Published in final edited form as:

Cancer. 2019 November 15; 125(22): 3993–4002. doi:10.1002/cnrc.32403.

Microbial characterization of esophageal squamous cell carcinoma and gastric cardia adenocarcinoma from a high-risk region of China

Dantong Shao, MD¹, Emily Vogtmann, PhD, MPH², Anqi Liu, MM¹, Junjie Qin, PhD³, Wen Chen, PhD¹, Christian C. Abnet, PhD, MPH², Wenqiang Wei, PhD^{4,*}

¹Department of Cancer Epidemiology, National Cancer Center/ National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College.

²Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, United States of America.

³Promogene Translational Research Institute, Shenzhen, Guangdong, China

⁴Cancer Registration Office, National Cancer Center/ National Clinical Research Center for Cancer/ Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College.

Abstract

Background: Little is known about the microbiota and upper gastrointestinal tumors.

Esophageal squamous cell carcinoma (ESCC) and gastric cardia adenocarcinoma (GCA) occur in adjacent organs, co-occur geographically, and share many risk factors, despite being of different tissue types.

Methods: We characterized the microbial communities of paired tumor and non-tumor samples from 67 ESCC and 36 GCA patients in Henan, China. DNA was extracted using the MOBIO PowerSoil kit. The V4 region of the 16S rRNA gene was sequenced using the MiniSeq and processed using QIIME1. LEfSe was used to identify differentially abundant microbes, the Wilcoxon rank-sum test was used to test alpha diversity differences, and the Permutational Multivariate Analysis of Variance was used to test for differences in beta diversity.

Results: The microbial environments of ESCC and GCA tissues were all comprised primarily of *Firmicutes*, *Bacteroidetes* and *Proteobacteria*. ESCC tumor tissues contained more *Fusobacteria* (3.9% vs. 1.7%) and less *Streptococcus* (12.0% vs. 30.2%) compared to non-tumor tissues. GCA non-tumor tissues had a greater abundance of *Helicobacter* (60.5% vs. 11.8%) which may be linked to the lower alpha diversity (58.0 vs. 102.5, $p=0.0012$) compared to tumor tissues.

*Corresponding author: weiwq@cicams.ac.cn; Tel: +86-10-87788661; Postal address: 17 South Panjiayuan Lane, Chaoyang District, Beijing, China, 100021.

Author Contributions: Dantong Shao: Conceptualization, methodology, investigation, data curation, writing-original draft. Emily Vogtmann: Conceptualization, writing-reviewing and editing. Anqi Liu: investigation, data curation. Junjie Qin: Methodology, software. Wen Chen: Supervision, funding acquisition. Christian C. Abnet: Writing-reviewing and editing. Wenqiang Wei: Conceptualization, supervision, project administration, funding acquisition

Conflict of Interest Statement: The authors declare no conflicts of interest.

Comparing ESCC and GCA non-tumor tissues, microbial composition ($p=0.0040$) and alpha diversity (87.0 vs. 58.0, $p=0.00052$) were significantly different. No significant differences were detected for alpha diversity within ESCC and GCA tumor tissues.

Conclusions: Our study showed differing microbial composition of ESCC and GCA paired tumor and non-tumor tissues and differences by organ site. Large-scale prospective cohort studies are needed to confirm our findings.

Abstract

Precis for use in the Table of Contents:

- This is the first study to investigate the microbial composition of ESCC tumor and non-tumor tissues through 16S rRNA gene sequencing and between ESCC and GCA tissues.
- We found that the levels of *Fusobacteria* increased in ESCC tumor tissues and there was a large decrease in the relative levels of *Helicobacter* in GCA tumor tissues compared to non-tumor tissues.

Keywords

Esophageal squamous cell carcinoma; Gastric cardia adenocarcinoma; 16S rRNA gene sequencing; microbiota; diversity

Introduction

Esophageal cancer and gastric cardia adenocarcinoma (GCA) are common types of upper gastrointestinal cancer and contribute to the worldwide cancer burden especially in developing countries. Esophageal cancer is the ninth most common cancer and the sixth leading cause of cancer death worldwide, with an estimated 572,034 new cases and 508,585 deaths in 2018¹. Esophageal squamous cell carcinoma (ESCC) accounts for approximately 90% of the overall incidence of esophageal cancer and China contributes nearly half of all esophageal cases worldwide². GCA is one of the major types of gastric cancer. High incidence rates of ESCC and GCA have been reported in the “Asian Esophageal Cancer Belt”, an area that stretches from northern Iran, to China in the east, and Russia to the north³, indicating that there may be some shared etiology between these two cancers.

These two cancers geographically co-occur at extraordinarily high rates in the same geographic regions suggesting a shared etiology. A number of factors are associated with ESCC² and GCA⁴ including tobacco smoking and consumption of high temperature liquids and foods. A single genomic locus at 10q23 in the gene *PLCE1*, is most strongly related to both cancers, suggesting that even genetic susceptibility is linked for these two distinct tumors⁵. However, much of the etiology of these two cancers is still poorly understood. Tooth loss and poor oral hygiene, which are associated with upper gastrointestinal cancer, are also related to changes in the oral microbiota, which suggests a potential role of the microbiota in the development of upper gastrointestinal cancers^{6,7}. And both cancers⁸ may be linked to gastric atrophy, often the result of *Helicobacter pylori* colonization of the gastric mucosa, which leads to neutralization of the stomach and possible changes in the microbiota

of the upper gastrointestinal tract. *H. pylori* is the only bacteria listed as a class I carcinogen⁹ and is the primary cause of non-cardia gastric cancer¹⁰ and cardia cancer in Asia¹¹. Emerging evidence has suggested that other microbes are associated with some types of cancer, including colorectal cancer (e.g., *Fusobacterium nucleatum* and *Escherichia coli*)^{12,13}. Furthermore, other microbes colonizing the human gastrointestinal tract may be related to several diseases, such as diabetes¹⁴, obesity¹⁵ and multiple sclerosis¹⁶.

