Evaluation of gastric microbiome and metagenomic function in patients with intestinal metaplasia using 16S rRNA gene sequencing

Chan Hyuk Park

Department of Medicine

The Graduate School, Yonsei University

Evaluation of gastric microbiome and metagenomic function in patients with intestinal metaplasia using 16S rRNA gene sequencing

Directed by Professor Sang Kil Lee

Doctoral Dissertation submitted to the Department of Medicine, the Graduate School of Yonsei University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Chan Hyuk Park

December 2017

연세대학교
YONSEI UNIVERSITY

This certifies that the Doctoral Dissertation of
Chan Hyuk Park is approved.

Thesis Supervisor : Sang Kil Lee

Thesis Committee Member#1 : Young Hoon Youn

Thesis Committee Member#2 : Chang Soo Eun

Thesis Committee Member#3 : Sung Jae Shin

Thesis Committee Member#4 : Ji-Hwan Ryu

The Graduate School
Yonsei University

December 2017

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Prof. Sang Kil Lee of the Graduate School at Yonsei University. He has led me to be interested in gastroenterology and begin research. The door to Prof. Lee office was always open whenever I ran into a trouble spot or had a question about my research.

I would also like to thank Prof. Dong Soo Han at Hanyang University College of Medicine who gave me great inspiration for this research. He gave me experience in microbiome research and instilled a sense of challenge.

I would also like to thank Prof. Chang Soo Eun at Hanyang University College of Medicine. He always supported me to adapt to the new lab and staring my own research. Without his passionate help and advice, my research could not have been successfully conducted.

I would also like to acknowledge A-reum Lee and Yu-ra Lee, researchers at Prof. Han's lab. I am gratefully indebted to their valuable comments on the experimental techniques.

Finally, I must express my very profound gratitude to my parents and to my spouse for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Author
Chan Hyuk Park

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Evaluation of gastric microbiome and metagenomic function in patients with intestinal metaplasia using 16S rRNA gene sequencing

Chan Hyuk Park

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Sang Kil Lee)

Despite recent advances in studies of the gastric microbiome, a specific role of non-*Helicobacter pylori* (HP) gastric microbiomes in gastric carcinogenesis has not been fully elucidated. The study aim was to identify potential functions of gastric microbiomes in gastric carcinogenesis by using metagenomic analysis. Participants were classified into six groups according to disease status (chronic superficial gastritis [CSG], intestinal metaplasia [IM], and cancer) and HP infection status (HP-positive and HP-negative). Gastric microbiomes were analyzed with mucosal tissues at the gastric antrum by using 16S rRNA gene sequencing. Metagenomic analysis was performed using the phylogenetic investigation of communities by reconstruction of unobserved states. Among 138 included patients, 48, 9, 23, 14, 12, and 32 were classified as HP(-) CSG, HP(-) IM, HP(-) cancer, HP(+) CSG, HP(+) IM, and HP(+) cancer groups, respectively. *Cyanobacteria* were

predominant in the HP(-) CSG group, whereas *Rhizobiales* were commonly observed in the HP(-) IM group. In the HP(-) IM group, genes encoding type IV secretion system proteins (T4SS), especially VirB4, VirB6, and VirB9, were prevalent among the metagenomes. These genes were mainly derived from *Rhizobiales* in the HP(-) IM group. Additionally, gastric microbiomes after HP eradication therapy were similar to microbiomes observed after spontaneous HP regression.

Key words : *Helicobacter pylori*; microbiome; gastric cancer; intestinal metaplasia; type IV secretion system

Evaluation of gastric microbiome and metagenomic function in patients with intestinal metaplasia using 16S rRNA gene sequencing

Chan Hyuk Park

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Sang Kil Lee)

I. INTRODUCTION

Helicobacter pylori (HP), which was classified as a "group I carcinogen" by the World Health Organization, is a well-known risk factor for gastric cancer.[1] In regions where gastric cancer is prevalent, such as Korea and Japan, almost all cases of gastric cancer develop in patients who have been infected with HP.[2,3] One of the initiating factors of gastric carcinogenesis is cytotoxin-associated gene A (*cagA*), which is a well-known bacterial virulence factor.[4] This factor causes inflammation and activates oncogenic pathways. Excessive oxidative stress from reactive oxygen and nitrogen species production in HP-infected stomachs may induce DNA damage in the gastric epithelium.[5]

However, there are several unsolved questions regarding HP-related gastric carcinogenesis. First, HP often disappears spontaneously in elderly patients due to progression of atrophic gastritis and intestinal metaplasia (IM),[6] while gastric cancer usually develops in elderly individuals rather than younger individuals.[7] If HP were the only direct risk factor for gastric carcinogenesis, gastric cancer risk might decrease as the relative abundance of HP decreases in elderly patients. Although IM has been suggested as one of the reasons for

gastric cancer development after HP regression, it is not known whether IM itself causes gastric cancer development. Various factors including bacterial overgrowth, nitrate reduction, *N*-nitroso carcinogens, and host susceptibility have been proposed as causes of gastric cancer development;[8] however, detailed pathogenic mechanisms of these factors have not been fully evaluated. Second, HP eradication therapy alone cannot completely prevent gastric cancer, especially in patients with precancerous lesions, including IM.[9] These results imply that specific changes that initiate gastric carcinogenesis may occur in patients with IM.

The first thing to consider is a bacterial overgrowth of a species other than HP in patients with IM, because many human diseases are affected by the microbiome.[10] Previous studies suggest that various bacteria including *Clostridium*, *Staphylococcus*, and *Neisseria* may play a role in the formation of *N*-nitroso compounds, which increase the risk of gastric cancer.[11-14] Moreover, metagenomes derived from intragastric bacteria other than HP might transfer to HP. Horizontal gene transfer is an important adaptive force in evolution, contributing to metabolic, physiological, and ecological innovation in most prokaryotes and some eukaryotes.[15]

Recently, there have been several attempts to determine features of the gastric microbiome using next-generation sequencing.[16,17] Previous studies, however, assessed only microbial distribution and did not elucidate the potential role of gastric microbiomes in gastric carcinogenesis.[16,17] The first aim of the study was to evaluate the characteristics of gastric microbiomes and metagenomic functions in patients with IM. The second aim was whether there were gastric microbiome differences between patients with spontaneous HP regression and those with HP eradication via antibiotics.

II. MATERIALS AND METHODS

1. Study design

Asymptomatic volunteers or patients with dyspepsia who were scheduled for upper gastrointestinal endoscopy were recruited for this study. Exclusion criteria were as follows: (a) patients who took medications including proton pump inhibitors, H_2 receptor antagonists, antacids, probiotics, mucosal protective agents, or antibiotics, within 3 months prior to enrollment, (b) patients with a history of gastric adenoma, carcinoma, or mucosa-associated lymphoid tissue lymphoma, and (c) patients who underwent gastrectomy.

After testing for HP infection and histologic examination of gastric tissue, patients who met one of the following criteria were finally enrolled and analyzed: (1) group A [HP(-) CSG; control], patients with CSG and without HP infection, (2) group B [HP(-) IM], patients with IM and without HP infection, (3) group C [HP(-) cancer], patients with gastric cancer and without HP infection, (4) group D [HP(+) CSG], patients with CSG and HP infection, (5) group E [HP(+) IM], patients with IM and HP infection, and (6) group F [HP(+) cancer], patients with cancer and HP infection. Group A represents the patients who had mild gastritis without prior HP infection. In this study, the group A was regarded as a control group. In contrast, group B and C represent the patients with prior infection and spontaneous regression of HP.

The study protocol was approved by the Institutional Review Board on Human Subjects Research and Ethics Committees Hanyang University Guri Hospital, Korea. All patients provided written informed consent before enrollment. The study was registered at the International Clinical Trials Registry Platform, no. KCT0001602 (https://cris.nih.go.kr).

2. Tissue sampling and serologic testing

For microbiome analysis, four pieces of gastric mucosal tissue were obtained by endoscopic biopsy from the greater curvature side of the antrum.

Then, the rapid urease test was performed to evaluate HP infection status. One additional piece of mucosal tissue was biopsied for histologic examination of gastric mucosa.

For serologic assessments of HP infection and atrophic gastritis, serum IgG anti-HP antibody and pepsinogen I/II testing were performed by enzyme immunoassay and latex agglutination turbidimetric immunoassay, respectively.

3. HP eradication therapy

In the HP(+) CSG group, patients underwent HP eradication therapy with the standard triple regimen for two weeks (rabeprazole 20 mg, amoxicillin 1 g, clarithromycin 500 mg, twice daily). One month after eradication therapy, four pieces of gastric mucosal tissue in the greater curvature side of the antrum were endoscopically biopsied to analyze microbiome changes after eradication. HP eradication was confirmed by ^{13}C-urea breath test. Patients who underwent HP eradication therapy were further classified into the successful eradication and failed eradication groups, according to the results of HP eradication therapy.

4. Extraction of bacterial DNA

Extraction of bacterial DNA was performed from mucosal biopsy samples as previously described.[16] Briefly, 100 mg of frozen gastric mucosal tissues was suspended in 750 μL of sterile bacterial lysis buffer (200 mmol/L NaCl, 100 mmol/L EDTA [pH 8.0], 20 mmol/L Tris base, 20 mg/mL lysozyme) and incubated at 37°C for 30 minutes. Then, 20 μL of proteinase K and 80 μL of 10% SDS were added to the mixture followed by incubation at 65℃ for 30 minutes. Finally, bead beating was performed for 90 seconds at 5,300 rpm (PRECELLYS 24; Bertin Technologies, Le Bretonneux, France) after adding 300 mg of 0.1-mm zirconium beads (BioSpec Products, Bartlesville, OK, USA) to finish homogenization. The homogenized mixture was cooled on ice and then centrifuged at 14,000 rpm for 5 minutes. Bacterial DNA was extracted from the

supernatant by phenol/chloroform/iso-amyl alcohol (25:24:1) followed by chloroform/iso-amyl alcohol (24:1), and precipitated by absolute ethanol at -20℃ for 1 hour. The precipitated DNA was suspended in DNase-free H_2O and cleaned up using a DNA clean-up kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Isolated DNA was stored at -80℃ until use for microbial characterization.

