity to produce mutagenic acetaldehyde from alcohol and formation of the potent animal carcinogen N-nitrosobenzylmethylamine (NBMA); certain *Candida* species have high nitrosation potential to produce potentially carcinogenic nitrosamines; *Candida* species may either induce dysplasia and colonize existing premalignant lesions, or promote the generation of precancerous conditions and their progression to cancer; and *Candida* species induce host immune responses to produce proteinases and pro-inflammatory mediators that may facilitate carcinogenesis. Thus, there are several biologically plausible mechanisms through which higher abundance of *Candida* might contribute to the aetiology of NPC.

We found that the species *Saccharomyces cerevisiae*, *Lodderomyces elongisporus*, and *Fusarium poae* were also over-represented in the oral cavity of NPC patients vs. controls. *Saccharomyces cerevisiae* is typically used for winemaking, baking, and brewing in the food industry, although it has been regarded as an opportunistic human pathogen with relatively low virulence.[47] *Lodderomyces elongisporus* is a pathogen that can cause human bloodstream infections,[48,49] while *Fusarium poae* is known as a cause of Fusarium head blight in cereal crops.[50] Our stratified analysis showed stronger associations among non-farmers and non-tea drinkers, but no evident heterogeneity by certain dietary factors, suggesting that associations are not necessarily more prominent among those with occupational (potentially high-dose) exposure to these species. More evidence is warranted to clarify whether these fungal species play a role in NPC carcinogenesis.

In our study, the most significantly enriched species among controls was from the genera *Cladosporium* and *Aspergillus*. In prior studies, *Cladosporium* and *Aspergillus* were core oral fungal taxa in healthy individuals.[3,51] Many species of *Cladosporium* are capable of producing secondary metabolites such as antibiotics, which are inhibitors of *Bacillus subtilis*, *Escherichia coli*, and *Candida albicans*.[52] Therefore, carrying these species may confer protective effects against the development of NPC. Future studies to explore these mechanisms are required to assess their potential for primary prevention and control strategies.

Recent data have shown that gut bacterial–fungal interactions may contribute to the pathogenesis of colorectal cancer.[53–55] Oral bacterial–fungal interactions have also been noted in the head and neck cancer[7] and oral tongue cancer.[8] However, the interaction of oral bacterial and fungal microbes in NPC carcinogenesis has not been studied previously. Our finding for the first time revealed a positive additive interaction between fungal and bacterial alpha diversities that was associated with remarkably increased NPC risk; and interactions between some fungal species and bacterial ASVs also resulted in much higher risk of NPC. Our study provides strong preliminary epidemiological evidence that oral bacterial–fungal interactions might play an important role in aetiology of NPC.

Furthermore, in addition to the dysbiosis of oral mycobiome observed in NPC cases, we noted a higher level of dysbiosis among cases with advanced cancer stage compared to cases in early stage. These findings suggest that applying anti-fungal treatment to modulate the oral mycobiome from the dysbiosis state to a balanced state might restore immune dysregulation, enhance immunity, and reduce the risk of oral fungal infection, thereby promoting the effect of NPC treatment and improving patients' prognosis and quality of life.

To our knowledge this is the largest population-based study in an NPC-endemic region to assess the potential association between oral mycobiome and NPC. Strengths of our study include the large sample size, stringent population-based study design, recruitment of NPC cases soon after diagnosis and before disease treatment, extensive quality control measures for participant data and sequencing data, inclusion of numerous covariates to adjust for confounding, and analysis of the oral bacterial and fungal microbiomes simultaneously.

However, some limitations must be acknowledged. First, we were unable to establish the temporal relationship between oral mycobiome dysbiosis and NPC risk due to the retrospective nature of the case–control

study design. Therefore, prospective cohort studies are required to verify and validate our observed associations. Second, despite FDR correction, false-positive significant findings might have arisen due to numerous statistical tests, particularly for the differential abundance analysis. Third, our results may be affected by certain unmeasured confounding, such as dietary factors. Finally, our results may not be generalized outside southern China, such as low-risk regions, given that the oral microbiome community is strongly affected by geographic region and dietary pattern.

In conclusion, our results reveal that a dysbiotic oral mycobiome characterized by reduced community richness and diversity, as well as an increased abundance in pathogenic or opportunistic pathogenic fungi and a decrease in symbiotic fungi, might increase the risk of NPC. We observed strong additive interactions between the fungal and bacterial microbiomes in relation to NPC risk. These findings, if confirmed by prospective studies, might offer insights into the NPC aetiology and pave a potential way for NPC prevention and control.

Contributors

E.C., G.H., H.O.A., R.K., and W.Y. conceived the study; Y.Z., Y.L., Y.Cai, and Z.Z. were responsible for sample collection and management; Y.C. and W.L. performed the lab work, supervised by W.Y.; Y.C. performed bioinformatic and biostatistical analyses, J.W.D. and L.M. contributed to the supervision and refinement of analyses; Y.C. wrote the manuscript, E.C., H.O.A. and W.Y. provided critical edits. All authors reviewed the manuscript and approved the final submission.