Only a few studies have focused on the effect of the upper gastrointestinal tract microbiota on ESCC and GCA. In the case of esophageal diseases, researchers have investigated the microbial characteristics of reflux-related esophagitis, Barrett's esophagus (BE) and esophageal adenocarcinoma, but the results have been inconsistent. Some studies suggested that patients with reflux-related esophagitis and BE harbored a greater proportion of Gram-negative bacteria, including *Veillonella*, *Prevotella*, *Neisseria*, *Campylobacter* and *Fusobacterium* compared to healthy groups¹⁷⁻²⁰, while another study²¹ failed to detect differences between normal mucosa and that of reflux-related esophagitis or BE. For ESCC, lower microbial richness in the upper digestive tract was shown to be independently associated with esophageal squamous dysplasia²² (the precursor lesion of ESCC), while both *F. nucleatum*²³ and *Porphyromonas gingivalis*^{24,25} have been found to be associated with ESCC. The microbiota of the oral cavity^{25,26} and gastric tissue²⁷ has been also investigated in ESCC patients which showed overall decreased oral microbial diversity and higher abundance of *Clostridiales* and *Erysipelotrichales* in the gastric corpus of ESCC patients compared to control subjects. Research related to the microbial environment of GCA patients is also limited. A previous study indicated that gastric microbial composition differed between paired tumor and non-malignant tissues in both Chinese and Mexican patients²⁸. In addition, GCA patients with a higher tumor grade had an increased relative abundance of *H. pylori* and lower alpha diversity compared to lower GCA tumor grade²⁹. However, further data regarding the microbial characteristics of normal and tumor tissues of ESCC and GCA is needed.

To investigate the microbiota of upper gastrointestinal tract cancer, we compared paired tumor and non-tumor tissues from ESCC and GCA patients from a high-incidence area of China and investigated microbial differences between esophageal and gastric cardia tissues.

Materials and Methods

Study population and sample collection

We recruited 103 individuals with a pathologic diagnosis of ESCC or GCA aged 40-69 years who underwent esophagectomy in 2015 at the Linzhou Esophageal Cancer Hospital in the Henan province of China. This study was approved by the Institutional Review Board of the Cancer Hospital of the Chinese Academy of Medical Sciences and written informed consent was obtained prior to participation. Participant information including sex, age, smoking and alcohol drinking status, diet style, the history of gastric atrophy and oral condition was collected by trained staff using a questionnaire and body mass index (BMI) was measured by the interviewers. Information related to pathologic diagnosis and tumor stage were obtained from the medical records from the hospital.

A total of 67 paired ESCC tissue samples (tumor and paired non-tumor tissues) and 36 paired GCA tissue samples were collected by biopsy from resection specimens immediately after surgery. We defined non-tumor tissue as the biopsy located 4 cm or more from the malignant tissue. Biopsy samples were transferred immediately to sterile cryogenic tubes and stored at -80°C .

Laboratory handling and bioinformatics

DNA was extracted from the tissue samples using the MOBIO PowerSoil® DNA Isolation Kit protocol and stored at -20°C in solution C6 (10mM Tris). Barcoded amplicons were generated covering the V4 region of the 16S rRNA gene using the 515F (5'-GTGYCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACNVGGGTWTCTAAT-3') primers. PCR thermal cycling consisted of an initial denaturation step at 95°C for 5 minutes, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds, and extension at 72°C for 40 seconds, with a final extension step at 72°C for 4 minutes. The PCR amplicons were quantified using the Qubit dsDNA HS Assay Kit (Thermo Fisher/Invitrogen Cat. no. Q32854) following the manufacturer's instructions. Then the amplicon libraries were pooled at an equal mass of 100 ng per sample and sequenced on the Illumina MiniSeq platform at the length of $2\times 150\text{bp}$. In DNA extraction, PCR amplification and sequencing steps, positive and negative quality controls were included along with the study samples in each batch. The positive quality control was a fecal sample which has been deeply sequenced previously and the negative quality control was a sample of water without DNA or contamination.

