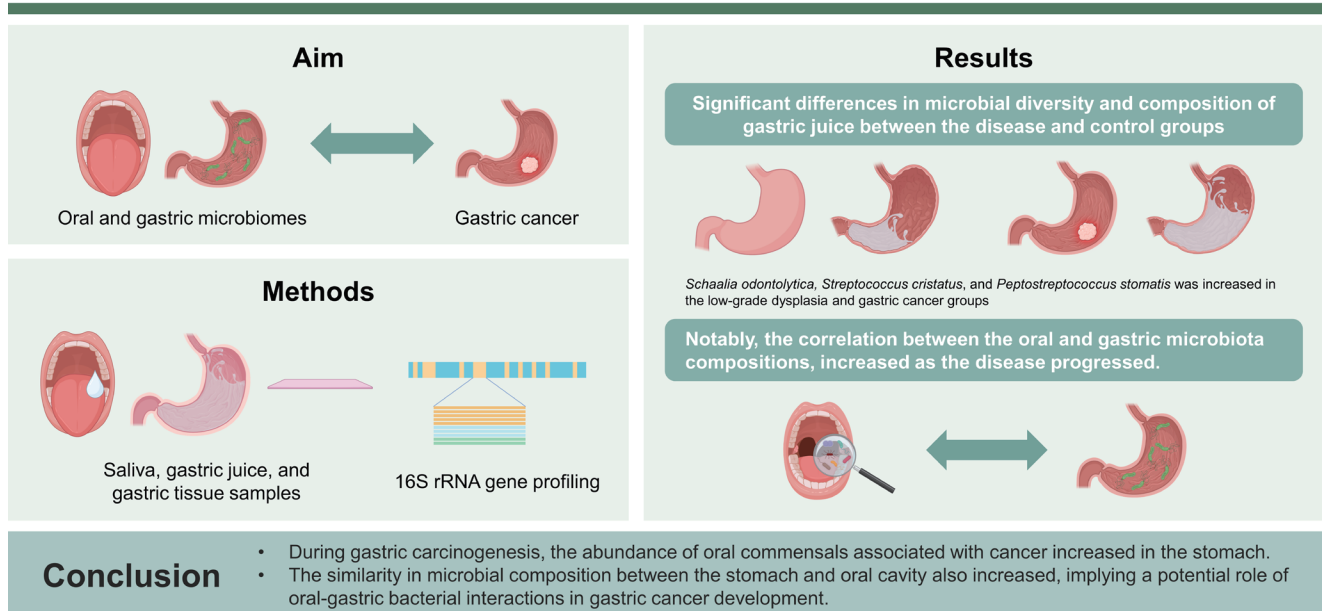KJIM

Increasing correlation between oral and gastric microbiota during gastric carcinogenesis

Hee Sang You[1,2], Jae Yong Park[2], Hochan Seo[1,2], Beom Jin Kim[2], and Jae Gyu Kim[2]

[1]Laboratory of Gastrointestinal Mucosal Immunology, Chung-Ang University College of Medicine, Seoul; [2]Department of Internal Medicine, Chung-Ang University College of Medicine, Seoul, Korea

Increasing correlation between oral and gastric microbiota during gastric carcinogenesis

Aim

Oral and gastric microbiomes — Gastric cancer

Methods

Saliva, gastric juice, and gastric tissue samples — 16S rRNA gene profiling

Results

Significant differences in microbial diversity and composition of gastric juice between the disease and control groups

Schaalia odontolytica, *Streptococcus cristatus*, and *Peptostreptococcus stomatis* was increased in the low-grade dysplasia and gastric cancer groups

Notably, the correlation between the oral and gastric microbiota compositions, increased as the disease progressed.

Conclusion
- During gastric carcinogenesis, the abundance of oral commensals associated with cancer increased in the stomach.
- The similarity in microbial composition between the stomach and oral cavity also increased, implying a potential role of oral-gastric bacterial interactions in gastric cancer development.

Background/Aims: Recent research has increasingly focused on the role of the gastric microbiome in the development of gastric cancer. We aimed to investigate the changes in the microbiome during gastric carcinogenesis in structural and functional aspects, with a specific focus on the association between oral and gastric microbiomes.

Methods: We collected saliva, gastric juice, and gastric tissue samples from 141 patients at different stages of gastric carcinogenesis and processed them for microbiome analysis using 16S rRNA gene profiling. The alpha and beta diversities were analyzed, and the differences in microbiome composition and function profiles were analyzed among the groups, as well as the correlation between changes in the oral and gastric microbiomes during carcinogenesis.

Results: We observed significant differences in microbial diversity and composition between the disease and control groups, primarily in the gastric juice. Specific bacterial strains, including *Schaalia odontolytica*, *Streptococcus cristatus*, and *Peptostreptococcus stomatis*, showed a significant increase in abundance in the gastric juice in the low-grade dysplasia and gastric cancer groups. Notably, the correlation between the oral and gastric microbiota compositions, increased as the dis-

You HS, et al. UGI tract micobiota in gastric cancer

KJIM✕

ease progressed. Predictive analysis of the metagenomic functional profiles revealed changes in functional pathways that may be associated with carcinogenesis (ABC transport and two-component systems).

Conclusions: During gastric carcinogenesis, the abundance of oral commensals associated with cancer increased in the stomach. The similarity in microbial composition between the stomach and oral cavity also increased, implying a potential role of oral-gastric bacterial interactions in gastric cancer development.

Keywords: Microbiome; Gastric cancer; Carcinogenesis; 16S rRNA

INTRODUCTION

Gastric cancer (GC) is a fatal disease that poses a significant health threat. Despite its decreasing incidence rate, it remains a significant global health concern and ranks as the fourth leading cause of cancer-related deaths worldwide [1]. Among the various risk factors associated with GC development, *Helicobacter pylori* infection is generally accepted as the most critical [2]. *H. pylori* infection plays a significant role in the perpetuation of chronic inflammation and irritation of the stomach lining, a condition known as gastritis, which leads to the development of premalignant gastric lesions [3].

Recently, there has been growing evidence that gastric bacteria other than *H. pylori* influence certain aspects of the development of GC [4,5]. In patients with GC, the composition of the gastric microbiota tends to shift in a direction similar to that of the intestinal or oral microbiota [6]. The oral cavity is a crucial route for the continuous supply of bacteria to the upper gastrointestinal tract, including the stomach. Non-*H. pylori* bacteria originating from the oral cavity may exert a more meaningful effect, especially in gastric environments with low acidity, such as severe mucosal atrophy. Nevertheless, although the gastric microbiome is suspected to play a role in GC development, few studies have focused on its association with the oral microbiome.

Furthermore, most previous studies have solely conducted structural analyses without considering functional aspects, which cannot provide significant insights into the mechanism of gastric carcinogenesis. To elucidate the role of bacteria in the occurrence of GC, concurrent metagenomic functional analyses should yield more revealing results [7].

Therefore, we aimed to characterize the changes in the microbiome during gastric carcinogenesis and assess its predictable functional profiles. Additionally, we aimed to investigate the correlation between the microbiomes of the oral cavity and stomach at different stages of gastric carcinogenesis.