5. 16S rRNA gene sequencing and analysis

Extracted gDNA was amplified using the following primers targeting the V3 to V4 regions of the 16S rRNA gene: forward, 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGC WGCAG-3′; reverse, 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGG TATCTAATCC-3′. High-throughput sequencing was performed on an Illumina MiSeq system (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Sequence data were processed using QIIME version 1.9.0.[18] Reads that showed low quality, contained incorrect primer sequences, or contained more than one ambiguous base were excluded. The remaining reads were classified into the groups based on their unique nucleotide barcodes. The read count was normalized to the corresponding copy number of 16S rRNA genes to reduce potential bias of abundance estimation due to copy number variation.[19] Taxonomic composition from phylum to species levels and bacterial diversity for each sample were evaluated based on 97% similarity with the GreenGenes database (version 13.5), using QIIME. Principal component analysis of bacterial abundance was performed based on the Euclidean distance to compare the microbial communities of subgroups. Differential abundant features of relative abundance of bacterial taxa or metagenomic functions were identified using the Linear Discriminant Analysis Effect Size, which is an algorithm for high-dimensional biomarker discovery and explanation that identifies genomic

features characterizing the differences between two or more biological conditions.[20]

DNA sequences obtained from this metagenomic project have been deposited in the NCBI Short Read Archive under the Accession No. SRP109017.

In patients with HP infection, additionally, *cagA* positivity was determined by semi-quantitative PCR using the following primers:[21] forward, 5′-GATAACAGGCAAGCTTTTGAGG-3′; reverse, 5′-CTGCAAAAGATTGTTTGGCAGA-3′.

6. Metagenome prediction

Functional composition of the microbiome was predicted using PICRUSt.[22] PICRUSt uses evolutionary modeling to predict metagenomes from 16S rRNA gene data compared with a reference genome database.[22] Metagenome inference based on the PICRUSt method was performed using the PICRUSt scripts (*normalize_by_copy_nubmer.py*, *predict_metagenomes.py*, and *metagenome_contributions.py*) with GreenGenes database (version 13.5).

7. Download of sequence data

To improve comparability and statistical power, microbiome sequences of gastric mucosa in East Asian patients with gastric cancer were downloaded from the NCBI (https://www.ncbi.nlm.nih.gov) following literature review. These data were analyzed in the same way as sequence data obtained from the current study.

8. Statistical analysis

Continuous and categorical variables were described as mean with standard deviation and number with proportion, respectively. Kruskal–Wallis test was used to compare bacterial or genomic abundance among groups.

Bray–Curtis distance with Wilcoxon rank-sum test was used to analyze dissimilarity among the groups. All statistical procedures were conducted using R (version 3.3.3; R Foundation for Statistical Computing, Vienna, Austria), except for the linear discriminant analysis, for which the analysis was performed using Galaxy, which is an open, web-based platform for computational biomedical research (http://huttenhower.sph.harvard.edu/galaxy).

III. RESULTS

1. Baseline characteristics and microbiome reads

Table 1 shows baseline patient characteristics and microbiome reads. Among 83 patients with chronic superficial gastritis (CSG) or IM, 48, 9, 14, and 12 were classified into the HP(-) CSG, HP(-) IM, HP(+) CSG, and HP(+) CSG groups, respectively, according to the HP testing and histologic examination results. All microbiome sequences of patients with CSG or IM were obtained and analyzed in Hanyang University Guri Hospital.

In addition, one patient in the HP(-) cancer group and two patients in the HP(+) cancer group were enrolled. To further analyze the differential characteristics of IM groups compared to cancer groups, microbiome sequences of 52 patients with gastric cancer were obtained from the National Center for Biotechnology Information (NCBI) (22 patients with HP(-) cancer, and 30 patients with HP(+) cancer; SRP038955,[16] GSE61493,[17] and SRP057951[23]). In total, sequence data from 138 patients were analyzed in this study.

Table 1. Baseline patient characteristics and microbiome reads of samples obtained from the current study

Variable	HP(-) CSG (Group A [control])	HP(-) IM (Group B)	HP(+) CSG (Group D)	HP(+) IM (Group E)
n	48	9	14	12
Age, year, mean±SD	32.3±11.0	67.4±7.7	31.4±8.8	60.9±13.4
Male, n (%)	23 (47.9)	4 (44.4)	3 (21.4)	9 (75.0)
Pepsinogen test				
Pepsinogen I, ng/mL, mean±SD	54.7±34.4	41.8±16.9	99.2±56.4	64.0±32.4
Pepsinogen II, ng/mL, mean±SD	12.0±9.7	11.4±4.8	29.1±15.4	25.5±12.7
Pepsinogen I/II ratio, mean±SD	4.9±1.5	4.4±2.3	5.3±7.7	2.5±0.9
Microbiome reads, mean±SD				
Read count	9134±5767	4542±2983	22430±5881	14948±5931
OTU	292±150	152±101	131±58	179±99
Chao1 estimator	161.6±58.8	95.2±29.5	85.2±27.8	105.7±32.9
Shannon's diversity index	3.84±0.33	3.30±0.46	0.74±0.41	1.74±1.23
Simpson's diversity index	0.95±0.03	0.93±0.02	0.22±0.14	0.51±0.35

HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia; OTU, operational taxonomic unit; SD, standard deviation

2. Relative bacterial abundance

Relative bacterial abundance at the phylum level is shown in **Figure 1**. The relative abundance of non-HP proteobacteria increased from the HP(-) CSG group, through the HP(-) IM group, to the HP(-) cancer group. In cases of patients with HP, proteobacteria (which mainly consisted of HP) decreased from the HP(+) CSG group, through the HP(+) IM group, to the HP(+) cancer group.

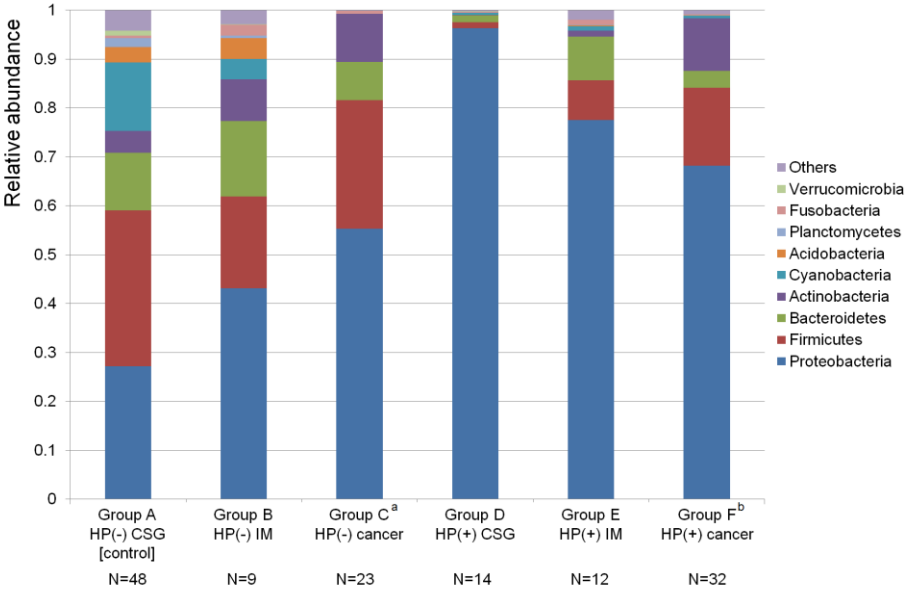

Figure 1. Relative abundance at the phylum level. [a]Sequence data in 22 of 23 patients were downloaded from the National Center for Biotechnology Information (SRP038955 and GSE61493). [b]Sequence data in 30 of 32 patients were downloaded from the National Center for Biotechnology Information (SRP038955, GSE61493, and SRP057951). HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia.

To compare microbiomes other than HP with disease status (CSG vs. IM vs. cancer), the relative bacterial abundance of the HP-negative groups were demonstrated by using cladograms **(Figure 2)**. Characteristically, the relative abundance of *Cyanobacteria* was high in the HP(-) CSG group (14.0%), while their abundance was relatively low in the HP(-) IM group (4.2%). *Cyanobacteria* were rarely found in the HP(-) cancer group (0.04%). On the contrary, the relative abundance of *Rhizobiales* was distinctly high in the HP(-) IM group (15.4%), while their abundance was very low in the HP(-) CSG and HP(-) cancer groups (1.9% and 2.8%, respectively).

3. Principal component analysis

The principal component analysis plot showed that the HP(-) CSG group was distinct from the HP(+) CSG, HP(+) IM, and HP(-) IM groups (**Figure 3A**). A linear transition was identified from the HP(+) CSG group, through the HP(+) IM group, to the HP(-) IM group, as the relative abundance of HP decreased. In **Figure 3B**, on the other hand, another linear transition was observed from the HP(-) CSG group, through the HP(-) IM group, to the HP(-) cancer group, which was mainly due to the decrease in relative abundance of *Cyanobacteria*.