Using Quantitative Insights into Microbial Ecology (QIIME 1.9.1)³⁰, the raw sequences were processed to remove low quality reads. We merged the forward and reverse reads and then clustered these reads into operational taxonomic units (OTUs) using a 97% similarity threshold. Taxonomy was assigned using the Greengenes database (Release 13.8)³¹ using the open reference method. One ESCC tumor tissue failed sequencing and 34 tissues (11 tumor tissues and 9 non-tumor tissues from ESCC cases and 4 tumor tissues and 10 non-tumor tissues from GCA cases) had fewer than 1000 reads and were excluded. The 1000 read cut point was selected after visualization of the Shannon index rarefaction curve (Supporting Information Figure 1). After exclusions, 61 ESCC tumor and non-tumor tissues with a mean of 18263 reads (standard deviation [SD]=12806) and 34 GCA tumor and non-tumor tissues with a mean of 13655 reads (SD=11834) were included in the analysis, consisting of 55 tumor tissues and 58 non-tumor tissues from ESCC patients (52 paired samples) as well as 32 tumor tissues and 26 non-tumor tissues from GCA patients (24 paired samples) (Figure. 1). We identified a total of 8068 OTUs. 2722 OTUs mapped to a reference sequence and 5346 OTUs were novel sequences. We generated alpha diversity metrics (i.e., observed OTUs and Shannon index) and beta diversity metrics (weighted and unweighted Unifrac distance) using QIIME after rarefaction to 1000 reads. Relative abundance tables from phylum to genus were generated using the full data in QIIME.

Statistical analysis

We used t-tests and chi-squared tests to compare demographic and other characteristics between ESCC and GCA patients. We used the Wilcoxon rank-sum test to determine alpha

diversity differences by tumor status or location. For beta diversity, we used Permutational Multivariate Analysis of Variance (PERMANOVA, R-vegan function *adonis*) to examine whether the overall bacterial community composition differed by tumor status or site. To visualize clustering between the different groups, principal coordinate analysis (PCoA) was performed based on weighted and unweighted Unifrac distances.

Linear discriminant analysis Effect Size (LEfSe) was performed to identify microbes associated with tumor status or site³². Using the non-parametric Kruskal-Wallis rank-sum test and Wilcoxon rank-sum test, LEfSe detects microbiota with significant differences between groups in multiple samples. Microbiota with an LDA score greater than 2.0³² were selected as potentially associated with tumor status or site. For tumor stage analyses, the Kruskal-Wallis rank-sum test was used to compare alpha diversity and the relative abundance of specific microbial taxa selected from LEfSe by tumor stage.

P-values less than 0.05 were considered statistically significant. All analyses were conducted using R 3.2.0.

Results

Demographic characteristics of ESCC and GCA patients

A total of 61 ESCC (including 52 paired sets) and 34 GCA (including 24 paired sets) patients were included in the final analysis after excluding samples with fewer than 1000 reads (Figure 1). The average age of the ESCC and GCA patients was 62.2 (SD 6.8) and 62.1 (SD 6.6) years, respectively. None of the ESCC patients nor GCA patients had a history or symptoms of gastric atrophy or periodontal ulcer. No statistically significant differences were detected for any of the baseline characteristics, including diet, between ESCC and GCA patients (Table 1, Supporting Information Table S1).

Microbiota associated with ESCC and GCA tumor tissues

For alpha diversity, in the ESCC tumor tissues, the Shannon index (4.2 vs. 3.9) was higher than non-tumor tissues and the observed OTUs (82.5 vs. 87.0) was lower than non-tumor tissues, but these differences were not statistically significant (Figure. 2a, 2b). In the GCA tumor tissues, the Shannon index (4.3 vs. 2.0, $p=0.00055$) and observed OTUs (102.5 vs. 58.0, $p=0.0012$) were significantly higher than non-tumor tissues (Figure. 2a, 2b). When considering microbial community composition (i.e., beta diversity) significant clustering was detected for the weighted Unifrac distance ($p=0.015$) between ESCC paired tumor and non-tumor tissues but not for the unweighted Unifrac distance (Figure. 3a, Supporting Information Figure. 2a, Supporting Information Table S2). GCA paired tumor and non-tumor tissues showed two distinct clusters by tumor status in both weighted and unweighted Unifrac distance ($p=0.0040$, $p=0.0020$, Figure. 3b, Supporting Information Figure. 2b).

Figure. 4 and Supporting Information Figure. 3 present the microbial relative abundances for ESCC and GCA patients at all taxonomic levels from phylum to genus. The microbiota associated with tumor status for ESCC and GCA patients from LEfSe is shown in Supporting Information Figure. 4a and 4b. Overall, ESCC tumor tissues tended to have microbial depletions compared to non-tumor tissues while GCA tumor tissues tended to

have multiple microbial enrichments compared to non-tumor tissues. ESCC tumor tissues had greater relative abundances of *Fusobacteria* at the phylum level (3.9% vs. 1.7%) with consistent associations at all taxonomic levels including *Fusobacterium* at the genus level (3.2% vs. 1.3%). ESCC tumor tissues also had lower abundances of *Firmicutes* at the phylum level (32.5% vs. 47.6%) and *Streptococcus* at the genus level (12.0% vs. 30.2%) compared to non-tumor tissues (Figure 4, Supporting Information Figure 4a). GCA tumor tissues had increased relative abundances of *Firmicutes* (27.8% vs. 11.1%), *Bacteroidetes* (33.8% vs. 19.0%), *Fusobacteria* (1.9% vs. 0.3%) and *Actinobacteria* (2.1% vs. 0.9%) but decreased relative abundances of *Proteobacteria* (32.1% vs. 67.5%) compared to non-tumor tissues. At the genus level, GCA tumor tissues had higher relative abundances of *Prevotella* (23.3% vs. 13.7%), *Streptococcus* (11.2% vs. 4.2%), *Veillonella* (8.6% vs. 3.7%), *Haemophilus* (4.3% vs. 1.0%) and *Neisseria* (4.8% vs. 1.5%) and lower relative abundance of the *Helicobacter* genus (11.8% vs. 60.5%) compared to non-tumor tissues (Figure. 4, Supporting Information Figure. 4b). The dramatically lower relative abundance of the *Helicobacter* genus in the GCA tumor tissues may have led to many of the other differences noted between the GCA tumor and non-tumor tissues.