METHODS

Study participants and sample collection

The participants were recruited at Chung-Ang University Hospital between 2017 and 2022. The inclusion criteria for the neoplasm group were patients newly diagnosed with histologically confirmed GC or adenoma, whereas the control group consisted of patients with gastritis only without evidence of GC or adenoma on endoscopic examination. Patients with a history of malignancy, those who underwent prior gastric surgery, those associated with a recent use of antibiotics or probiotics within the last 3 months, or those under 20 years of age were excluded. The final analysis group of 141 participants comprised 132 patients and nine controls. This case-control study was approved by the Institutional Review Board of the Chung-Ang University Hospital (IRB No. C2016047[1790]) and was conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants.

The Vienna Classification System was used to diagnose and categorize gastric neoplasms [8]. Category 3 (non-invasive low-grade adenoma/dysplasia) was classified as the low-grade dysplasia (LGD) group. Category 4.1 (high-grade adenoma/dysplasia) and category 4.2 (non-invasive carcinoma [carcinoma in situ]) were classified as the high-grade dysplasia (HGD) group. Category 4.3 (suspicion of invasive carcinoma) and 5 (invasive neoplasia) were classified as the GC group.

A minimum of 10 mL of saliva was collected by allowing it to accumulate on the floor of the mouth, followed by spitting it into a specimen tube. Samples were stored at -80°C until use. Gastric juice samples were obtained according to the following protocol: Patients fasted for more than 8 h before sample collection. For patients who underwent endoscopic examination or resection, a trap tube was connected between the endoscope and the suction tube before the

procedure. Gastric juice (7–30 mL) was collected by suction through the endoscope at the beginning of the procedure. Gastric tissues were collected from the same gastric sites in both the patient and control groups, following the same protocol. Two samples were obtained from the antrum and two from the body, resulting in a total of four tissue samples, which were then stored at -80°C. The endoscopes were washed and disinfected to avoid contamination by immersion in a detergent solution containing 7% proteolytic enzymes and 2% glutaraldehyde before use. Following the collection of gastric juice, the samples were kept at -20°C and immediately transferred to the nearby laboratory, without the use of preservative reagents. The collected gastric juice was diluted to 40 mM by the addition of 1 M Tris base and centrifuged at 600 × g for 10 minutes at 4°C to collect the supernatant. After transfer to a new 15 mL tube, it was again centrifuged at 1,500 × g for 10 minutes at 4°C. The supernatant was subsequently harvested and stored at -80°C.

DNA extraction and sequencing
All samples were diluted in 10 mL of PBS and incubated for 24 hours prior to DNA extraction. After the centrifugation step at 10,000 × g for 10 minutes at 4°C, the pellet containing the bacteria was diluted with 200 µL of PBS to create a suspension. Microbial genomic DNA was extracted using a DNeasy PowerSoil kit (QIAGEN, Hilden, Germany) according to the standard protocol provided in the manufacturer's instructions. The isolated genomic DNA was amplified by targeting the V3-V4 hypervariable regions of the 16S rRNA gene. This amplification was carried out using specific primers (16S_V3_F: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG- 3' and 16S_V4_R: 5'- GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC- 3'). Subsequently, amplicon libraries were prepared from the amplified DNA fragments. The MiSeq Reagent Kit v3 (600-cycle) (Illumina, San Diego, CA, USA) was used to sequence the library reagents. The Nextra XT Index Kit v2 Set A (96 indices, 384 samples) (Illumina) was used for barcodes and adapters. Library preparation for sequencing followed the 16S Metagenomic Sequencing Library Preparation (Part # 15,044,223 Rev. B). All amplicons were sequenced using MiSeq (Illumina) according to the manufacturer's instructions. The bacterial DNA in each sample was quantified using a QIAxpert system (QIAGEN).

Taxonomic assignment and profiling
Paired-end 16S rRNA gene sequences were input into the Quantitative Insights into Microbial Ecology software (QIIME2 v2021. 4) [9]. Adapter sequences were removed using cutadapt. Reads were filtered for quality and chimeric reads using dada2 with manual parameters (trim-left-f 0, trim-left-r 0, trunc-len-f 260, trunc-len-r 200, trunc-q 2,

Table 1. Clinical characteristics of patients in each sample group

Sample type	Variable	Total	Control	Low-grade dysplasia	High-grade dysplasia	Gastric cancer
Oral wash	Numbers	129	8	58	33	30
	Age (yr)	64.2 ± 11.3	34.9 ± 11.0	66.2 ± 8.3	66.5 ± 7.9	66.2 ± 8.4
	Sex (M:F)	96:33	3:5	42:16	24:9	27:3
Gastric juice	Numbers	141	9	58	33	41
	Age (yr)	64.4 ± 11.9	33.2 ± 11.4	66.2 ± 8.3	66.4 ± 83.0	67.0 ± 9.1
	Sex (M:F)	100:41	3:6	42:16	24:9	31:10
Tissue (antrum)	Numbers	130	9	58	33	30
	Age (yr)	63.8 ± 11.9	33.2 ± 11.4	66.2 ± 8.3	66.5 ± 7.9	65.5 ± 8.4
	Sex (M:F)	96:34	3:6	42:16	24:9	27:3
Tissue (body)	Numbers	130	9	58	33	30
	Age (yr)	63.8 ± 11.9	33.2 ± 11.4	66.2 ± 8.3	66.5 ± 8.0	65.5 ± 8.4
	Sex (M:F)	96:34	3:6	42:16	24:9	27:3

Values are presented as number only or mean ± standard deviation.
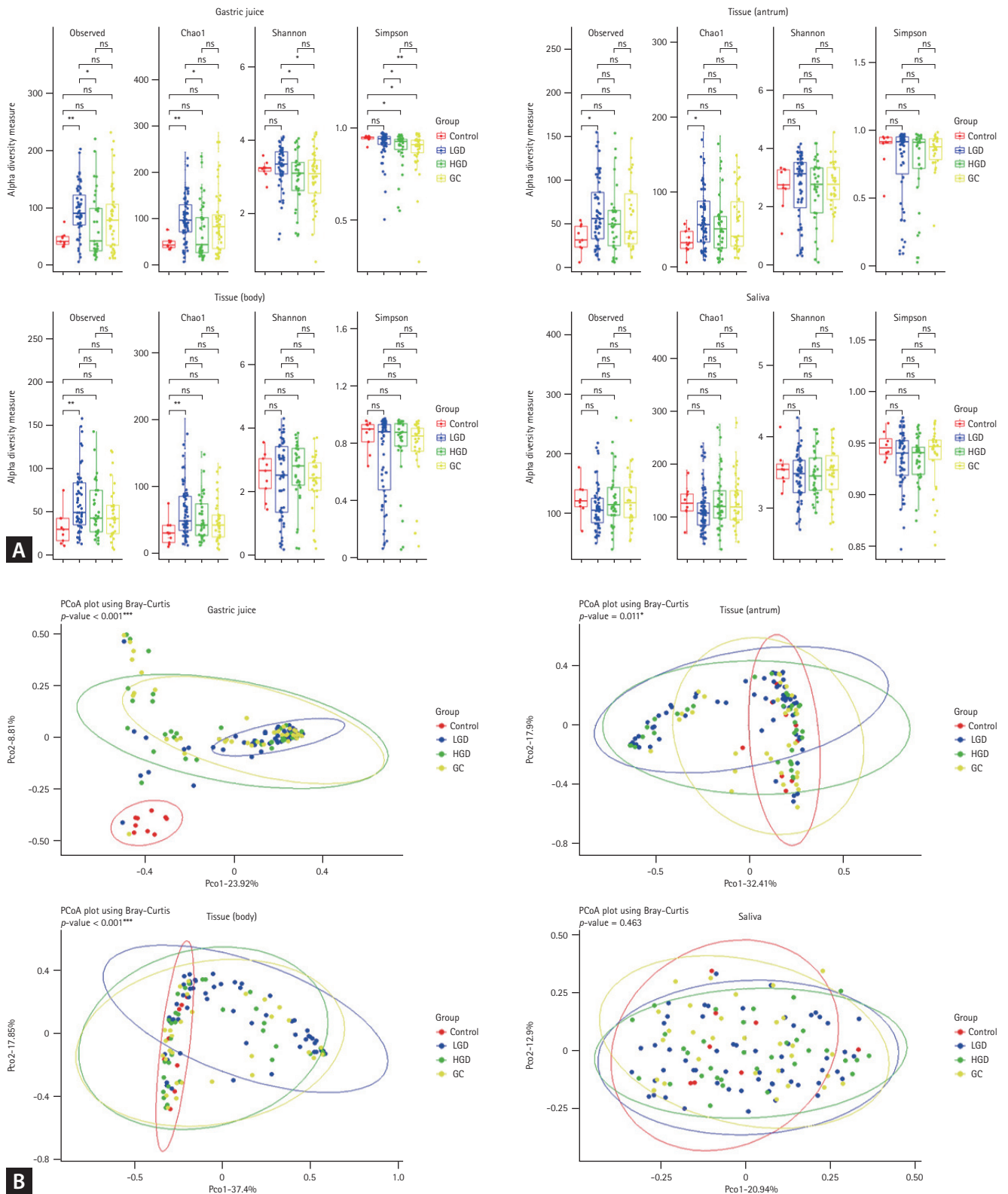M, male; F, female.