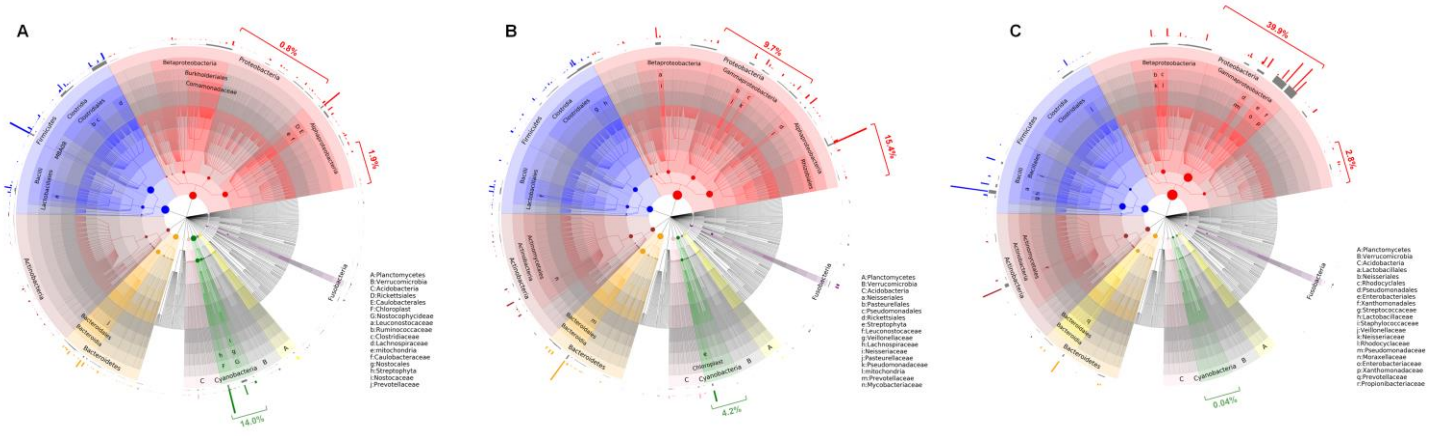

Figure 2. Cladograms for bacterial abundance in the HP-negative groups. (A) HP(-) CSG group, (B) HP(-) IM group, (C) HP(-) cancer group. Values represent the relative abundance of the specified bacterial taxa. Clade marker size in the first and second inner circles represents the relative abundance of bacteria at the phylum and class levels, respectively. Height of the inner and outer rings represents the relative abundance of bacteria at the family and genus levels, respectively. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia.

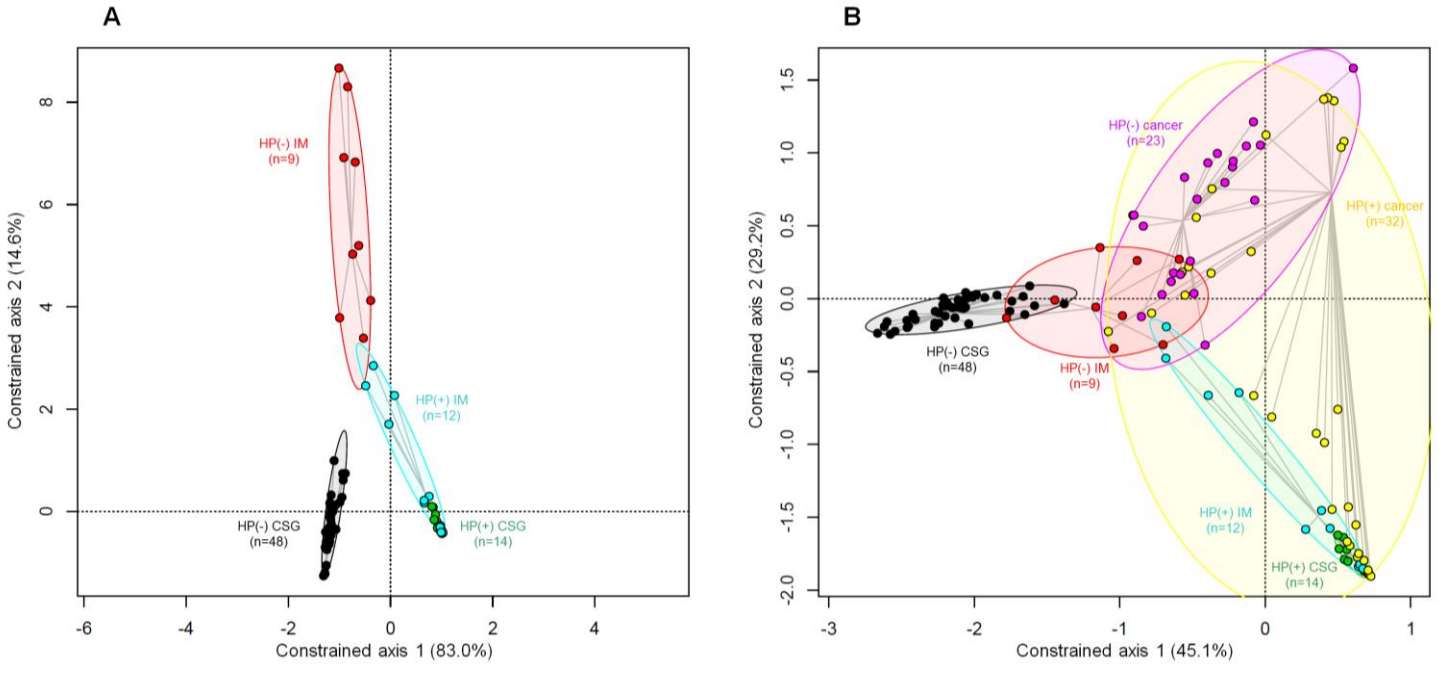

Figure 3. Principal component analysis plot. (A) CSG and IM groups, (B) All groups including cancer groups. A linear transition was identified from the HP(+) CSG group, through the HP(+) IM group, and to the HP(-) IM group, as the relative abundance of HP decreased. On the other hand, another linear transition was observed from the HP(-) CSG group, through the HP(-) IM group, to the HP(-) cancer group, mainly due to the decreased relative abundance of *Cyanobacteria*.

HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia.

4. Linear discriminant analysis for relative bacterial abundance

Figure 4 shows the results of the linear discriminant analysis for the relative bacterial abundance among the HP-negative groups. In the HP(-) CSG group, the relative abundance of *Firmicutes* and *Cyanobacteria* were significantly higher compared to those in other groups. Relative abundance of *Rhizobiales* and *Bacteroidetes* was higher in the HP(-) IM group than in the other two groups. Various bacterial taxa including *Xanthomonadaceae*, *Streptococcaceae*, *Moraxellaceae*, and *Pseudomonadaceae* were increased in the HP(-) cancer group compared to their abundances in the HP(-) CSG and HP(-) IM groups.

5. Metagenomic function

Metagenomic analysis was performed through 16S rRNA gene sequencing data and the phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt) method. Among various functions that were significantly higher in each group, three genes, namely, K03199, K03201, and K03204 (which encode type IV secretion system [T4SS] protein VirB4, VirB6, and VirB9, respectively) were noticed, because the T4SS proteins are essential for transferring CagA into the human gastric epithelium (**Figure 5**).[24] Although the gene *VirD4* (K03205) that encodes a T4SS protein was the most common in the HP(-) CSG group, it was also commonly identified in the HP(-) IM group, compared to the HP(-) cancer group. In addition, other genes encoding T4SS protein subunits, including K03194, K03197, K03198, K03200, K03202, K03203, K03195, and K03196, tended to be higher in the HP(-) IM group, compared to the HP(-) CSG and HP(-) cancer groups (**Figure 6**).

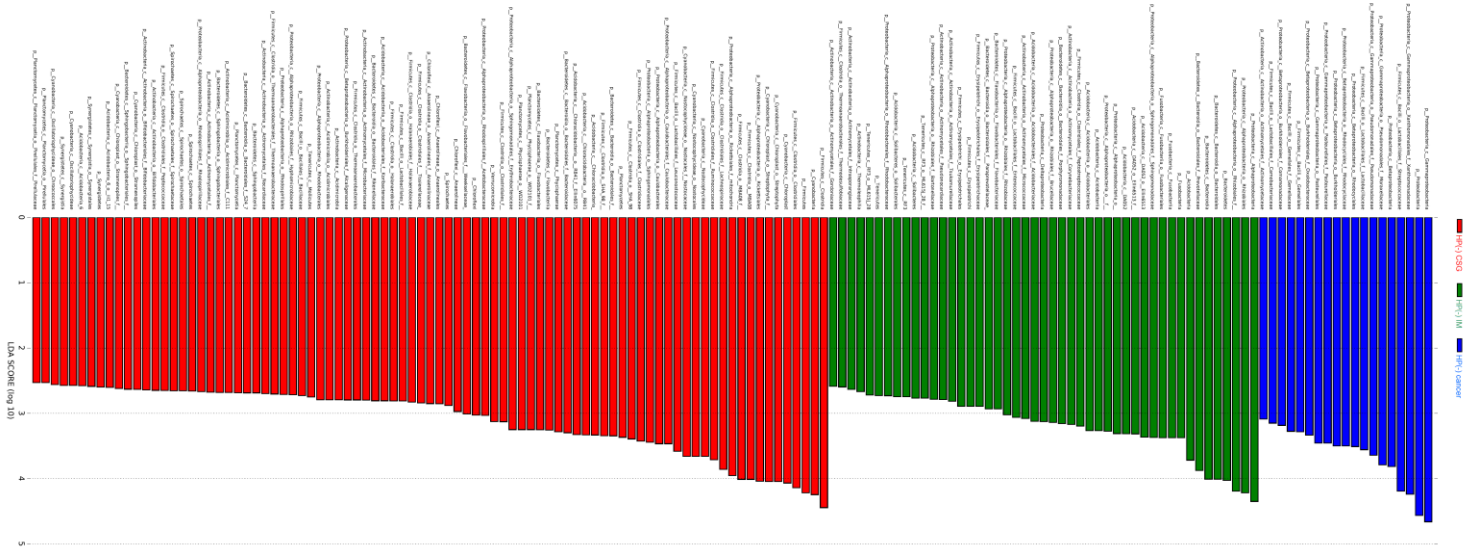

Figure 4. Linear discriminant analysis for relative abundance of bacteria in the HP-negative groups. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia; LDA, linear discriminant analysis