Y.C. and W.Y. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Data sharing statement

The data that support the findings of this study can be obtained upon request. Requests should be directed towards the corresponding author (weimin.ye@ki.se). The data are not publicly available due to privacy or ethical restrictions.

Declaration of interests

E.C. is an employee of GRAIL, LLC, which played no role in the development of this manuscript. The other authors declare no competing interests.

Acknowledgements

The original field work was supported by a grant from the National Cancer Institute at the US National Institutes of Health (R01 CA115873). This work was supported by the Swedish Research Council (2015-02625, 2015-06268, 2017-05814, 2019-01429) and the High-level Talents Research Start-up Project of Fujian Medical University (No. XRCZX2017035 and No. XRCZX2020034). Yufeng Chen was also partly supported by a scholarship from China Scholarship Council (201600160071).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2023.104813.

References

1 Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;486:207–214.
2 Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE. Defining the normal bacterial flora of the oral cavity. *J Clin Microbiol*. 2005;43:5721–5732.
3 Ghannoum MA, Jurevic RJ, Mukherjee PK, et al. Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. *PLoS Pathog*. 2010;6:e1000713.
4 Nilsson RH, Anslan S, Bahram M, Wurzbacher C, Baldrian P, Tedersoo L. Mycobiome diversity: high-throughput sequencing and identification of fungi. *Nat Rev Microbiol*. 2018;17:95–109.
5 Bandara HMHN, Panduwawala CP, Samaranayake LP. Biodiversity of the human oral mycobiome in health and disease. *Oral Dis*. 2019;25:363–371.
6 Perera M, Al-Hebshi NN, Perera I, et al. A dysbiotic mycobiome dominated by Candida albicans is identified within oral squamous-cell carcinomas. *J Oral Microbiol*. 2017;9:1385369.
7 Shay E, Sangwan N, Padmanabhan R, Lundy S, Burkey B, Eng C. Bacteriome and mycobiome and bacteriome-mycobiome interactions in head and neck squamous cell carcinoma. *Oncotarget*. 2020;11:2375–2386.
8 Mukherjee PK, Wang H, Retuerto M, et al. Bacteriome and mycobiome associations in oral tongue cancer. *Oncotarget*. 2017;8:97273–97289.
9 Peters BA, Wu J, Hayes RB, Ahn J. The oral fungal mycobiome: characteristics and relation to periodontitis in a pilot study. *BMC Microbiol*. 2017;17.
10 Li Y, Wang K, Zhang B, et al. Salivary mycobiome dysbiosis and its potential impact on bacteriome shifts and host immunity in oral lichen planus. *Int J Oral Sci*. 2019;11:13.
11 Baraniya D, Chen T, Nahar A, et al. Supragingival mycobiome and inter-kingdom interactions in dental caries. *J Oral Microbiol*. 2020;12:1729305.
12 Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209–249.
13 Chang ET, Ye W, Zeng Y-X, Adami H-O. The evolving epidemiology of nasopharyngeal carcinoma. *Cancer Epidemiol Biomarkers Prev*. 2021;30:1035–1047.
14 Chang ET, Ye W, Ernberg I, Zeng YX, Adami HO. A novel causal model for nasopharyngeal carcinoma. *Cancer Causes Control*. 2022;33:1013–1018.
15 Chien YC, Chen JY, Liu MY, et al. Serologic markers of Epstein-Barr virus infection and nasopharyngeal carcinoma in Taiwanese men. *N Engl J Med*. 2001;345:1877–1882.
16 Raab-Traub N. Epstein-Barr virus in the pathogenesis of NPC. *Semin Cancer Biol*. 2002;12:431–441.
17 Liu Z, Chang ET, Liu Q, et al. Oral hygiene and risk of nasopharyngeal carcinoma-A population-based case-control study in China. *Cancer Epidemiol Biomarkers Prev*. 2016;25:1201–1207.
18 Zhu X-X, Yang X-J, Chao Y-L, et al. The potential effect of oral microbiota in the prediction of mucositis during radiotherapy for nasopharyngeal carcinoma. *eBioMedicine*. 2017;18:23–31.
19 Debelius JW, Huang T, Cai Y, et al. Subspecies niche specialization in the oral microbiome is associated with nasopharyngeal carcinoma risk. *mSystems*. 2020;5:e00065.
20 Du Y, Feng R, Chang ET, et al. Influence of pre-treatment saliva microbial diversity and composition on nasopharyngeal carcinoma prognosis. *Front Cell Infect Microbiol*. 2022;12:831409.
21 Qiao H, Tan X-R, Li H, et al. Association of intratumoral microbiota with prognosis in patients with nasopharyngeal carcinoma from 2 hospitals in China. *JAMA Oncol*. 2022;8:1301.
22 Ye W, Chang ET, Liu Z, et al. Development of a population-based cancer case-control study in southern China. *Oncotarget*. 2017;8:87073–87085.
23 Tedersoo L, Lindahl B. Fungal identification biases in microbiome projects. *Environ Microbiol Rep*. 2016;8:774–779.
24 Tedersoo L, Bahram M, Põlme S, et al. Global diversity and geography of soil fungi. *Science*. 2014;346:6696.
25 Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;37:852–857.
26 Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17:10–12.
27 Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581–583.
28 Nilsson RH, Larsson K-H, Taylor AFS, et al. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res*. 2019;47:D259–D264.
29 R Core Development Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2020.