Figure 1. The alpha diversity (A) and beta diversity (B, C) at each stage of gastric carcinogenesis in the gastric juice, gastric tissue (antrum, body), and saliva samples. The alpha diversity was analyzed using the observed Chao1, Shannon, and Simpson indices, while the beta diversity was assessed using the Bray–Curtis method. The disease stages were categorized as control, low-grade dysplasia (LGD), high-grade dysplasia (HGD), and gastric cancer (GC). *p < 0.05, **p < 0.01, ***p < 0.001.
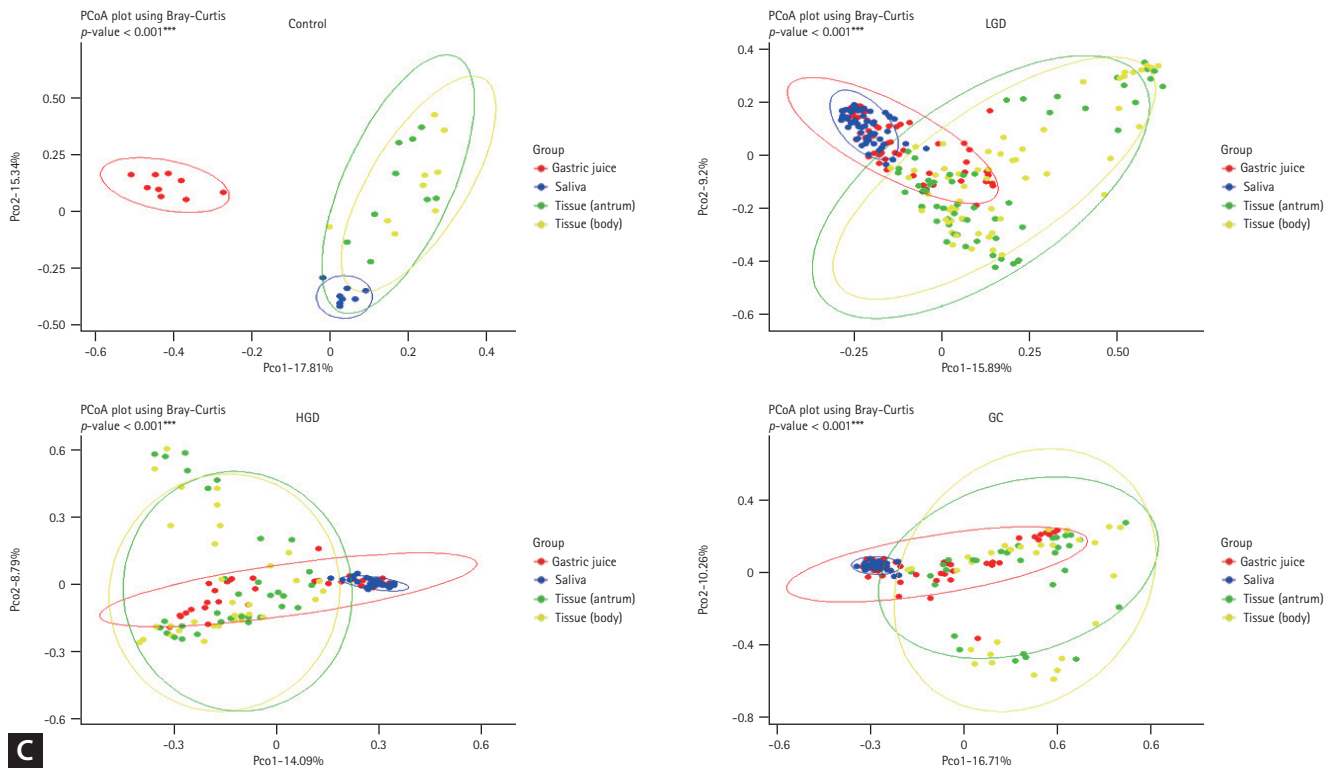
Figure 1. Continued

max-ee-f 3, max-ee-r 3) [10]. Taxonomic classification was performed using a Naïve Bayes classifier trained on the extracted V3-V4 region from the SILVA 138 database. All sequences classified as either chloroplasts or mitochondria were removed.

Statistical analysis

For alpha diversity analysis, samples were rarefied to the minimum read number (1,969) to normalize the read counts. Alpha diversity was further examined by adjusting for age and sex using analysis of variance (ANOVA). Principal coordinate analysis (PCoA) was conducted to determine the individual taxa-level clustering of groups based on the Bray–Curtis dissimilarity distance and weighted UniFrac distance. The p value for the PCoA was calculated by permutational multivariate analysis of variance (PERMANOVA) using dissimilarity matrices. Additionally, we verified the PCoA after adjusting for age and sex using PERMANOVA. To analyze the differences in microbiome composition and functional profiles between groups, the Wilcoxon rank-sum test was performed. Linear discriminant analysis (LDA) Effect Size (LEfSe) was also used to determine the differentially abundant taxa between the groups for the selection of