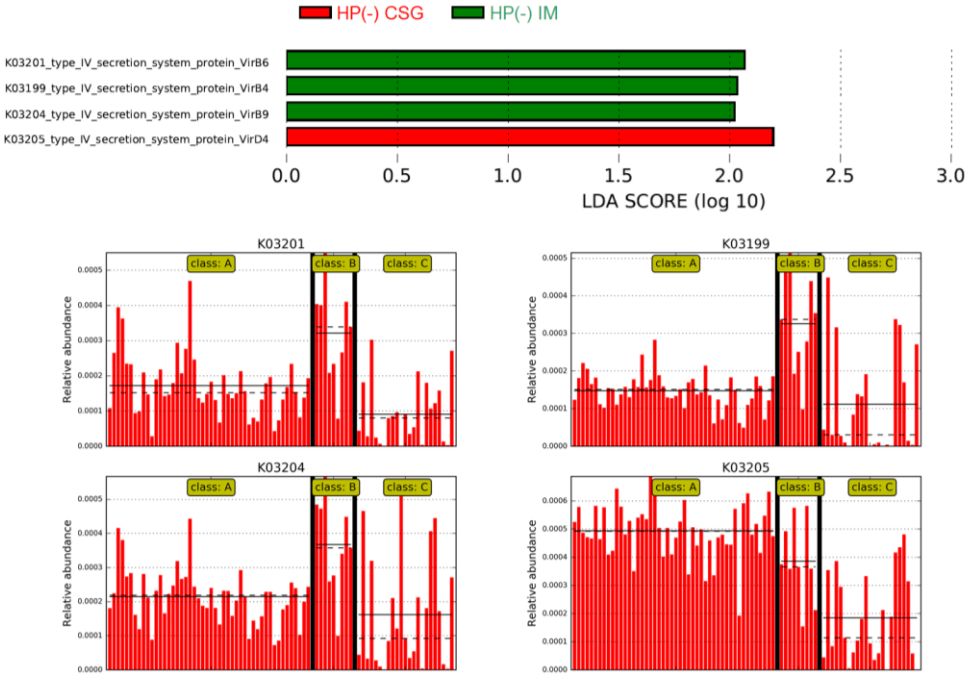

Figure 5. Highly identified genes encoding type IV secretion system protein subunits in the HP-negative CSG and IM groups. Class A, B, and C represent the HP(-) CSG, HP(-) IM, and HP(-) cancer groups, respectively. Red bar graphs represent the relative abundance of genes in each sample. Solid and dotted lines in the bar graphs represent the mean and median values of relative abundance, respectively, in each group. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia; LDA, linear discriminant analysis

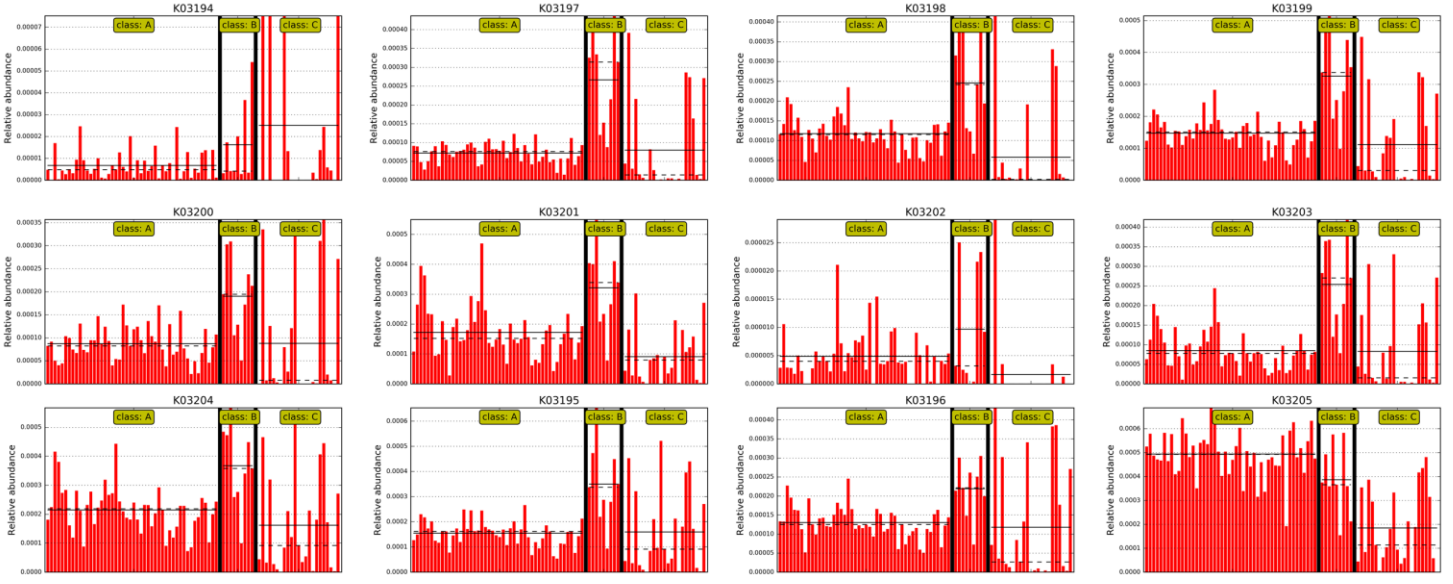

Figure 6. Relative abundance of type IV secretion system protein genes in the HP-negative groups. Class A, B, and C represent the HP(-) CSG, HP(-) IM, and HP(-) cancer groups, respectively. Solid and dotted lines in the bar graphs represent the mean and median values of relative abundance of each gene, respectively. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia; K03194, VirB1; K03197, VirB2; K03198, VirB3; K03199; VirB4; K03200, VirB5; K03201, VirB6; K03202, VirB7; K03203, VirB8; K03204, VirB9; K03195, VirB10; K03196, VirB11; K03205, VirD4.

6. Metagenome contribution

The bacterial contributions to the T4SS protein *VirB6* gene are shown in **Figure 7**. The *VirB6* gene in the HP(-) IM group was mainly derived from *Rhizobiales* (53.6% of *VirB6* genes)and *Neisseriaceae* (10.5% of *VirB6* genes), while that in the HP(-) cancer group was derived from various taxa including *Xanthomonadaceae* (30.6% of *VirB6* genes), *Neisseriaceae* (22.0% of *VirB6* genes), and *Enterobacteriaceae* (13.2% of *VirB6* genes). The analyses for the other T4SS protein subunit genes showed similar results, that *Rhizobiales* and *Neisseriaceae* were the main taxa that contributed to those T4SS genes in the HP(-) IM group (**Figure 8**).

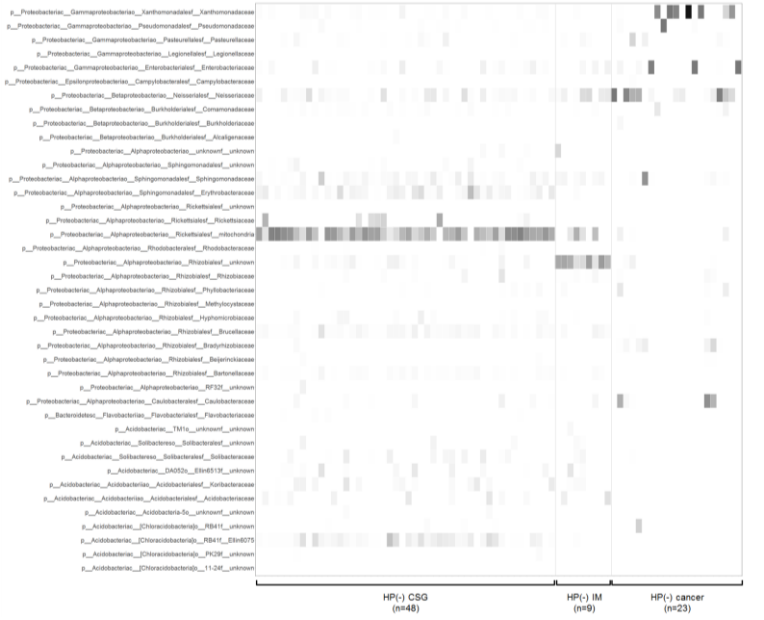

Figure 7. Metagenome contributions for the type IV secretion system protein VirB6 in the HP-negative groups. Type IV secretion system protein *VirB6* gene in the HP(-) IM group was mainly derived from *Rhizobiales* and *Neisseriaceae*, while the *VirB6* gene in the HP(-) cancer group was derived from various taxa including *Xanthomonadaceae*, *Neisseriaceae*, and *Enterobacteriaceae*. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia.