Microbiota associated with tumor stage within ESCC and GCA patients

ESCC patients consisted of tumor stages 1, 2 and 3 while GCA patients were composed of tumor stages 2, 3 and 4 (Table 2). There were no significant differences by ESCC tumor stage for alpha diversity (both Shannon Index and Observed OTUs). Marginally significant differences were seen in alpha diversity by GCA tumor stage ($p=0.075$ and 0.066 for Shannon Index and Observed OTUs, respectively). We found a positive association between the relative abundance of *Fusobacterium* with increasing tumor stage in ESCC tissues. Within GCA patients, the relative abundance of *Helicobacter* and *Streptococcus* decreased and *Prevotella* increased with more advanced tumor stage although these were not statistically significant.

Microbial differences between the ESCC and GCA tissues

When comparing the microbial differences between ESCC and GCA tissues, the relative abundance of the *Proteobacteria* (32.1% vs. 13.4% for tumor tissues, 67.5% vs. 14.7% for non-tumor tissues), particularly the genus *Helicobacter* (11.8% vs. 0.04% for tumor tissues, 60.5% vs. 0.1% for non-tumor tissues), was higher in GCA patients compared to ESCC patients for both tumor and non-tumor tissues (Figure. 4, Supporting Information Figure. 4c, 4d). ESCC tumor tissues had higher relative abundances of phylum-level *Fusobacteria* (3.9% vs. 1.9%) and *Tenericutes* (3.6% vs. 0.3%) and genus-level *Mycoplasma* (3.6% vs. 0.3%) and *Fusobacterium* (3.2% vs. 1.3%) compared to GCA tumor tissues (Figure. 4, Supporting Information Figure. 4c). ESCC non-tumor tissues tended to have greater relative abundances of phylum-level *Firmicutes* (47.6% vs. 11.1%), *Bacteroidetes* (32.1% vs. 19%) and *Fusobacteria* (1.7% vs. 0.3%) and genus-level *Streptococcus* (30.2% vs. 4.2%), *Prevotella* (23.5% vs. 13.7%), *Veillonella* (7.0% vs. 3.7%) and *Haemophilus* (4.0% vs. 1.0%) compared to GCA non-tumor tissues (Figure. 4, Supporting Information Figure. 4d).

There were no statistically significant differences in alpha diversity as measured by the Shannon index (4.3 vs. 4.4) and observed OTUs (83.0 vs. 104.5) for ESCC and GCA tumor

tissues, however ESCC non-tumor tissues had significantly increased alpha diversity, as measured by the Shannon index (3.8 vs. 2.0, $p=0.00021$) and observed OTUs (87.0 vs. 58.0, $p=0.00052$) compared to GCA non-tumor tissues (Figure. 2a, 2b). PCoA plots based on the weighted and unweighted Unifrac distance suggested clustering of ESCC and GCA tumor tissues ($p=0.017$, $p=0.0040$, Figure. 5a, Supporting Information Figure. 5a). In addition, distinct clustering was detected for ESCC and GCA non-tumor tissues as measured by both weighted and unweighted Unifrac distances ($p=0.0010$ for both, Figure. 5b, Supporting Information Figure. 5b).

Discussion

In this study, we compared the microbial composition of tumor and non-tumor tissues for ESCC and GCA. The microbial environments of ESCC and GCA tumor and non-tumor tissues were comprised primarily of *Firmicutes*, *Bacteroidetes* and *Proteobacteria*. The microbial composition of paired tumor and non-tumor tissues were distinct for both ESCC and GCA patients. In ESCC tumor tissues, the relative abundance of multiple taxa decreased, while *Fusobacteria* increased. In GCA tumor tissues, there was a large decrease in the relative abundance of *Helicobacter*, and increases in multiple taxa including *Fusobacteria*, *Firmicutes*, and *Bacteroidetes*. *Helicobacter* appeared to dominate the GCA non-tumor tissue, with a mean relative abundance of 60.5%, which was also reflected by the increased alpha diversity in GCA tumor tissues compared to non-tumor tissues, likely due to the dominance of *Helicobacter* in non-tumor tissues. For non-tumor tissues, the esophagus and gastric cardia appeared to be distinct communities, but in the tumor, the microbial communities appeared to be slightly more similar, particularly in comparisons of alpha diversity.