biomarkers with statistical and biological significance. When converting the feature table to LEfSe format, 100,000 was used as the normalization value. The LEfSe algorithm utilized the Kruskal–Wallis test with a cutoff LDA score (log10) of 3.5 for bacteria and 0.25 for functional profiles. During the LEfSe analysis, the false discovery rate (FDR) was applied to further validate the p values. This value represents the q-value and was examined to determine the significance of the results and assess false positives. Additionally, an analysis was conducted using thse multivariate association with linear models (MaAsLin2) package in R, adjusting for age and sex as covariates. A q-value of less than 0.25, adjusted for the FDR using the Benjamini–Hochberg method, was considered statistically significant. This threshold is the default significance threshold in MaAsLin2, and is commonly used in microbiome studies [11]. To reduce noise, taxa present in less than 10% of all samples with a mean relative abundance of less than 0.01% were removed before conducting LEfSe and MaAsLin2 analyses. For the analysis of the relative abundance and functional profile prediction, the Total-Sum Scaling method was used for normalization. Spearman's rank correlation coefficient was used to analyze the correlation between the microbiota compositions of all
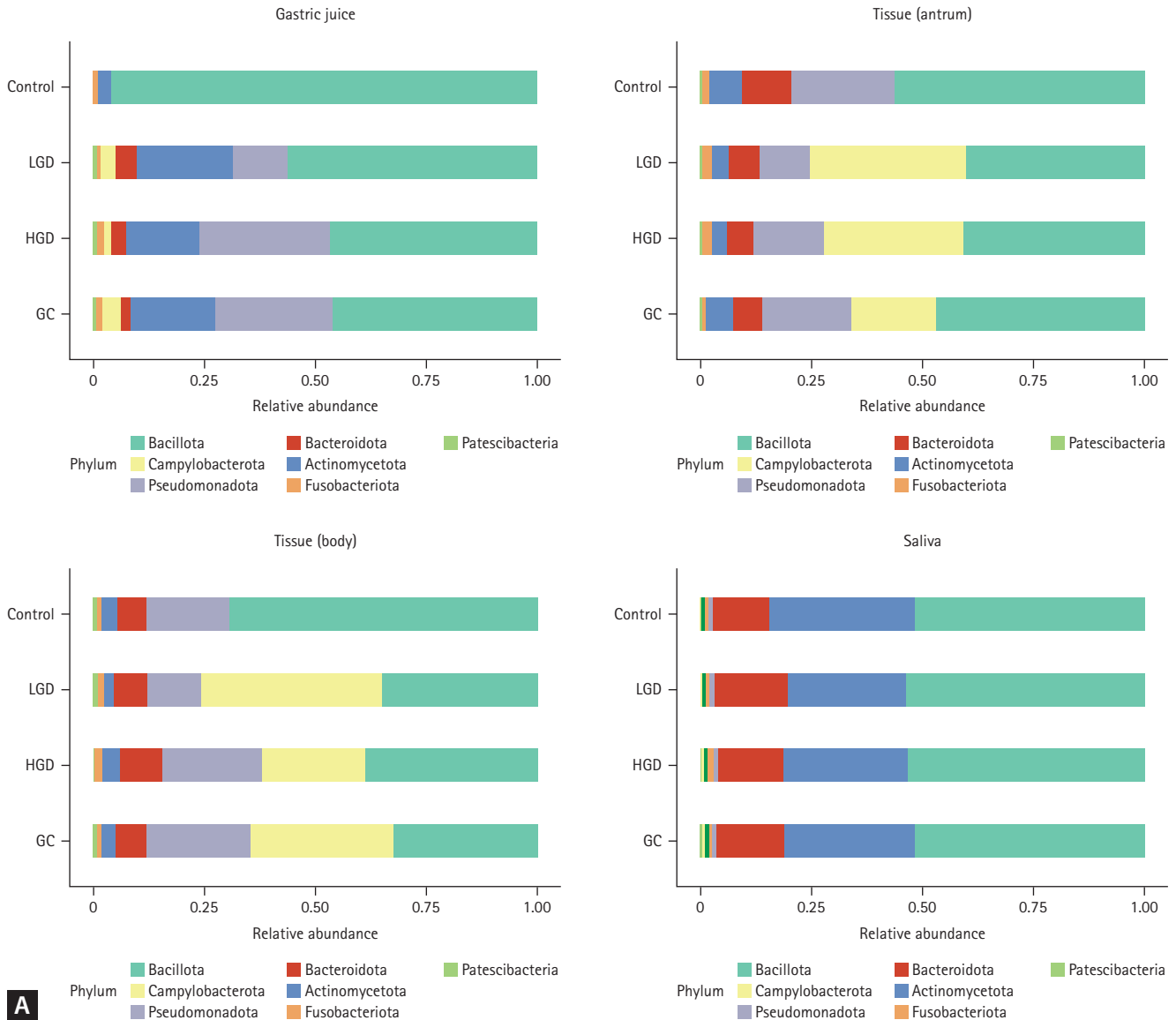
Figure 2. Microbial composition and linear discriminant analysis Effect Size (LEfSe) analysis across gastric cancer stages for the gastric juice, gastric tissue (antrum, body), and saliva samples. (A) The microbial composition changes at the phylum level in response to gastric carcinogenesis in the tissues (antrum and body) and saliva. (B) LEfSe analysis employing a cutoff linear discriminant analysis (LDA) score of 3.5 or higher for further analysis. After using the non-parametric Kruskal–Wallis test with a significance threshold (FDR-adjusted p value; q-value) of 0.05, we selected significant discriminative groups using the LDA within the LEfSe algorithm. Subsequently, data normalization for LEfSe analysis was conducted through relative log expression. Microbial taxa with high proportions in the control, low-grade dysplasia (LGD), high-grade dysplasia (HGD), and gastric cancer (GC) groups were color-coded as follows: red > orange > light blue > blue. The indicated strains have a value of $p < 0.05$, i.e., Pseudomonadota (formerly Proteobacteria), Bacillota (formerly Firmicutes), and Actinomycetota (formerly Actinobacteria).

samples. Spearman's rho and p values were calculated using the cor.test in R [12]. The analyzed correlation coefficient value (rs) was found to be statistically significant when the p value was < 0.05, indicating a meaningful relationship. We performed multiple testing corrections for the Wilcoxon rank-sum test using the Benjamini–Hochberg procedure to