30 McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 2013;8:e61217.
31 Sørensen T. *A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons*. Copenhagen, Denmark: I kommission hos E. Munksgaard, København. The Royal Danish Society of Scientific Publications; 1948.
32 Oksanen J, Blanchet F, Friendly M, et al. *vegan: community ecology package*. Vienna, Austria: R Foundation for Statistical Computing; 2020.
33 Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol*. 2008;26:32–46.
34 Anderson MJ. *Permutational multivariate analysis of variance (PERMANOVA)*. Wiley StatsRef: Statistics Reference Online, 2017 John Wiley & Sons, Ltd; 2017.
35 Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. *Nat Commun*. 2020;11:3514.
36 VanderWeele TJ, Knol MJ. A tutorial on interaction. *Epidemiol Methods*. 2014;3:33.
37 Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*. 2010;17:1471–1474.
38 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*. 2018;4:255–262.
39 Robinson S, Peterson CB, Sahasrabhojane P, et al. Observational cohort study of oral mycobiome and interkingdom interactions over the course of induction therapy for leukemia. *mSphere*. 2020;5:e00048.
40 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*. 2020;10:16582.
41 Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. *Nature*. 2012;489:220–230.
42 Belkaid Y, Hand Timothy W. Role of the microbiota in immunity and inflammation. *Cell*. 2014;157:121–141.
43 Tomkovich S, Jobin C. Microbiota and host immune responses: a love-hate relationship. *Immunology*. 2016;147:1–10.
44 Ramirez-Garcia A, Rementeria A, Aguirre-Urizar JM, et al. Candida albicans and cancer: can this yeast induce cancer development or progression? *Crit Rev Microbiol*. 2016;42:181–193.
45 Sanjaya PR, Gokul S, Gururaj Patil B, Raju R. Candida in oral pre-cancer and oral cancer. *Med Hypotheses*. 2011;77:1125–1128.
46 Andrea H, Anett D, Gabor D, et al. Infectious agents associated with head and neck carcinomas. In: Donelli G, ed. *Advances in microbiology, infectious diseases and public healthed*. vol. 1. Switzerland: Springer International Publishing; 2016:63–80.
47 Alan M, Kevin K. Emergence of Saccharomyces cerevisiae as a human pathogen: implications for biotechnology. *Enzyme Microb Technol*. 1999;25:551–557.
48 Lockhart SR, Messer SA, Pfaller MA, Diekema DJ. Lodderomyces elongisporus masquerading as Candida parapsilosis as a cause of bloodstream infections. *J Oral Microbiol*. 2008;46:374–376.
49 Al-Obaid K, Ahmad S, Joseph L, Khan Z. Lodderomyces elongisporus: a bloodstream pathogen of greater clinical significance. *New Microbes New Infect*. 2018;26:20–24.
50 Vanheule A, De Boevre M, Moretti A, et al. Genetic divergence and chemotype diversity in the Fusarium head blight pathogen Fusarium poae. *Toxins*. 2017;9:255.
51 Andes DR, Dupuy AK, David MS, et al. Redefining the human oral mycobiome with improved practices in amplicon-based taxonomy: discovery of Malassezia as a prominent commensal. *PLoS One*. 2014;9:e90899.
52 Gallo ML, Seldes AM, Cabrera GM. Antibiotic long-chain and α,β-unsaturated aldehydes from the culture of the marine fungus Cladosporium sp. *Biochem Syst Ecol*. 2004;32:545–551.
53 Coker OO, Nakatsu G, Dai RZ, et al. Enteric fungal microbiota dysbiosis and ecological alterations in colorectal cancer. *Gut*. 2019;68:654–662.
54 Lin Y, Lau HC-H, Liu Y, et al. Altered mycobiota signatures and enriched pathogenic Aspergillus rambellii are associated with colorectal cancer based on multicohort fecal metagenomic analyses. *Gastroenterology*. 2022;163:908–921.
55 Liu N-N, Jiao N, Tan J-C, et al. Multi-kingdom microbiota analyses identify bacterial–fungal interactions and biomarkers of colorectal cancer across cohorts. *Nat Microbiol*. 2022;7:238–250.