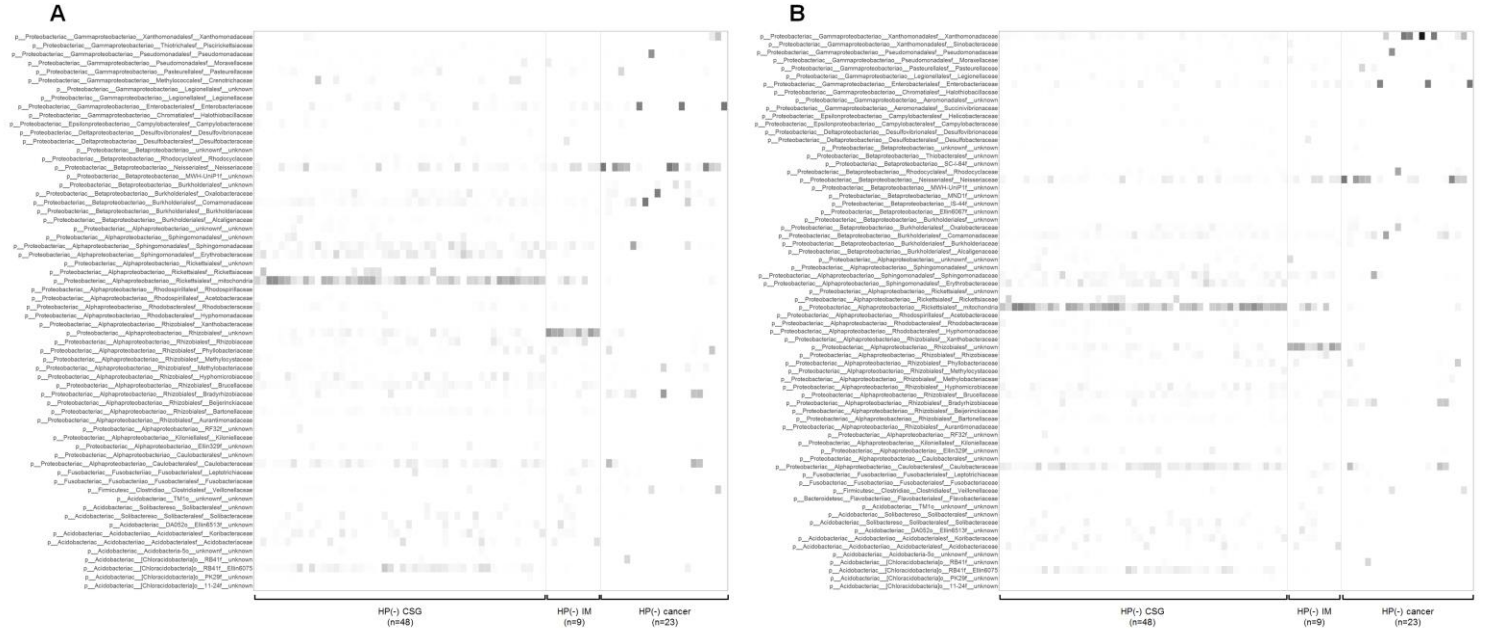

Figure 8. Metagenome contributions for type IV secretion system protein VirB4 (A) and VirB9 (B) in the HP-negative groups. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia.

7. Relationship between HP and *Rhizobiales*

Figure 9 shows the relationship of relative abundance between HP and other T4SS protein gene-contributing bacteria (*Rhizobiales* and *Neisseriaceae*) in the HP(+) CSG, HP(+) IM, and HP(-) IM groups. The relative abundance of *Rhizobiales* and *Neisseriaceae* tended to gradually increase as the abundance of HP decreased. Several samples, such as numbers 23–27 in **Figure 9,** showed a relatively high abundance of *Rhizobiales* and *Neisseriaceae* with a concurrent amount of HP abundance. In sample numbers 28–35, HP was rarely found, while *Rhizobiales* were commonly observed.

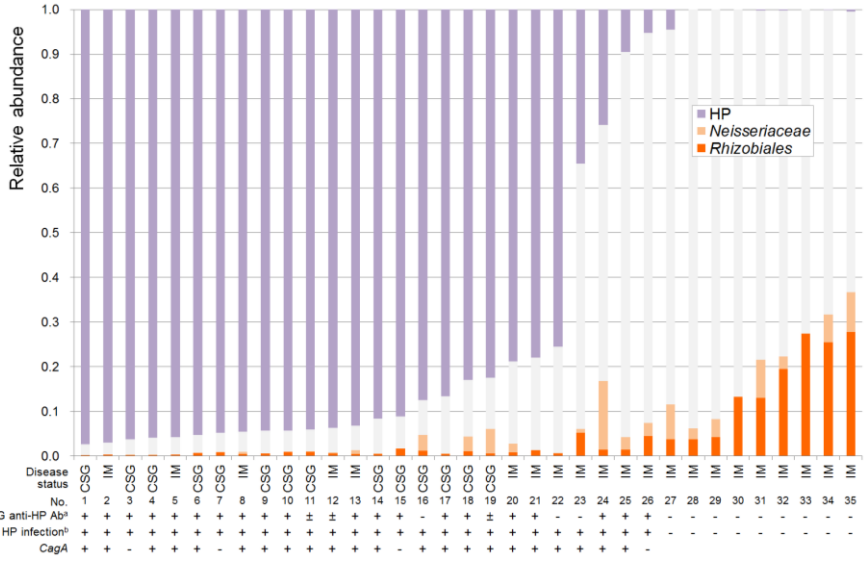

Figure 9. Relative abundance of HP, *Rhizobiales*, and *Neisseriaceae* in each sample of the HP(+) CSG, HP(+) IM, and HP(-) IM groups. The relative abundance of *Rhizobiales* and *Neisseriaceae* tended to gradually increase as the abundance of HP decreased. [a]Serum IgG anti-HP antibody was evaluated via enzyme immunoassay. [b]HP infection was determined based on conventional methods including rapid urease test, urea breath test, and histologic examination. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia.

8. HP eradication therapy in patients with CSG

Figure 10 shows relative bacterial abundance at the phylum level, before and after HP eradication therapy in patients with CSG. Before eradication therapy, the relative abundance of *Proteobacteria* was above 90% in all samples. After eradication therapy, the relative abundance of *Proteobacteria* (mainly non-HP *Proteobacteria*) decreased in the successful eradication group.

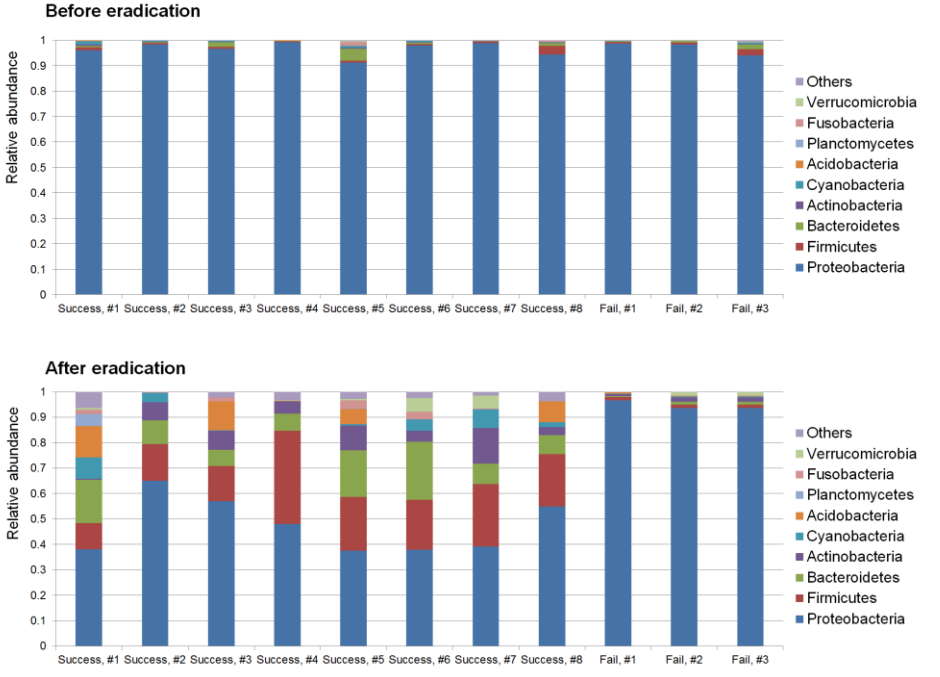

Figure 10. Relative abundance of each sample at the phylum level in the HP(+) CSG group. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis.

Distribution of bacterial abundance in the successful eradication group was similar to that in the HP(-) IM group (**Figure 11**). The successful eradication group showed a high abundance of *Rhizobiales* (16.2%) and some degree of *Cyanobacteria* (3.3%). Principal component analysis also showed similar clustering between the successful eradication and HP(-) IM groups (**Figure 12**). For quantitative assessment of group similarity, the Bray–Curtis distance between samples in the successful eradication and HP(-) IM groups was plotted as shown in **Figure 13**. Dissimilarity was even lower in samples between the successful eradication and HP(-) IM groups than those within HP(-) IM groups.

Finally, the relative abundance of T4SS protein genes in the successful eradication group was analyzed (**Figure 14**). The relative abundance of T4SS protein genes did not differ between the successful eradication and HP(-) IM groups.