control for the FDR. The functional profiles of the gastric and oral microbiomes of the control and each disease stage were predicted using Tax4Fun2 [7]. All analyses were conducted using R Statistical Software (version 3.6.1; R Foundation for Statistical Computing, Vienna, Austria).
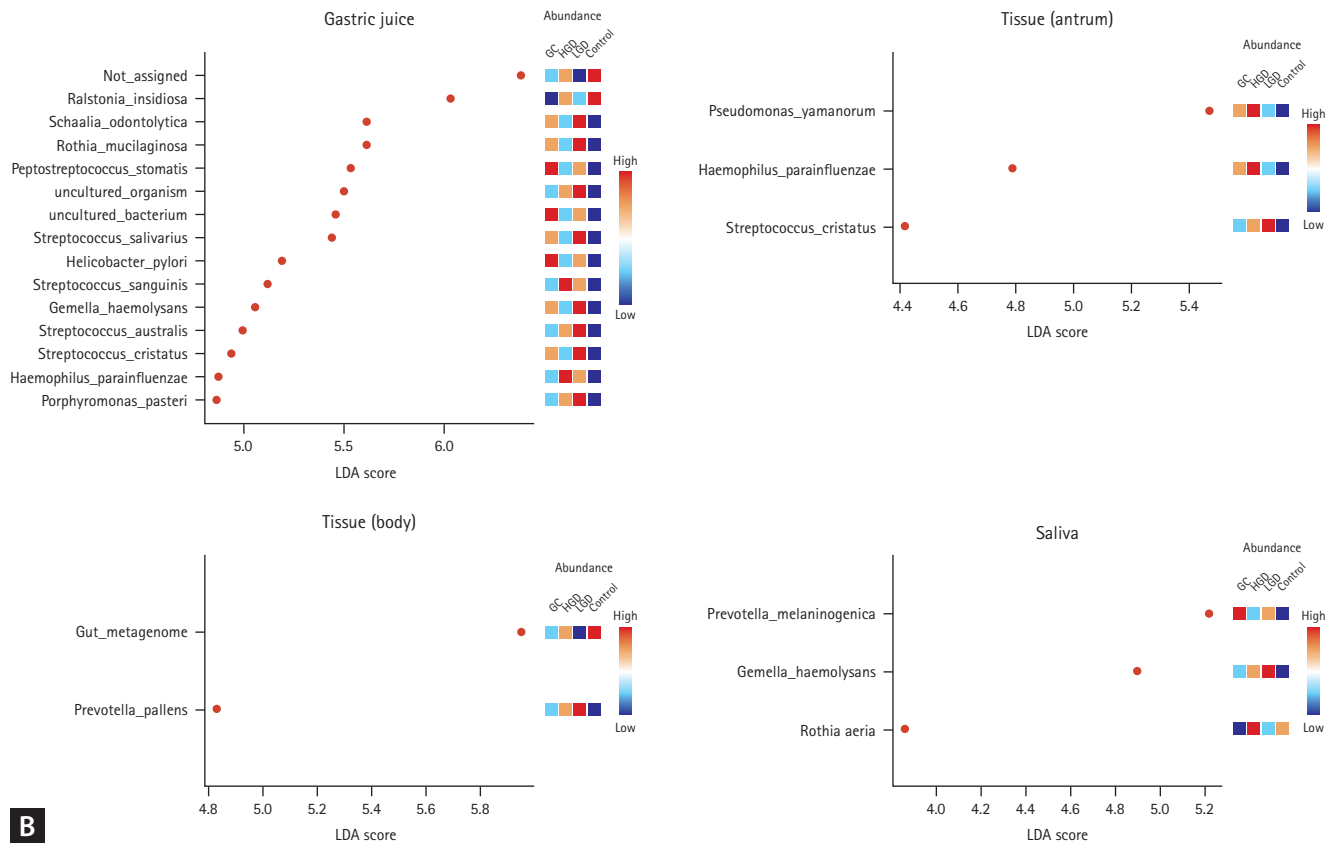
B

Figure 2. Continued

RESULTS

Clinical characteristics of the patients

Samples were collected from 141 patients (nine controls, 58 with LGD, 33 with HGD, and 41 with GC). The study participants had a mean age of 64.4 ± 11.9 years, with a male proportion of 70.9% (100/141). The clinical characteristics of the enrolled patients according to each sample type are summarized in Table 1.

Changes in the diversity of gastric and oral microbiota during gastric carcinogenesis

We analyzed the microbial diversity in the microbiota of the gastric juice among the four groups. In the alpha diversity analysis of the gastric juice, the Simpson indices were significantly lower in the HGD and GC groups than in the LGD and control groups ($p < 0.05$; Fig. 1A). In contrast, there was no significant difference in the alpha diversity of the microbial composition within the gastric tissue (both antrum and body) and oral samples among the four groups.

The beta diversity was assessed between the different stages of GC in the gastric juice, tissue (antrum, body), and saliva samples (Fig. 1B). A significant difference in microbiota composition was confirmed among the four groups of gastric juice samples, and the clustering distinction of the control group was particularly prominent ($p < 0.001$; Fig. 1B). Regarding the beta diversity, there was also a significant difference in the microbial composition of the gastric tissue among the four groups (antrum, $p < 0.05$; body, $p < 0.001$; Fig. 1B). Moreover, no significant differences were found in the beta diversity of the oral microbiota among the four groups ($p = 0.463$, Fig. 1B).

After adjusting for age and sex, the alpha and beta diversity analyses exhibited numerical fluctuations in certain p-values; however, the overall significance remained consistent both before and after the adjustment. Additionally, to clearly demonstrate the differences in beta diversity based on location corresponding to each stage of gastric carcinogenesis, schematic diagrams were constructed (Fig. 1C). Significant differences were observed in microbi-

You HS, et al. UGI tract micobiota in gastric cancer

KJIM✕

ome composition among the different sample types at all stages of gastric carcinogenesis. In particular, the difference in microbial composition between the saliva and gastric juice was more pronounced in the control group than in the disease groups (LGD, HGD, and GC groups).

Changes in the gastric and oral microbiota composition during gastric carcinogenesis

A difference was observed in the gastric juice microbial community between the control and disease groups (Fig. 2A). The abundance of Pseudomonadota (formerly Proteobacteria) decreased, while that of Bacillota (formerly Firmicutes) and Actinomycetota (formerly Actinobacteria) increased in the gastric juice of the disease group compared to that of the control group. Moreover, a difference in the microbial composition of the tissues between the control and disease groups (both the antrum and the body) was observed. In the tissues, the abundance of Campylobacterota increased in the disease groups compared to that in the control group. In contrast, no significant difference in the oral microbial composition between the control and disease groups was observed. In the saliva, Bacillota, Actinomycetota, and Bacteroidota predominated in both the control and disease groups, and the composition fractions of each group were very similar, with few differences between the groups.

The LEfSe results for gastric juice samples confirmed significant differences in the abundance of specific strains, including *Ralstonia insidiosa*, *Schaalia odontolytica*, *Rothia mucilaginosa*, *Peptostreptococcus stomatis*, *Streptococcus salivarius*, *H. pylori*, *Streptococcus sanguinis*, *Gemella haemolysans*, *Streptococcus australis*, *Streptococcus cristatus*, *Haemophilus parainfluenza,* and *Prophyromonas pasteri*, among others, within each group (Fig. 2B, q < 0.05). Among these characteristic bacteria, the abundance of *R. insidiosa* decreased in the disease groups compared to the control group (q < 0.01). However, the abundance of other strains increased in the disease groups than in the control group, with the highest proportion observed in the LGD group (q < 0.05). In the gastric fluid sample analysis, adjusting for sex and age resulted in a loss of significance for some bacterial strains. However, the strains *S. odontolytica*, *P. stomatis*, and *S. cristatus* generally maintained their significance even after adjustment for age and sex (Supplementary Table 1, 2). In the tissues (antrum), the abundance of *Pseudomonas yamanorum*, *H. parainfluenza*, and *S. cristatus* differed significantly between each disease group

(q < 0.05). Among the three bacterial species, *P. yamanorum* and *H. parainfluenza* were more abundant in the HGD group than in the control group, whereas *S. cristatus* was the most abundant in the LGD group. In the body tissues, the abundance of *Prevotella pallens* significantly differed in each disease group (q < 0.05). The abundance increased in the disease group compared to the control group, with the highest observed in the LGD group. In the oral samples, the abundance of *Prevotella melanionogenica*, *G. haemolysans*, and *Rothia aeria* showed significant differences among the disease groups (q < 0.05), with *G. haemolysans* exhibiting a significantly lower proportion in the control group than in the other disease groups (q < 0.05). Moreover, the bacterial strains exhibited significant changes (q < 0.05) following the LEfSe analysis (Supplementary Table 3-8).