We did not find any significant differences in alpha diversity for ESCC tumor and non-tumor tissues. In one previous study of the upper gastrointestinal microbiota, alpha diversity of esophageal squamous dysplasia patients was found to be lower than in healthy individuals, but this study utilized sampling from the entire upper gastrointestinal tract²² and not a comparison of tumor compared to non-tumor tissues. Similar to the variation of the microbial environment between tumor and non-tumor tissues in colorectal cancer³³, we discovered differing relative abundances of specific taxa in ESCC tumor and non-tumor tissues in this Chinese population. The relative abundance of *Fusobacterium* was significantly increased in ESCC tumor tissues compared to non-tumor tissues in our study. Within ESCC tumors, the relative abundance of *Fusobacterium* increased with more advanced stage, although this finding was not statistically significant which may be related to the limited sample size in the earlier tumor stages. But given the increasing relative abundance with more advanced stage, it is possible that the esophageal tumor microenvironment is more conducive for the growth of *Fusobacteria*, a gram-negative bacterium, versus *Fusobacterium* playing a role in etiology. In Japanese esophageal cancer patients, esophageal cancer tissues harbored more *Fusobacterium nucleatum* DNA than matched normal esophageal mucosa. Furthermore, *F. nucleatum*-positive patients had significantly shorter esophageal cancer survival²³, suggesting a potential role as a prognostic biomarker for esophageal cancer, which has also been seen for colorectal cancer³⁴. *F. nucleatum*, normally colonizing the human oral cavity, has been recognized as a pathogen in

periodontal diseases³⁵, ulcerative colitis³⁶ and colorectal cancer¹². *F. nucleatum* can stimulate secretion of MMP-9 and MMP-13 from epithelial cells³⁷, which could be involved in the development of several human malignancies, since degradation of collagen IV in the basement membrane and extracellular matrix facilitates tumor progression, including invasion, metastasis, growth and angiogenesis in the esophagus³⁸. Infection by *F. nucleatum* in gingival epithelial cells activates NLRP3 inflammasome. NLRP3 inflammasome activation activates caspase-1 and induces secretion of mature IL-1 β . Also, *F. nucleatum* infection can induce other DAMPs (damage-associated molecular patterns) that mediate inflammasomes including HMGB1 (high-mobility group box-1 protein), apoptosis-associated speck-like protein (ASP) with a similar time-course as caspase-1 activation³⁹. Studies have found that *P. gingivalis* was frequently detected in esophageal cancer and dysplasia tissues but rarely in noncancerous and normal tissues^{24,40}. *P. gingivalis* infection was also positively associated with multiple clinical pathologic characteristics, including differentiation status, metastasis, and overall survival rate²⁴. However, these studies investigated the presence of *P. gingivalis* rather than the relative abundance. We did not find significant differences in the relative abundance of *Porphyromonas* between ESCC tumor and non-tumor tissues which may also be due to decreased sample size or population differences. At least one other previous study also did not find a significant difference in the presence of *Porphyromonas* between ESCC patients and normal controls ($p=0.09$), but did not test for difference between the tumor and non-tumor tissues²⁵. Differences in *Streptococcus* in the esophagus have been previously observed and it was suggested that the esophageal microbiota can be classified into two types¹⁷: Type I dominated by *Streptococcus* and concentrated in the phenotypically normal esophagus; and Type II with a greater proportion of gram-negative bacteria and correlated with esophagitis and Barrett's esophagus. We also found higher levels of *Streptococcus* in ESCC non-tumor tissues compared to tumor tissues. In ESCC tumor analyses, the relative abundance of *Streptococcus* decreased with more advanced stage disease. This suggests that *Streptococcus*, a gram-positive bacterium commonly found in the oral cavity, which may be better able to colonize the non-diseased esophagus.

We found greater alpha diversity in GCA tumor tissues compared to non-tumor tissues, while Yu et al. also found increased alpha diversity in tumor tissues, but only in a Chinese population, not a Mexican population²⁸. *H. pylori* is a known cause of gastric cancer⁹. A meta-analysis of mostly Western studies showed that the association with *H. pylori* was found with gastric non-cardia cancers (OR 3.0; 95% CI 2.3–3.8) but not with GCA (OR 1.0; 95% CI 0.7–1.4)¹⁰. However, *H. pylori* seropositivity is associated with increased risks for GCA in China¹¹ and Iran⁴¹, but not in other countries, such as the US, Finland, Sweden and Norway¹⁰, which suggests that *H. pylori* may have a distinct association with GCA in China and Iran, perhaps through interactions with other microbiota. Because of these geographic differences, IARC now includes Asian cardia cancer in their estimates of *H. pylori* disease burden⁴². Few previous studies have considered the microbial composition of GCA outside of *H. pylori*. However, a recent study including 80 Chinese GCA patients found *H. pylori* to be dominant in both GCA tumor and non-tumor tissues, but that the relative abundance of *H. pylori* was higher in non-tumor tissues compared to matched tumor tissues²⁸. In our study, we also found that GCA tumor tissues had an overall distinct microbial composition

compared to non-tumor tissues. In our study, compared to GCA non-tumor tissues, tumor tissues had less *Proteobacteria* (the phylum which includes *Helicobacter*), and more *Bacteroidetes*, *Firmicutes* and *Fusobacteria*, which is consistent with the previous study²⁸. The relative abundance of *Helicobacter* ranked second in our GCA tumor tissues after the relative abundance of *Prevotella*, which is in contrast with the previous study²⁸. We did not find any significant difference in alpha diversity and the relative abundance of specific taxa by GCA stage, which is in accordance with a previous study²⁹. However, we observed a non-significant decrease in *Helicobacter* and increase in *Prevotella* with more advanced GCA tumor stage.