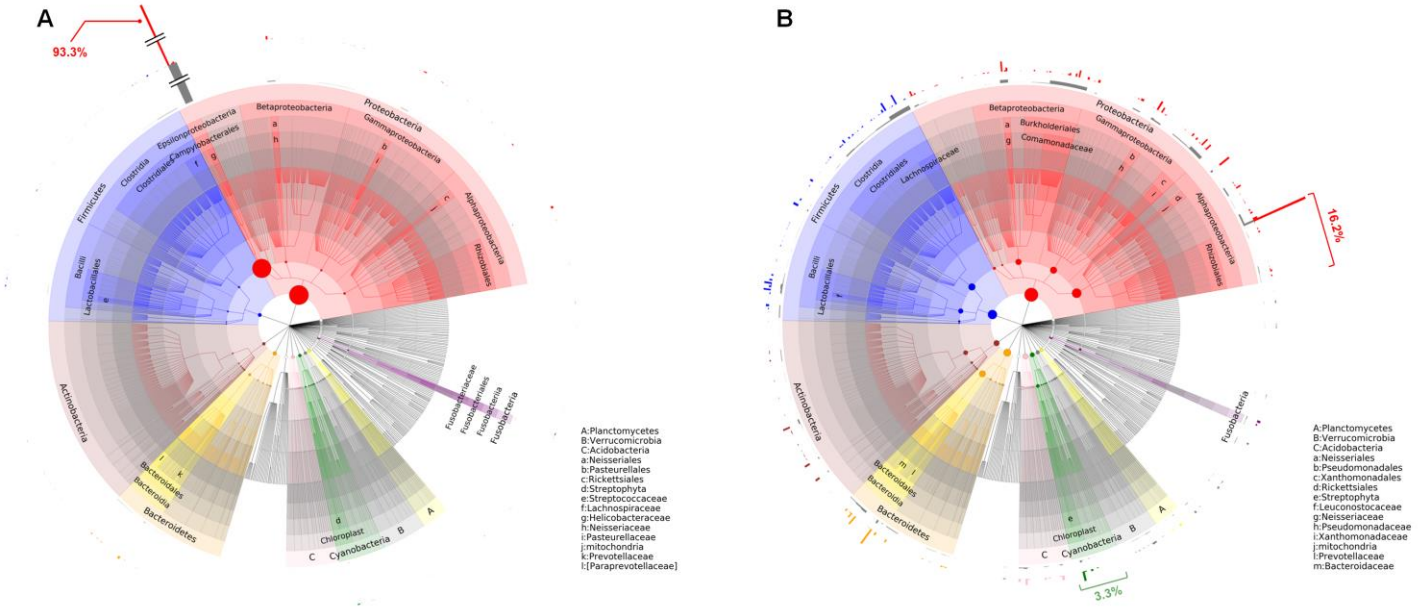

Figure 11. Cladograms for bacterial abundance before and after HP eradication in the HP(+) CSG group. (A) Before eradication in the HP(+) CSG group, (B) After eradication in the HP(+) CSG group. Clade marker size in the first and second inner circles represents the relative abundance of bacteria at the phylum and class levels, respectively. Height of the inner and outer ring represents the relative abundance of bacteria at the family and genus levels, respectively. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis.

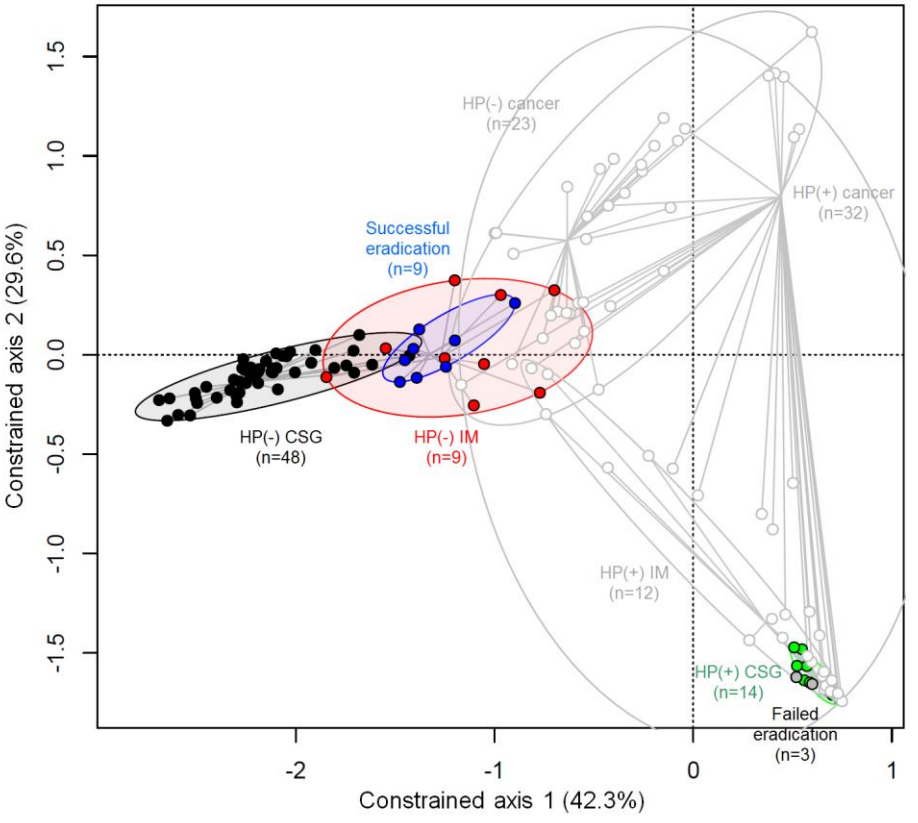

Figure 12. Principal component analysis plot for all groups including samples after HP eradication therapy. Samples in the successful eradication group were not visually distinct from those in the HP(-) IM group. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia.

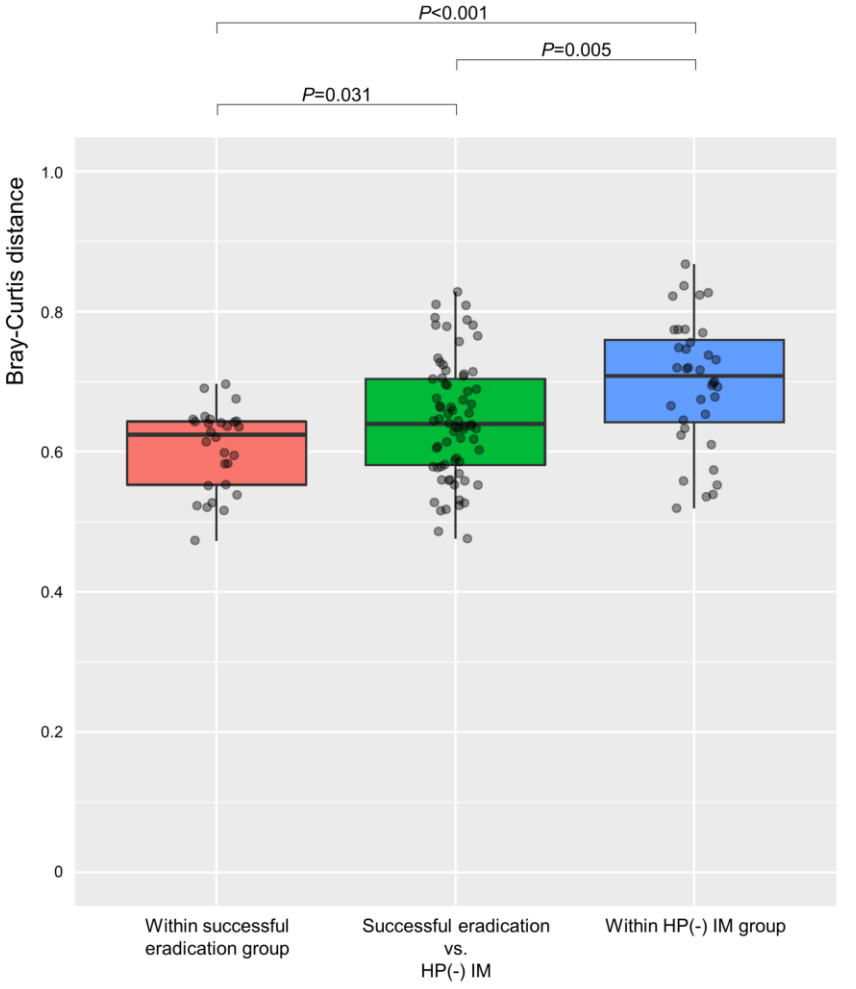

Figure 13. Bacterial community structure dissimilarity index based on Bray–Curtis distance. Dissimilarity was even lower in samples between the successful eradication and HP(-) IM groups (green box) than those within HP(-) IM groups (blue box). Dissimilarity was measured using Bray–Curtis distance of beta diversity (at family level) and group difference was determined by Wilcoxon rank-sum test. HP, *Helicobacter pylori*; CSG, chronic superficial gastritis; IM, intestinal metaplasia.

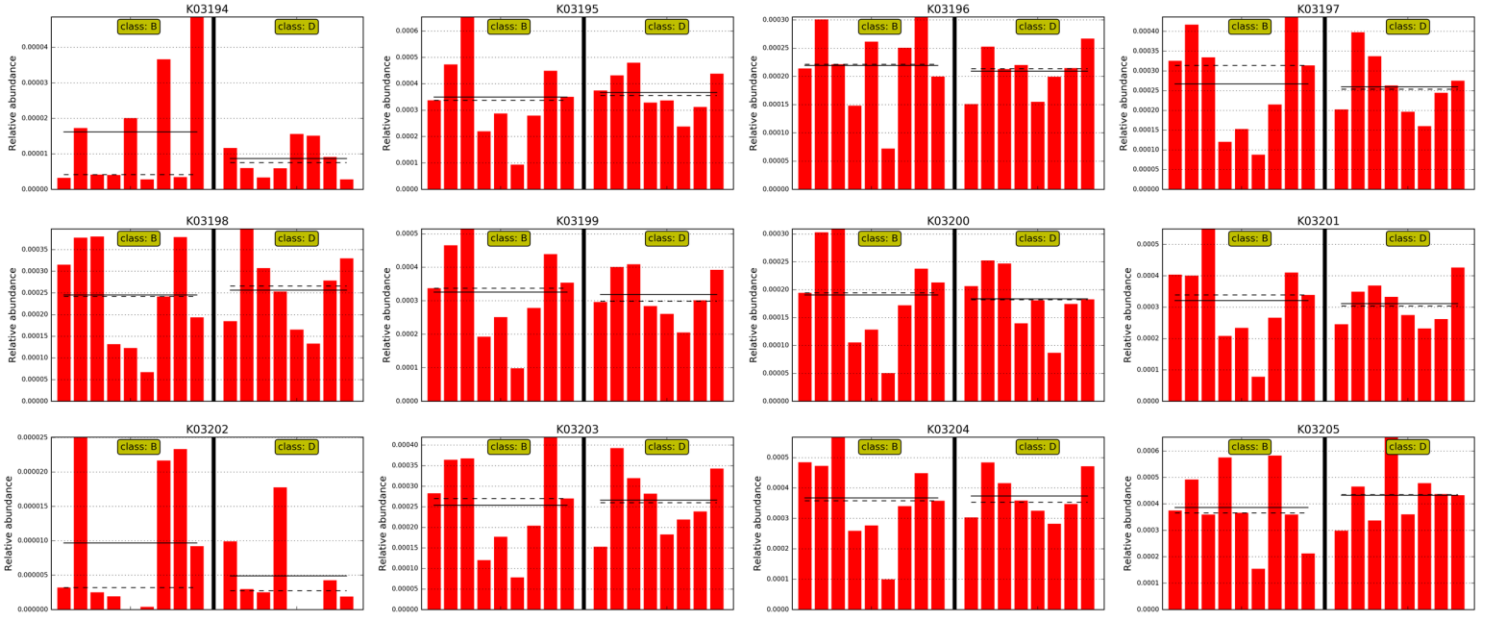

Figure 14. Relative abundance of type IV secretion system protein genes between the successful eradication and HP(-) IM groups. Class B and D represent the HP(-) IM group and successful eradication group, respectively. Solid and dotted lines in the bar graphs represent the mean and median values of relative abundance of each gene, respectively. HP, *Helicobacter pylori*; IM, intestinal metaplasia; K03194, VirB1; K03197, VirB2; K03198, VirB3; K03199; VirB4; K03200, VirB5; K03201, VirB6; K03202, VirB7; K03203, VirB8; K03204, VirB9; K03195, VirB10; K03196, VirB11; K03205, VirD4.