Increased similarity between the oral and gastric microbiotas during gastric carcinogenesis

The degree of similarity in bacterial composition among the oral, gastric juice, and gastric tissue samples within each indi-
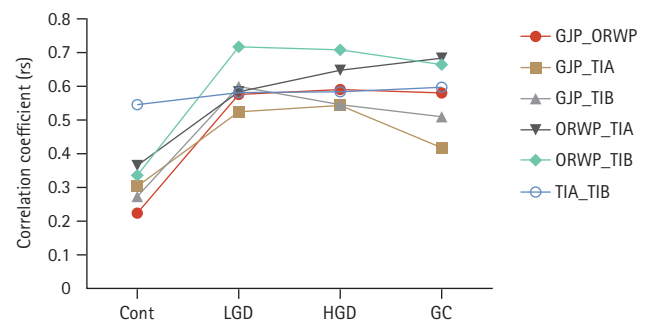

Figure 3. Correlation of the microbial composition among different sample types and disease groups. The composition and ratio of strains of each microbial composition in the oral cavity, gastric juice, gastric antrum tissue, and gastric body tissue were confirmed by analyzing the similarity in Spearman's rank correlation according to the disease stage between each sample. The graph illustrates changes in the degree of similarity based on gastric carcinogenesis. This enabled us to determine how the similarity levels among gastric juice, gastric tissue, and saliva changed in relation to gastric carcinogenesis. The similarity between the samples increased as the disease progressed. Using the Spearman rank test, the correlation coefficient for the relative microbial composition between the two types of specimens during gastric carcinogenesis was calculated. Spearman's rho and *p* value were calculated using the cor.test in R. All results displayed in the graph indicate significant positive correlations (correlation coefficient > 0.2; *p* < 0.05). GJP, gastric juice; ORWP, saliva; TIA, tissue (antrum); TIB, tissue (body); LGD, low-grade dysplasia; HGD, high-grade dysplasia; GC, gastric cancer.
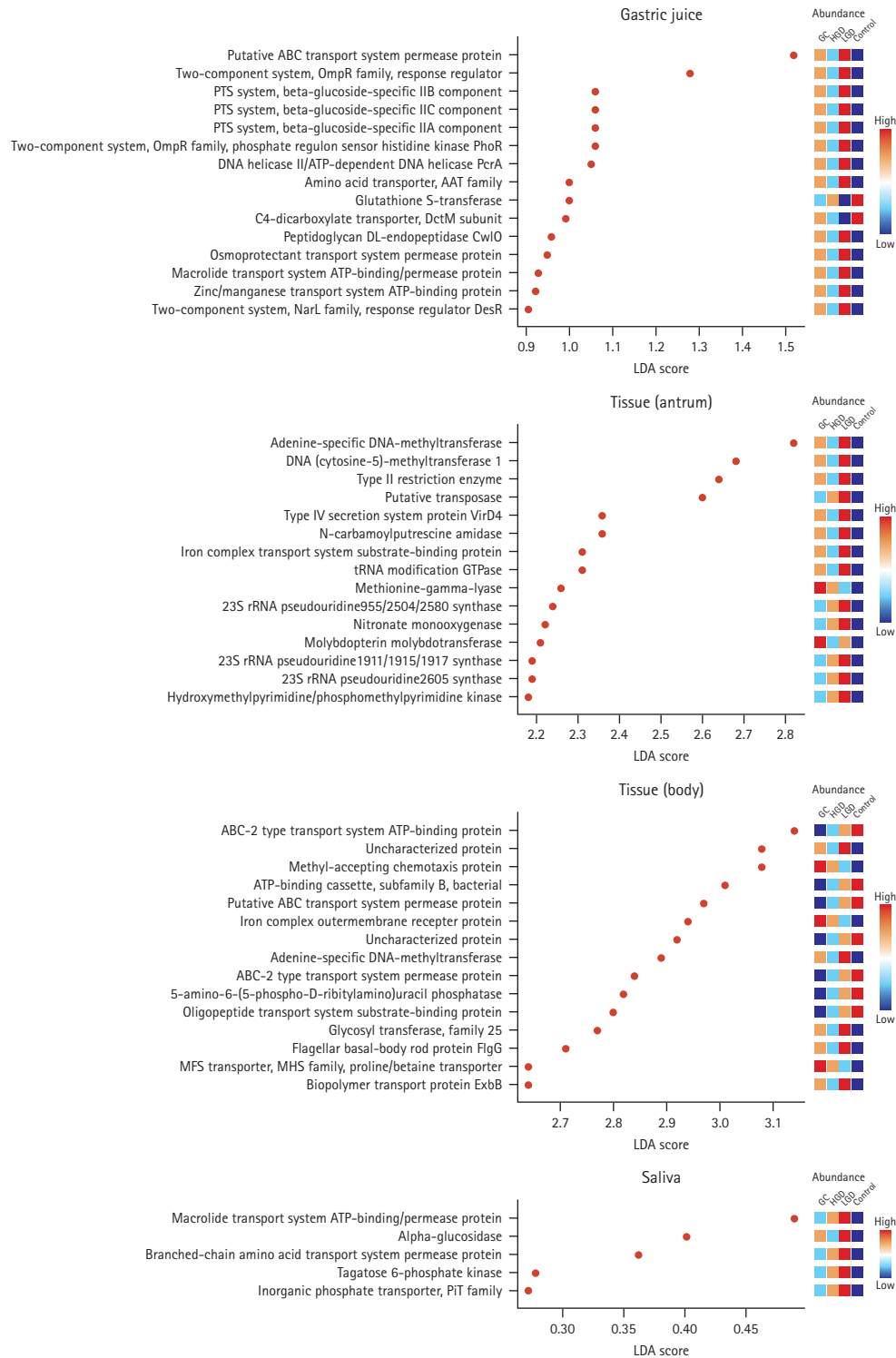
Figure 4. Metagenomic analysis reveals functional differences among the different sample types and disease groups. Linear discriminant analysis Effect Size (LEfSe) analysis identified functional profiles that exhibited changes with a linear discriminant analysis (LDA) score cut-off of 0.25 in response to gastric carcinogenesis. After employing the non-parametric Kruskal–Wallis test with a significance threshold (FDR-adjusted *p* value; q-value) of 0.05, we selected significant discriminative groups using the LDA within the LEfSe algorithm. Subsequently, data normalization for LEfSe analysis was conducted using relative log expression. Functional profiles with high proportions in the control, low-grade dysplasia (LGD), high-grade dysplasia (HGD), and gastric cancer (GC) groups were color-coded as follows: red > orange > light blue > blue.

You HS, et al. UGI tract micobiota in gastric cancer

KJIM

vidual disease group was assessed using correlation analysis (Fig. 3, Supplementary Table 9). The similarity between the samples increased as the disease progressed. Moreover, the similarity in microbial composition was relatively low among the different sample types in the control group. The similarity in microbial composition between the saliva and gastric juice samples increased from 0.22 in the control group to 0.58, 0.60, and 0.58 in the LGD, HGD, and GC groups, respectively. Similar tendencies were also found between the gastric juice and tissues (antrum and body), as well as between the saliva and tissues (antrum and body), with higher similarity of microbial composition in the GC groups than in the control group (0.42–0.51 vs. 0.30–0.27 and 0.68–0.66 vs. 0.36–0.34, respectively). Meanwhile, the microbial similarity between the antrum and body tissues consistently displayed a stable level of similarity regardless of the disease status. Furthermore, there was a tendency for an increase in the similarity between samples with respect to the stages of gastric carcinogenesis. However, some samples showed a slight decrease in similarity as they progressed from the LGD group, which exhibited the highest similarity, to the HGD and GC groups (GJP-TIA, GJP-TIB, and ORWP-TIB).