Although the esophagus is adjacent to the gastric cardia, the microbial composition of both tumor and non-tumor tissues between these two sites appeared to be different. ESCC tissues are squamous and GCA tissues are glandular and the different tissue types may contribute to the growth of distinct microbiota. However, the existence of *Helicobacter* in GCA tumor and non-tumor tissues also was a large source of this difference. Except for *Helicobacter*, other bacteria that colonized the GCA and ESCC non-tumor tissues are commonly found in the oral cavity, including *Streptococcus*, *Prevotella*, *Veillonella* and *Neisseria*, which is consistent with a previous study²⁸. Interestingly, the alpha and beta diversity of GCA and ESCC tumor tissues were more similar than observed for non-tumor tissue. It is possible that the tumor microenvironment creates a niche for similar microbial communities to colonize and thrive even in unique tissue types.

There are some limitations in our study. First, we did not have tissue specimens from healthy individuals, so our comparisons were restricted to tumor and non-tumor samples from individuals with cancer. Therefore, our conclusions about how the esophageal or gastric cardia microbiota differ is restricted to cancer patients and may not apply to healthy individuals. Second, our sample size was relatively small, particularly for the stage-specific analyses. Larger studies will be needed to confirm our associations. Third, the bacterial communities we measured likely include bacteria resident on the tissue and that of swallowed saliva. Finally, we only examined data cross-sectionally and prospective data is required to understand whether the observed differences between tumor and non-tumor tissues are due to the presence of the tumor, related to tumor development, progression, or patient survival.

We compared the microbiota of paired tumor and non-tumor tissues for ESCC and GCA as well as microbial differences between the two sites. We found the microbial communities were different between tumor and non-tumor tissues both from ESCC and GCA patients and that the relative abundance of some taxa may change due to advanced stage disease. In non-tumor tissues, the esophagus and gastric cardia had distinct microbial compositions, but the tumor environment may be more conducive to the growth of specific microbial communities since the esophageal and gastric cardia tumor communities were more similar. To our knowledge, this is the first study to describe the microbial environment of ESCC paired tumor and non-tumor tissues using next generation sequencing. This is also the first time that microbial communities of tissues were compared between the esophagus and the gastric cardia in the upper gastrointestinal tract. Our findings indicate that some specific

microbial taxa may be used in the future for early detection efforts for upper gastrointestinal cancer, although this requires additional studies including healthy individuals.

These findings are intriguing, but additional case-control studies as well as large-scale prospective cohorts are needed to verify our associations with *Fusobacterium* and ESCC. In addition, investigations of the biological mechanisms of the microbiota are needed to elucidate the association between complex microbial environments and upper gastrointestinal cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This work was supported by National Key Research and Development Program of precision medicine (2016YFC0901400), Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (CIFMS) (2016-I2M-3-001), National Natural Science Foundation of China (31570116), and the Intramural Research Program of the National Cancer Institute of the National Institutes of Health.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018.
2. Abnet CC, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology*. 2018;154(2):360–373. [PubMed: 28823862]
3. Glenn TF. Esophageal cancer. Facts, figures, and screening. *Gastroenterology nursing : the official journal of the Society of Gastroenterology Nurses and Associates*. 2001;24(6):271–273; quiz 274–275. [PubMed: 11837211]
4. Olefson S, Moss SF. Obesity and related risk factors in gastric cardia adenocarcinoma. *Gastric Cancer*. 2015;18(1):23–32. [PubMed: 25209115]
5. Abnet CC, Freedman ND, Hu N, et al. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nature genetics*. 2010;42(9):764–767. [PubMed: 20729852]
6. Abnet CC, Qiao YL, Dawsey SM, Dong ZW, Taylor PR, Mark SD. Tooth loss is associated with increased risk of total death and death from upper gastrointestinal cancer, heart disease, and stroke in a Chinese population-based cohort. *Int J Epidemiol*. 2005;34(2):467–474. [PubMed: 15659476]
7. Abnet CC, Kamangar F, Islami F, et al. Tooth loss and lack of regular oral hygiene are associated with higher risk of esophageal squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev*. 2008;17(11):3062–3068. [PubMed: 18990747]
8. Ye W, Held M, Lagergren J, et al. Helicobacter pylori infection and gastric atrophy: risk of adenocarcinoma and squamous-cell carcinoma of the esophagus and adenocarcinoma of the gastric cardia. *J Natl Cancer Inst*. 2004;96(5):388–396. [PubMed: 14996860]
9. IARC. Monographs on the evaluation of carcinogenic risks to humans volume 100b: a review of human carcinogens: biological agents. Lyon: International Agency for Research on Cancer 2012;100(Pt B):1–441.
10. Helicobacter, Cancer Collaborative G. Gastric cancer and Helicobacter pylori: a combined analysis of 12 case control studies nested within prospective cohorts. *Gut*. 2001;49(3):347–353. [PubMed: 11511555]
11. Limburg P, Qiao Y, Mark S, et al. Helicobacter pylori seropositivity and subsite-specific gastric cancer risks in Linxian, China. *J Natl Cancer Inst*. 2001;93(3):226–233. [PubMed: 11158192]