IV. DISCUSSION

Recent advances in high-throughput sequencing technologies aid in a better understanding of the gastric microbiome.[16,17] Gastric microbiome compositions may differ depending on various factors including mucosal atrophy or intestinal metaplasia,[16] use of medication such as proton pump inhibitors,[25] and previous gastrectomy.[23] Until now, however, gastric microbiome studies based on next-generation sequencing have focused on the phenotype of the gastric microbiome, rather than metagenomic function. Eun *et al.* demonstrated that gastric microbial compositions from patients with gastric cancer are different to patients with CSG or IM.[16] Jo *et al.* evaluated nitrosating or nitrate-reducing bacteria in gastric microbiome in patients with gastric cancer; however, they could not identify the significant role of bacteria other than HP in the gastric carcinogenesis.[17] The current study, however, has identified the potential role of the microbiome in the gastric carcinogenesis through the analysis of gastric microbiome in patients with IM.

In the stomach of patients with IM, T4SS protein genes derived from *Rhizobiales* and *Neisseriaceae* were highly abundant. T4SS protein typically consists of 11 VirB proteins (VirB1–11) and the coupled protein (VirD4).[26] These proteins enable the secretion of CagA from the bacterial cytoplasm directly into the cytoplasm of gastric epithelial cells.[24] Then, membrane-tethered CagA initiates gastric carcinogenesis by recruiting a fraction of cytoplasmic SHP2 to the membrane and dissociating β-catenin from E-cadherin.[27] Based on the findings of the current study, it may be proposed that genes encoding T4SS proteins may be transferred from gastric bacteria including *Rhizobiales* and *Neisseriaceae* to HP in patients with IM, thereby facilitating gastric carcinogenesis.

There is much evidence that horizontal gene transfer may occur between bacteria, including HP.[28,29] Especially, it was shown that genes from HP may be lost or gained during progression from atrophic gastritis to gastric

adenocarcinoma in the same patient.[29] In addition, it has been known that there are several plasticity zones in the HP genome that show high genetic variation.[30] They generally contain complete sets of genes required to produce T4SS protein, as well as genes encoding different DNA processing proteins.[29,31-33] A complete T4SS plasticity zone cluster has been known to be a virulence factor that may be important for the colonization of HP and to the development of severe outcomes of the infection with cagA-positive strains.[34] Additionally, genes encoding T4SS proteins can be transferred by conjugation between *Rhizobium radiobacter* and *Bartonella rattaustraliani*.[35] Network analysis of gene sharing among Proteobacterial genomes also showed that higher horizontal gene transfer rates are inferred within the phylum level than the species level.[36] Although there is no direct evidence of horizontal gene transfer between HP and *Rhizobiales* yet, the results may suggest the possibility of T4SS protein gene transfer between HP and other microbiome components.

Based on the results of the current study, it is possible that gastric carcinogenesis initiation may require HP as well as other microbiomes such as *Rhizobiales* and *Neisseriaceae*. From this point of view, patients with HP-predominant CSG are not regarded as being at a high-risk for gastric cancer. As HP-infected stomach tissue progressed from gastritis to IM, however, the abundance of gastric microbiomes other than HP, including *Rhizobiales* and *Neisseriaceae*, increased. Therefore, patients with progressed IM, who exhibit abundance of HP and other microbiomes, may be high-risk individuals for gastric cancer development. If CagA translocation occurs in these patients, it will not be possible to prevent gastric cancer even if HP has been eradicated.

Before the study began, it was expected that the gastric microbiome in patients with CSG after HP eradication (the successful eradication group) may be similar to in the HP(-) CSG group, because HP eradication therapy may reduce the risk of gastric cancer development in patients with CSG.[9] However, the gastric microbiome in the successful eradication group resembled that in the

HP(-) IM group which is regarded as a high risk group. The results implied that the gastric microbiome composition did not depend on the cause of HP regression (eradication by antibiotics vs. spontaneous regression after IM progression). Nevertheless, a change in the *Rhizobiales*-abundant microbiome after HP eradication therapy does not mean an increase in gastric cancer risk, because the source of CagA has already been removed through HP eradication. In other words, the risk-lowering effect of HP eradication therapy was because of HP eradication itself, rather than changes in gastric microbiome components besides HP.

One of the interesting findings of the study was that *Cyanobacteria* were highly identified in the HP(-) CSG group. The relative abundance of *Cyanobacteria* decreased from the HP(-) CSG group, through the HP(-) IM group, to the HP(-) cancer group. *Cyanobacteria* abundance was not recovered in patients with a history of HP infection, even after HP eradication. These findings implied that *Cyanobacteria*-abundant microbiomes may be a cause, rather than a result, of an absence of HP infection. Several previous studies showed that polysaccharides derived from *Cyanobacteria* prevented HP attachment to gastric mucosa.[37,38] Considering both the potentially preventive effect of *Cyanobacteria* for HP infection and the relatively low abundance of *Cyanobacteria* in HP-eradicated patients, re-infection risk of HP should be kept in mind in patients that have undergone eradication therapy.

The current study firstly demonstrated bacterial taxa that affect gastric carcinogenesis through next-generation sequencing. Previous studies did not identify *Rhizobiales* as a bacteria that contributes to gastric carcinogenesis.[16,17] There are two main reasons for these discrepant results between studies. First, in the initial analysis of microbiome, the HP-negative groups and HP-positive groups were evaluated separately. Because HP is the most abundant taxa in HP-infected stomachs, differences in other microbiome species may not be perceived if HP-positive and HP-negative groups are analyzed together. After

identifying *Rhizobiales* and *Neisseriaceae* as taxa that potentially contribute to gastric carcinogenesis, the relationship among these bacterial taxa in both HP-positive and HP-negative groups was analyzed together. Second, gastric microbiome in patients with IM, rather than cancer, was focused, because gastric carcinogenesis initiation events, such as CagA translocation, may occur in this stage. Considering the bacterial driver-passenger model in carcinogenesis,[39] the abundance of various taxa increased in the HP(-) cancer group, including *Xanthomonadaceae*, *Streptococcaceae*, *Moraxellaceae*, and *Pseudomonadaceae*, which may be a result, rather than a cause, of cancer development.

Although this study was the first 16S rRNA gene sequencing analysis to propose bacteria that potentially contribute to gastric carcinogenesis, it had several limitations. First, the sample size was relatively small in several groups such as the HP(+) IM and HP(-) IM groups. Therefore, only *VirB4*, *VirB6*, and *VirB9* genes of T4SS protein subunits were significantly different in abundance among the groups. However, other T4SS subunit genes also tended to increase in the HP(-) IM group compared to the HP(-) CSG and HP(-) cancer groups. Larger studies may provide a definitive conclusion and identify additional metagenomic functions that affect gastric carcinogenesis. Second, most sequence data in the cancer group were downloaded from the NCBI, although all samples in patients with CSG or IM were obtained from the one institute (Hanyang University Guri Hospital). Different study settings and sample collections may result in a bias. However, the current study mainly focused on gastric microbiome in patients with IM compared to those with CSG, rather than cancer. In addition, 12 of 52 downloaded samples were deposited by the research group of Hanyang University Guri Hospital through the previous study (SRP038955).[16] Another 34 samples were collected from a geographically adjacent hospital from the institute (GSE61493).[17] Third, direct evidence of horizontal gene transfer between HP and other bacteria including *Rhizobiales*

and *Niserriaceae* was not demonstrated through the current study. Although we hypothesized that T4SS gene transfer would occur between HP and other bacteria, horizontal gene transfer is more likely to occur in the same class than in other class. Bacteria including *Rhizobiales* may affect the gastric carcinogenesis through other mechanisms rather than T4SS gene transfer. Subsequent experimental studies are needed to validate the hypothesis derived from the study.

V. CONCLUSION

Despite these limitations, the current study provides a better understanding of the potential role of the gastric microbiome in gastric carcinogenesis. As HP-infected stomachs progressed from gastritis to IM, the relative abundance of *Rhizobiales* increased. In the metagenome of patients with IM, T4SS protein genes, which are mainly derived from *Rhizobiales* and *Neisseriaceae*, were highly observed. In addition, the gastric microbiome after HP eradication therapy was similar to the microbiome observed after spontaneous HP regression.