Functional profile analysis of the microbiome

When comparing the functional profiles based on microbiome clustering in four distinct samples across the disease stages, we anticipated that changes would primarily occur in signal transduction (Gastric juice: putative ABC transport system permease protein - K02004, two-component system, OmpR family, response regulator - K02483, PTS system, beta-glucoside-specific IIB component - K02756, PTS system, beta-glucoside-specific IIC component - K02757, PTS system, beta-glucoside-specific IIA component - K02755, two-component system, OmpR family, phosphate regulon sensor histidine kinase PhoR - K07636, two-component system, NarL family, response regulator DesR - K07693, and two-component system, chemotaxis family, chemotaxis protein CheY - K03413; Body tissue: ABC-2 type transport system ATP-binding protein - K01990 and ABC-2 type transport system permease protein - K01992), DNA replication and expression (Gastric juice: glycine cleavage system transcriptional repressor - K03567; Antrum tissue: adenine-specific DNA-methyltransferase - K07316 and DNA (cytosine-5)-methyltransferase 1 - K00558; Body tissue: adenine-specific DNA-methyltransferase - K07316), and ATP-related pathways (Gastric juice: macrolide transport

system ATP-binding/permease protein - K18230 and zinc/manganese transport system ATP-binding protein - K02074; Body tissue: ATP-binding cassette, subfamily B, bacterial - K06147; Saliva: macrolide transport system ATP-binding/permease protein - K05685) (Fig. 4, q < 0.05). The functional profiles exhibited significant changes (q < 0.05) in the LEfSe analysis, and this significance was maintained by the application of q-values (Supplementary Table 10-13).

DISCUSSION

In this study, we investigated changes in microbial composition and function in the oral cavity and stomach throughout the gastric carcinogenesis pathway. We observed differences in bacterial diversity and composition between the disease and control groups. Furthermore, changes in the metagenomic functional profiles potentially associated with carcinogenesis were observed. We also found that the bacterial composition in the oral cavity and stomach gradually became more similar during gastric carcinogenesis.

Similar to previous studies, the GC group exhibited a decrease in microbial diversity in the gastric juice compared with the control group, with a consistent trend of dominance by specific minority species [5]. The decrease in microbial diversity in various diseases compared to that in the healthy state is well established [13]. Moreover, the beta diversity was significantly different between the healthy and disease groups in the microbial communities of the gastric juice and tissues, indicating variations in the intergroup microbial composition. Specifically, at the phylum level, there was a decrease in the abundance of Pseudomonadota and an increase in that of Bacillota in the gastric juice with disease progression. Moreover, certain species of bacteria showed differences between the control and disease groups. At the phylum level, the abundance of Campylobacterota increased in the disease group compared to the control group. The results observed at the phylum level are consistent with those of previous studies on microbial clustering in GC [14]. Among the specific bacterial species whose abundance increased in the disease group compared to the control group, some have been widely reported to be associated with cancer. *S. odontolytica* has been identified as a potential pathogen capable of triggering malignant oral conditions [15]. It has also been reported to be associated with higher cancer prevalence [16]. Similar to *S. odontolyt-*

ica, which is commonly found in the oral cavity, *S. cristatus* and *P. stomatis* are associated with oral cancer [17-19]. Furthermore, *P. stomatis* has been identified as a potent potential carcinogenesis biomarker not only for oral cancer but also for colorectal cancer [20,21]. In our study, we observed that the abundance of these strains increased in the stomach during gastric carcinogenesis, raising the suspicion that these strains may play a significant role in GC development.

Inspired by the increasing trend of oral microbial strains in the disease group, we further analyzed the correlation between the oral and gastric microbiomes. We compared and analyzed the characteristics of the microbial composition of samples obtained from different anatomical locations in the upper gastrointestinal tract of the patient and control groups. Disparities in the correlations between microbial compositions based on anatomical location were observed between the disease and control groups. In the control group, the oral microbial composition showed relatively low similarity with that of the gastric juice and tissue samples. In contrast, in the disease group, there was an increase in the similarity between the oral and gastric samples, as well as between the gastric juice and tissues. Considering the absence of significant differences in the oral microbial composition among the groups, we suspect that differences in the gastric environment might have altered the process by which oral bacteria influence the microbial composition upon entering the stomach. The oral cavity serves as a continuous route for bacterial influx into the upper GI tract. The bacterial population in the esophagus originates from the oral cavity, and oral bacteria may influence the development of esophageal cancer. Similarly, it can be inferred that oral bacteria can significantly impact the gastric microbial composition [22-25].

Infection with *H. pylori*, a well-recognized risk factor for GC, leads to persistent chronic inflammation of the gastric mucosa, eventually causing mucosal atrophy and worsening atrophic gastritis. As atrophic gastritis progresses, gastric acid secretion decreases, which can substantially affect the gastric microbial composition [26]. Bacterial strains that would typically struggle to thrive in the highly acidic environment of the stomach could proliferate more easily as the acidity decreases. This could also account for the increased similarity between the floating bacteria in the gastric juice and the clustered bacteria adhering to the gastric mucosa in the stomach of patients with gastric neoplasms. Previous studies have demonstrated this phenomenon, and the altered gastric microbial composition in patients taking proton pump inhibitors further supports this observation. Therefore, although a significant proportion of orally ingested microbial strains may persist in a healthy stomach, these strains may more readily establish clusters and microbial compositions within the stomach as gastric carcinogenesis progresses, coupled with reduced gastric acid secretion. Recent studies have actively explored the potential impacts of non-*H. pylori* gastric bacteria on GC development, which aligns with this hypothesis.

Consequently, additional analyses were conducted to analyze the functional profile prediction and assess the relevance of the oral-gastric microbiome cluster. The functional profiles of microorganisms known to be associated with cancer, such as the ATP-binding cassette transport system, two-component system, and phosphoenolpyruvate-carbohydrate phosphotransferase system, showed significant changes in the LEfSe analysis when compared between the disease and control groups. These profiles were higher in the disease group than in the control group, suggesting a positive correlation, similar to the observed increase in *P. stomatis* and *S. cristatus* with disease progression. This indicates that increased microbial metabolism may facilitate the survival and growth of these organisms. The aforementioned functional profiles, such as the ABC transport system and two-component system, operate in bacteria. The ABC transport system can also function in eukaryotic cells, aiding nutrient uptake, toxin secretion, and sometimes chemotherapy drug resistance in tumor cells [27-29]. The two-component system serves as a central player in bacterial signal transduction, catalyzing DNA binding, and biochemical reactions. It is involved in various physiological functions related to survival, such as sporulation and antibiotic resistance, making it play a role in inducing cancer [30,31].