12. Zhang X, Zhu X, Cao Y, Fang JY, Hong J, Chen H. Fecal *Fusobacterium nucleatum* for the diagnosis of colorectal tumor: A systematic review and meta-analysis. *Cancer Med*. 2019;8(2):480–491. [PubMed: 30636375]
13. Arthur JC, Perez-Chanona E, Muhlbauer M, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science*. 2012;338(6103):120–123. [PubMed: 22903521]
14. Pedersen HK, Gudmundsdottir V, Nielsen HB, et al. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature*. 2016;535(7612):376–381. [PubMed: 27409811]
15. Liu R, Hong J, Xu X, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat Med*. 2017;23(7):859–868. [PubMed: 28628112]
16. Cekanaviciute E, Yoo BB, Runia TF, et al. Gut bacteria from multiple sclerosis patients modulate human T cells and exacerbate symptoms in mouse models. *Proceedings of the National Academy of Sciences of the United States of America*. 2017;114(40):10713–10718. [PubMed: 28893978]
17. Yang L, Lu X, Nossa CW, Francois F, Peek RM, Pei Z. Inflammation and intestinal metaplasia of the distal esophagus are associated with alterations in the microbiome. *Gastroenterology*. 2009;137(2):588–597. [PubMed: 19394334]
18. Liu N, Ando T, Ishiguro K, et al. Characterization of bacterial biota in the distal esophagus of Japanese patients with reflux esophagitis and Barrett's esophagus. *BMC infectious diseases*. 2013;13:130. [PubMed: 23496929]
19. Blackett KL, Siddhi SS, Cleary S, et al. Oesophageal bacterial biofilm changes in gastro-oesophageal reflux disease, Barrett's and oesophageal carcinoma: association or causality? *Alimentary pharmacology & therapeutics*. 2013;37(11):1084–1092. [PubMed: 23600758]
20. Gall A, Fero J, McCoy C, et al. Bacterial Composition of the Human Upper Gastrointestinal Tract Microbiome Is Dynamic and Associated with Genomic Instability in a Barrett's Esophagus Cohort. *PloS one*. 2015;10(6):e0129055. [PubMed: 26076489]
21. Amir I, Konikoff FM, Oppenheim M, Gophna U, Half EE. Gastric microbiota is altered in oesophagitis and Barrett's oesophagus and further modified by proton pump inhibitors. *Environmental microbiology*. 2014;16(9):2905–2914. [PubMed: 24112768]
22. Yu G, Gail MH, Shi J, et al. Association between upper digestive tract microbiota and cancer-predisposing states in the esophagus and stomach. *Cancer Epidemiol Biomarkers Prev*. 2014;23(5):735–741. [PubMed: 24700175]
23. Yamamura K, Baba Y, Nakagawa S, et al. Human Microbiome *Fusobacterium Nucleatum* in Esophageal Cancer Tissue Is Associated with Prognosis. *Clin Cancer Res*. 2016;22(22):5574–5581. [PubMed: 27769987]
24. Gao S, Li S, Ma Z, et al. Presence of *Porphyromonas gingivalis* in esophagus and its association with the clinicopathological characteristics and survival in patients with esophageal cancer. *Infectious agents and cancer*. 2016;11:3. [PubMed: 26788120]
25. Peters BA, Wu J, Pei Z, et al. Oral Microbiome Composition Reflects Prospective Risk for Esophageal Cancers. *Cancer Res*. 2017;77(23):6777–6787. [PubMed: 29196415]
26. Chen X, Winckler B, Lu M, et al. Oral Microbiota and Risk for Esophageal Squamous Cell Carcinoma in a High-Risk Area of China. *PloS one*. 2015;10(12):e0143603. [PubMed: 26641451]
27. Nasrollahzadeh D, Malekzadeh R, Ploner A, et al. Variations of gastric corpus microbiota are associated with early esophageal squamous cell carcinoma and squamous dysplasia. *Sci Rep*. 2015;5:8820. [PubMed: 25743945]
28. Yu G, Torres J, Hu N, et al. Molecular Characterization of the Human Stomach Microbiota in Gastric Cancer Patients. *Front Cell Infect Microbiol*. 2017;7:302. [PubMed: 28730144]
29. Yu G, Hu N, Wang L, et al. Gastric microbiota features associated with cancer risk factors and clinical outcomes: A pilot study in gastric cardia cancer patients from Shanxi, China. *Int J Cancer*. 2017;141(1):45–51. [PubMed: 28319273]
30. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7(5):335–336. [PubMed: 20383131]
31. DeSantis TZ, Hugenholtz P, Larsen N, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 2006;72(7):5069–5072. [PubMed: 16820507]

32. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011;12(6):R60. [PubMed: 21702898]
33. Burns MB, Lynch J, Starr TK, Knights D, Blekhman R. Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment. *Genome Med.* 2015;7(1):55. [PubMed: 26170900]
34. Flanagan L, Schmid J, Ebert M, et al. *Fusobacterium nucleatum* associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. *Eur J Clin Microbiol Infect Dis.* 2014;33(8):1381–1390. [PubMed: 24599709]
35. Griffen AL, Beall CJ, Campbell JH, et al. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *ISME J.* 2012;6(6):1176–1185. [PubMed: 22170420]
36. Tahara T, Shibata T, Kawamura T, et al. *Fusobacterium* detected in colonic biopsy and clinicopathological features of ulcerative colitis in Japan. *Dig Dis Sci.* 2015;60(1):205–210. [PubMed: 25102986]
37. Gursoy UK, Kononen E, Uitto VJ. Stimulation of epithelial cell matrix metalloproteinase (MMP-2, -9, -13) and interleukin-8 secretion by fusobacteria. *Oral Microbiol Immunol.* 2008;23(5):432–434. [PubMed: 18793368]
38. Groblewska M, Siewko M, Mroczko B, Szmitkowski M. The role of matrix metalloproteinases (MMPs) and their inhibitors (TIMPs) in the development of esophageal cancer. *Folia Histochem Cytobiol.* 2012;50(1):12–19. [PubMed: 22532131]
39. Bui FQ, Johnson L, Roberts J, et al. *Fusobacterium nucleatum* infection of gingival epithelial cells leads to NLRP3 inflammasome-dependent secretion of IL-1 β and the danger signals ASC and HMGB1. *Cell Microbiol.* 2016;18(7):970–981. [PubMed: 26687842]
40. Yuan X, Liu Y, Kong J, et al. Different frequencies of *Porphyromonas gingivalis* infection in cancers of the upper digestive tract. *Cancer letters.* 2017;404:1–7. [PubMed: 28705771]
41. Shakeri R, Malekzadeh R, Nasrollahzadeh D, et al. Multiplex *H. pylori* Serology and Risk of Gastric Cardia and Noncardia Adenocarcinomas. *Cancer Res.* 2015;75(22):4876–4883. [PubMed: 26383162]
42. Plummer M, de Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers attributable to infections in 2012: a synthetic analysis. *The Lancet Global health.* 2016;4(9):e609–616. [PubMed: 27470177]