REFERENCES

1. Schistosomes, liver flukes and Helicobacter pylori. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994. IARC Monogr Eval Carcinog Risks Hum 1994;61:1-241.
2. Yoon H, Kim N, Lee HS, Shin CM, Park YS, Lee DH, et al. Helicobacter pylori-Negative Gastric Cancer in South Korea: Incidence and Clinicopathologic Characteristics. Helicobacter 2011;16:382-8.
3. Matsuo T, Ito M, Takata S, Tanaka S, Yoshihara M, Chayama K. Low prevalence of Helicobacter pylori-negative gastric cancer among Japanese. Helicobacter 2011;16:415-9.
4. Nishizawa T, Suzuki H. Gastric Carcinogenesis and Underlying Molecular Mechanisms: Helicobacter pylori and Novel Targeted Therapy. Biomed Res Int 2015;2015:794378.
5. Suzuki H, Nishizawa T, Tsugawa H, Hibi T. Molecular approaches and modern clinical strategies for the management of Helicobacter pylori infection in Japan. Keio J Med 2012;61:109-19.
6. Chen S, Ying L, Kong M, Zhang Y, Li Y. The Prevalence of Helicobacter pylori Infection Decreases with Older Age in Atrophic Gastritis. Gastroenterol Res Pract 2013;2013:494783.
7. Chon HJ, Hyung WJ, Kim C, Park S, Kim JH, Park CH, et al. Differential Prognostic Implications of Gastric Signet Ring Cell Carcinoma: Stage Adjusted Analysis From a Single High-volume Center in Asia. Ann Surg 2017;265:946-53.
8. Fox JG, Wang TC. Inflammation, atrophy, and gastric cancer. J Clin Invest 2007;117:60-9.
9. Wong BC, Lam SK, Wong WM, Chen JS, Zheng TT, Feng RE, et al. Helicobacter pylori eradication to prevent gastric cancer in a high-risk region of China: a randomized controlled trial. JAMA 2004;291:187-94.
10. Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. Curr Opin Gastroenterol 2015;31:69-75.
11. Calmels S, Dalla Venezia N, Bartsch H. Isolation of an enzyme catalysing nitrosamine formation in Pseudomonas aeruginosa and Neisseria mucosae.

Biochem Biophys Res Commun 1990;171:655-60.

12. Mowat C, Williams C, Gillen D, Hossack M, Gilmour D, Carswell A, et al. Omeprazole, Helicobacter pylori status, and alterations in the intragastric milieu facilitating bacterial N-nitrosation. Gastroenterology 2000;119:339-47.
13. Forsythe SJ, Dolby JM, Webster AD, Cole JA. Nitrate- and nitrite-reducing bacteria in the achlorhydric stomach. J Med Microbiol 1988;25:253-9.
14. Williams C, McColl KE. Review article: proton pump inhibitors and bacterial overgrowth. Aliment Pharmacol Ther 2006;23:3-10.
15. Ragan MA, Beiko RG. Lateral genetic transfer: open issues. Philos Trans R Soc Lond B Biol Sci 2009;364:2241-51.
16. Eun CS, Kim BK, Han DS, Kim SY, Kim KM, Choi BY, et al. Differences in gastric mucosal microbiota profiling in patients with chronic gastritis, intestinal metaplasia, and gastric cancer using pyrosequencing methods. Helicobacter 2014;19:407-16.
17. Jo HJ, Kim J, Kim N, Park JH, Nam RH, Seok YJ, et al. Analysis of Gastric Microbiota by Pyrosequencing: Minor Role of Bacteria Other Than Helicobacter pylori in the Gastric Carcinogenesis. Helicobacter 2016;21:364-74.
18. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;7:335-6.
19. Kembel SW, Wu M, Eisen JA, Green JL. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. PLoS Comput Biol 2012;8:e1002743.
20. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol 2011;12:R60.
21. Yamaoka Y, Kodama T, Kita M, Imanishi J, Kashima K, Graham DY. Relationship of vacA genotypes of Helicobacter pylori to cagA status, cytotoxin production, and clinical outcome. Helicobacter 1998;3:241-53.
22. Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA

marker gene sequences. Nature biotechnology 2013;31:814-21.

23. Tseng CH, Lin JT, Ho HJ, Lai ZL, Wang CB, Tang SL, et al. Gastric microbiota and predicted gene functions are altered after subtotal gastrectomy in patients with gastric cancer. Sci Rep 2016;6:20701.
24. Backert S, Selbach M. Role of type IV secretion in Helicobacter pylori pathogenesis. Cell Microbiol 2008;10:1573-81.
25. Tsuda A, Suda W, Morita H, Takanashi K, Takagi A, Koga Y, et al. Influence of Proton-Pump Inhibitors on the Luminal Microbiota in the Gastrointestinal Tract. Clin Transl Gastroenterol 2015;6:e89.
26. Kwok T, Zabler D, Urman S, Rohde M, Hartig R, Wessler S, et al. Helicobacter exploits integrin for type IV secretion and kinase activation. Nature 2007;449:862-6.
27. Hatakeyama M. Helicobacter pylori CagA and gastric cancer: a paradigm for hit-and-run carcinogenesis. Cell Host Microbe 2014;15:306-16.
28. Fischer W, Breithaupt U, Kern B, Smith SI, Spicher C, Haas R. A comprehensive analysis of Helicobacter pylori plasticity zones reveals that they are integrating conjugative elements with intermediate integration specificity. BMC Genomics 2014;15:310.
29. Oh JD, Kling-Backhed H, Giannakis M, Xu J, Fulton RS, Fulton LA, et al. The complete genome sequence of a chronic atrophic gastritis Helicobacter pylori strain: evolution during disease progression. Proc Natl Acad Sci U S A 2006;103:9999-10004.
30. Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, et al. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature 1999;397:176-80.
31. Fischer W, Windhager L, Rohrer S, Zeiller M, Karnholz A, Hoffmann R, et al. Strain-specific genes of Helicobacter pylori: genome evolution driven by a novel type IV secretion system and genomic island transfer. Nucleic Acids Res 2010;38:6089-101.
32. Kersulyte D, Velapatino B, Mukhopadhyay AK, Cahuayme L, Bussalleu A, Combe J, et al. Cluster of type IV secretion genes in Helicobacter pylori's plasticity zone. J Bacteriol 2003;185:3764-72.

33. Kersulyte D, Lee W, Subramaniam D, Anant S, Herrera P, Cabrera L, et al. Helicobacter Pylori's plasticity zones are novel transposable elements. PLoS One 2009;4:e6859.

34. Silva B, Nunes A, Vale FF, Rocha R, Gomes JP, Dias R, et al. The expression of Helicobacter pylori tfs plasticity zone cluster is regulated by pH and adherence, and its composition is associated with differential gastric IL-8 secretion. Helicobacter 2017;22.

35. Saisongkorh W, Robert C, La Scola B, Raoult D, Rolain JM. Evidence of transfer by conjugation of type IV secretion system genes between Bartonella species and Rhizobium radiobacter in amoeba. PLoS One 2010;5:e12666.

36. Kloesges T, Popa O, Martin W, Dagan T. Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. Molecular biology and evolution 2011;28:1057-74.

37. Ascencio F, Gama NL, De Philippis R, Ho B. Effectiveness of Cyanothece spp. and Cyanospira capsulata exocellular polysaccharides as antiadhesive agents for blocking attachment of Helicobacter pylori to human gastric cells. Folia Microbiol (Praha) 2004;49:64-70.

38. Loke MF, Lui SY, Ng BL, Gong M, Ho B. Antiadhesive property of microalgal polysaccharide extract on the binding of Helicobacter pylori to gastric mucin. FEMS Immunol Med Microbiol 2007;50:231-8.

39. Tjalsma H, Boleij A, Marchesi JR, Dutilh BE. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. Nature Reviews Microbiology 2012;10:575-82.

ABSTRACT(IN KOREAN)

16S rRNA 유전자 염기서열 분석법을 이용한 장상피화생 환자의 위 미생물총과 메타게놈 기능 분석

<지도교수 이상길>

연세대학교 대학원 의학과

박 찬 혁

최근 위 내 미생물총에 관한 연구의 발전에도 불구하고, 위암 발병 단계에 있어 헬리코박터 파일로리 (HP) 이외의 위 내 미생물총의 역할에 대해서는 충분히 밝혀져 있지 않다. 본 연구에서는 메타게놈 분석을 통하여 위암 발병 단계에 있어 위 내 미생물총이 갖고 있는 잠재적 기능을 알아보고자 하였다. 대상자를 위 내 질환의 종류 (만성표재성위염 [CSG], 장상피화생 [IM], 암 [cancer])와 HP 감염 여부 (HP 양성, HP 음성)에 따라 총 6개 군으로 분류하였다. 위 내 미생물총은 위 전정부의 점막 조직을 이용하여 16S rRNA 유전자 시퀀싱 기법을 통해 분석하였다. 메타게놈 분석은 phylogenetic investigation of communities by reconstruction of unobserved states 방법을 통해 시행하였다. 총 138명의 대상자 중, HP(-) CSG 군, HP(-) IM 군, HP(-) cancer 군, HP(+) CSG 군, HP(+) IM 군, HP(+) cancer 군은 각각 48, 9, 23, 14, 12, 32명 이었다. HP(-) CSG 군에서는 *Cyanobacteria* 가 가장 우세하였던 반면, HP(-) IM 군에서는 *Rhizobiales* 가 가장 우세하였다. HP(-) IM 군의 메타게놈에서는 type IV secretion system protein 을 암호화하는 유전자 (특히 type IV secretion system protein의 VirB4, VirB6, VirB9 하위단위를 암호화하는 유전자)가 우세하게 발견되었다. HP(-) IM 군에서 type IV secretion system protein 을 암호화하는 유전자는 대부분 *Rhizobiales*에서 유래되었다. 이와 더불어, HP 제균치료 후 위 내 미생물총은 HP가 자연소실 된 후의 위 내 미생물총과 유사하게 변화함을 확인하였다.

핵심되는 말 : 헬리코박터 파일로리; 미생물총; 위암; 장상피화생; type IV secretion system protein