Our study had some limitations. In some sample groups, there was an imbalance in the distribution of sex and age, with a bias towards one side. Consequently, we included groups with sex and age imbalances in the analysis. After conducting the analysis, additional verification was performed after adjusting for age and sex. Although there were no significant changes in alpha and beta diversity, some bacterial strains lost significance in the LEfSe results when adjusted for sex and age. However, in gastric juice samples, the significance of *S. odontolytica*, *P. stomatis*, and *S. cristatus* strains was maintained. This may have influenced the results of the metagenomic analyses between

You HS, et al. UGI tract micobiota in gastric cancer

KJIM

the control and disease groups. In the analysis of functional profiles, we conducted a predictive analysis rather than an actual functional analysis. The current body of evidence often lacks comprehensive analyses of both structural and functional aspects of microbial communities. Therefore, it is not possible to conclusively demonstrate the precise role and mechanism of microbial communities in GC development. Despite these limitations, our study revealed some interesting findings.

Our findings suggest that compositional changes occur in the gastric microbiome, accompanied by functional implications during carcinogenesis. This indicates that microbes other than *H. pylori*, a prominent contributor to GC, play a role in altering the balance between the host and commensal microbial flora within the stomach. The increased similarity between the microbiota of the oral cavity and stomach implies that the oral microbiome might play a role in carcinogenesis, sustaining the inflammatory process and heightening the risk of GC development through its distinctive genotoxic potential.

KEY MESSAGE

1. The correlation between oral and gastric microbiota increases during gastric carcinogenesis, implying a possible link between oral bacteria and GC.
2. The abundance of certain specific oral commensals, known to be associated with cancer development, was increased in the gastric juice of patients with LGD and GC compared to controls.
3. Functional profile analysis of the microbial composition of the gastric juice further supported the potential role of the microbiome in gastric carcinogenesis.

REFERENCES

1. 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.
2. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Schistosomes, liver flukes and *Helicobacter pylori*. Lyon: International Agency for Research on Cancer, 1994

3. Correa P, Shiao YH. Phenotypic and genotypic events in gastric carcinogenesis. Cancer Res 1994;54(7 Suppl):1941s-1943s.
4. Eun CS, Kim BK, Han DS, et al. Differences in gastric mucosal microbiota profiling in patients with chronic gastritis, intestinal metaplasia, and gastric cancer using pyrosequencing methods. Helicobacter 2014;19:407-416.
5. Park JY, Seo H, Kang CS, et al. Dysbiotic change in gastric microbiome and its functional implication in gastric carcinogenesis. Sci Rep 2022;12:4285.
6. Zhang X, Li C, Cao W, Zhang Z. Alterations of gastric microbiota in gastric cancer and precancerous stages. Front Cell Infect Microbiol 2021;11:559148.
7. Wemheuer F, Taylor JA, Daniel R, et al. Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. Environ Microbiome 2020;15:11.
8. Schlemper RJ, Riddell RH, Kato Y, et al. The Vienna classification of gastrointestinal epithelial neoplasia. Gut 2000;47:251-255.
9. 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.
10. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 2016;13:581-583.
11. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol 2021;17:e1009442.
12. R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2011.
13. Mosca A, Leclerc M, Hugot JP. Gut microbiota diversity and human diseases: should we reintroduce key predators in our ecosystem? Front Microbiol 2016;7:455.
14. Liu D, Zhang R, Chen S, Sun B, Zhang K. Analysis of gastric microbiome reveals three distinctive microbial communities associated with the occurrence of gastric cancer. BMC Microbiol 2022;22:184.
15. Herreros-Pomares A, Hervás D, Bagan-Debón L, Jantus-Lewintre E, Gimeno-Cardona C, Bagan J. On the oral microbiome of oral potentially malignant and malignant disorders: dysbiosis, loss of diversity, and pathogens enrichment. Int J Mol Sci 2023;24:3466.
16. Huang X, Chen C, Xie W, et al. Metagenomic analysis of intratumoral microbiome linking to response to neoadjuvant chemoradiotherapy in rectal cancer. Int J Radiat Oncol Biol

Phys 2023;117:1255-1269.

17. Pignatelli P, Romei FM, Bondi D, Giuliani M, Piattelli A, Curia MC. Microbiota and oral cancer as a complex and dynamic microenvironment: a narrative review from etiology to prognosis. Int J Mol Sci 2022;23:8323.

18. Zhang M, Zhang Y, Han Y, Zhao X, Sun Y. Characteristics of pathogenic microbes in lung microenvironment of lung cancer patients without respiratory infection. J BUON 2021;26:1862-1870.

19. Chattopadhyay I, Verma M, Panda M. Role of oral microbiome signatures in diagnosis and prognosis of oral cancer. Technol Cancer Res Treat 2019;18:1533033819867354.

20. Zhang L, Liu Y, Zheng HJ, Zhang CP. The oral microbiota may have influence on oral cancer. Front Cell Infect Microbiol 2020;9:476.

21. Osman MA, Neoh HM, Ab Mutalib NS, et al. *Parvimonas micra*, *Peptostreptococcus stomatis*, *Fusobacterium nucleatum* and *Akkermansia muciniphila* as a four-bacteria biomarker panel of colorectal cancer. Sci Rep 2021;11:2925.

22. Ahn J, Chen CY, Hayes RB. Oral microbiome and oral and gastrointestinal cancer risk. Cancer Causes Control 2012;23:399-404.

23. Sun JH, Li XL, Yin J, Li YH, Hou BX, Zhang Z. A screening method for gastric cancer by oral microbiome detection. Oncol Rep 2018;39:2217-2224.

24. Olsen I, Yamazaki K. Can oral bacteria affect the microbiome of the gut? J Oral Microbiol 2019;11:1586422.

25. Zhang Y, Niu Q, Fan W, Huang F, He H. Oral microbiota and gastrointestinal cancer. Onco Targets Ther 2019;12:4721-4728.

26. Parck CH. Changes in gastric microbial composition before and after *Helicobacter pylori* eradication therapy. Korean J Helicobacter Up Gastrointest Res 2020;20:177-186.

27. Domenichini A, Adamska A, Falasca M. ABC transporters as cancer drivers: potential functions in cancer development. Biochim Biophys Acta Gen Subj 2019;1863:52-60.

28. Davidson AL, Chen J. ATP-binding cassette transporters in bacteria. Annu Rev Biochem 2004;73:241-268.

29. Dawson RJ, Locher KP. Structure of a bacterial multidrug ABC transporter. Nature 2006;443:180-185.

30. Boroughs LK, DeBerardinis RJ. Metabolic pathways promoting cancer cell survival and growth. Nat Cell Biol 2015;17:351-359.

31. Mitrophanov AY, Groisman EA. Signal integration in bacterial two-component regulatory systems. Genes Dev 2008;22:2601-2611.

Correspondence to
Jae Gyu Kim, M.D., Ph.D.
Department of Internal Medicine, Chung-Ang University College of Medicine, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Korea
Tel: +82-2-6299-3147, Fax: +82-2-749-9150
E-mail: jgkimd@cau.ac.kr
https://orcid.org/0000-0002-4841-9404

CRedit authorship contributions
Hee Sang You: methodology, data curation, formal analysis, validation, software, writing - original draft, writing - review & editing, visualization; Jae Yong Park: resources, data curation, validation, writing - original draft, writing - review & editing, visualization, funding acquisition; Hochan Seo: methodology, investigation, data curation, formal analysis, visualization; Beom Jin Kim: formal analysis, validation, writing - review & editing; Jae Gyu Kim: conceptualization, writing - review & editing, supervision, project administration, funding acquisition

Conflicts of interest
The authors disclose no conflicts.