Figure 1.
Diagram of study enrollment and eligible paired and total samples.

Figure 2. Microbial comparison between ESCC and GCA patients for alpha diversity.

Alpha diversity of 52 paired ESCC tumor and non-tumor tissues, 24 paired GCA tumor and non-tumor tissues, 55 ESCC and 32 GCA unpaired tumor tissues, 58 ESCC and 26 GCA unpaired non-tumor tissues assessed using the Shannon index (a) and observed OTUs (b).

Figure 3. Microbial comparison between paired tumor and non-tumor tissues from ESCC and GCA patients for beta diversity.

Comparison of beta diversity based on weighted Unifrac distance using a PCoA plot of 52 paired tumor and non-tumor tissues in ESCC patients (a), and of 24 paired tumor and non-tumor tissues in GCA patients (b).

Figure 4. Microbial relative abundances at the phylum and genus level for ESCC and GCA patients.

Microbial relative abundance from 55 ESCC tumor tissues, 58 ESCC non-tumor tissues, 32 GCA tumor tissues and 26 GCA non-tumor tissues at the phylum and genus level.

*[*Prevotella*] indicates unclassified genus similar to *Prevotella*.

Figure 5. Comparison of the microbial communities of unpaired ESCC and GCA tissues. Comparison of beta diversity based on weighted Unifrac distance of 55 ESCC and 32 GCA unpaired tumor tissues using a PCoA plot (a), and of 58 ESCC and 26 GCA unpaired non-tumor tissues (b).

Table 1.

Baseline characteristics of ESCC and GCA patients

Characteristics	N (%) / Mean $\pm$ SD	
	ESCC	GCA
Sex (Male)	45 (74 %)	25 (74 %)
Age (Mean $\pm$ SD)	62.2 $\pm$ 6.8 years	62.1 $\pm$ 6.6 years
Education		
No education	17 (28.3 %)	14 (42.4 %)
Primary education	19 (31.7 %)	12 (36.4 %)
Secondary education or more	24 (40 %)	7 (21.2 %)
BMI (Mean $\pm$ SD)	22.5 $\pm$ 2.7	22.5 $\pm$ 2.8
Household income (Mean $\pm$ SD)	6361.4 $\pm$ 5992.1 RMB	5440.7 $\pm$ 5310.9 RMB
Tobacco smoking		
Current	31 (51.7 %)	13 (39.4 %)
Former	5 (8.3 %)	5 (15.2 %)
Never	24 (40 %)	15 (45.5 %)
Alcohol intake (Yes)	26 (43.3 %)	8 (24.2 %)
Family history of cancer (Yes)	20 (33.9 %)	9 (27.3 %)

* There were no significant differences for any comparisons.

[†] BMI: Body Mass Index; ESCC: Esophageal Squamous Cell Carcinoma; GCA: Gastric Cardia Adenocarcinoma

Table 2.

Comparison of alpha diversity and the relative abundance of specific microbial taxa by tumor stage for ESCC and GCA patients.

Tumor stage [†]	ESCC				GCA			
	T1	T2	T3	p value	T2	T3	T4	p value
Number of patients	6	8	41		2	19	11	
Alpha diversity								
Shannon Index	4.1	3.9	4.3	0.77	1.4	4.7	4.3	0.075
Observed OTUs	82	80	84	0.47	51	107	103	0.066
Relative abundance								
<i>Fusobacterium</i> (%)	0.34	1.1	1.9	0.29	-	-	-	-
<i>Streptococcus</i> (%) [*]	11	17	5.9	0.030	1.1	7.4	6.3	0.15
<i>Helicobacter</i> (%)	-	-	-	-	46	0.39	0	0.50
<i>Prevotella</i> (%)	-	-	-	-	0.88	22	23	0.080

^{*} Comparison between tumor stage 2 and 3 for ESCC patients; p=0.040

[†] The tumor stage describes the size of the original tumor and whether it has invaded nearby tissue. ESCC patients consisted of tumor stages 1, 2 and 3 while GCA patients were composed of tumor stages 2, 3 and 4.

[‡] ESCC: Esophageal Squamous Cell Carcinoma; GCA: Gastric Cardia Adenocarcinoma