iScience, Volume *25*

Supplemental information

Multi-omics reveals microbiome, host gene expression, and immune landscape in gastric carcinogenesis

Chan Hyuk Park, Changjin Hong, A-reum Lee, Jaeyun Sung, and Tae Hyun Hwang

Table S1. Baseline characteristics of 30 participants who underwent endoscopic biopsy, Related to Figure 1.

Variable	Healthy stomach	Post-HP eradication	HP-associated gastritis	Gastric cancer	*P*-value
Number	10	6	11	3	
Age, year, median (IQR)	26 (24.8-29.8)	28 (26.3-33.5)	32 (27.0-63.0)	65 (63.0-69.0)	0.013
Male, n (%)	9 (90.0)	0 (0.0)	6 (54.5)	2 (66.7)	0.003
Comorbidity, n (%)					
Hypertension	0 (0.0)	0 (0.0)	3 (27.3)	0 (0.0)	0.258
Diabetes	0 (0.0)	0 (0.0)	2 (18.2)	0 (0.0)	0.457
Cerebrovascular accident history	0 (0.0)	0 (0.0)	1 (9.1)	0 (0.0)	1.000
Current HP infection, n (%)	0 (0.0)	0 (0.0)	9 (81.8)	1 (33.3)	<0.001
Serologic testing[a]					
IgG anti-HP antibody, n (%)					<0.001
Negative	9 (0.0)	0 (0.0)	3 (27.3)	0 (0.0)	
Equivocal	1 (10.0)	2 (40.0)	0 (0.0)	0 (0.0)	
Positive	0 (0.0)	3 (60.0)	8 (72.7)	2 (100.0)	
Pepsinogen, median (IQR)					
Pepsinogen I, ng/mL	52.8 (38.8-64.3)	87.4 (68.3-118.7)	90.2 (67.8-105.2)	190.9 (107.7-274.0)	0.052
Pepsinogen II, ng/mL	10.3 (8.0-14.3)	37.2 (19.2-46.1)	28.8 (16.3-43.8)	70.7 (32.5-108.9)	0.009
Pepsinogen I/II ratio	5.0 (4.0-5.4)	3.1 (2.3-3.7)	3.3 (2.3-4.9)	2.9 (2.5-3.3)	0.016
Histology (gastric cancer), n (%)					
Moderately differentiated adenocarcinoma	N/A	N/A	N/A	3 (100.0)	N/A

[a]Values were missing for two patients in the cohort.

HP, Helicobacter pylori; IQR, interquartile range; N/A, not applicable

Table S2. Baseline characteristics of 40 participants with gastric cancer who underwent gastrectomy, Related to Figure 1.

Variable	Value
Number	40
Patient characteristics	
Age, year, median (IQR)	64 (58.3-69.0)
Male, n (%)	30 (75.0)
Comorbidity, n (%)	
Hypertension	16 (40.0)
Diabetes	10 (25.0)
Cerebrovascular accident history	1 (2.5)
Current HP infection, n (%)	32 (80.0)
Lesion characteristics (gastric cancer)	
Location, n (%)	
Upper third	4 (10.0)
Middle third	19 (47.5)
Lower third	17 (42.5)
Size, cm, median (IQR)	7.0 (5.5-9.5)
Histology, n (%)	
Well-differentiated adenocarcinoma	3 (7.5)
Moderately differentiated adenocarcinoma	11 (27.5)
Poorly differentiated adenocarcinoma	18 (45.0)
Signet ring cell carcinoma	7 (17.5)
Mucinous adenocarcinoma	1 (2.5)
TNM stage, n (%)	
Stage I	1 (2.5)
Stage II	8 (20.0)
Stage III	27 (67.5)
Stage IV	4 (10.0)

HP, Helicobacter pylori; TNM, tumor-node-metastasis; IQR, interquartile range

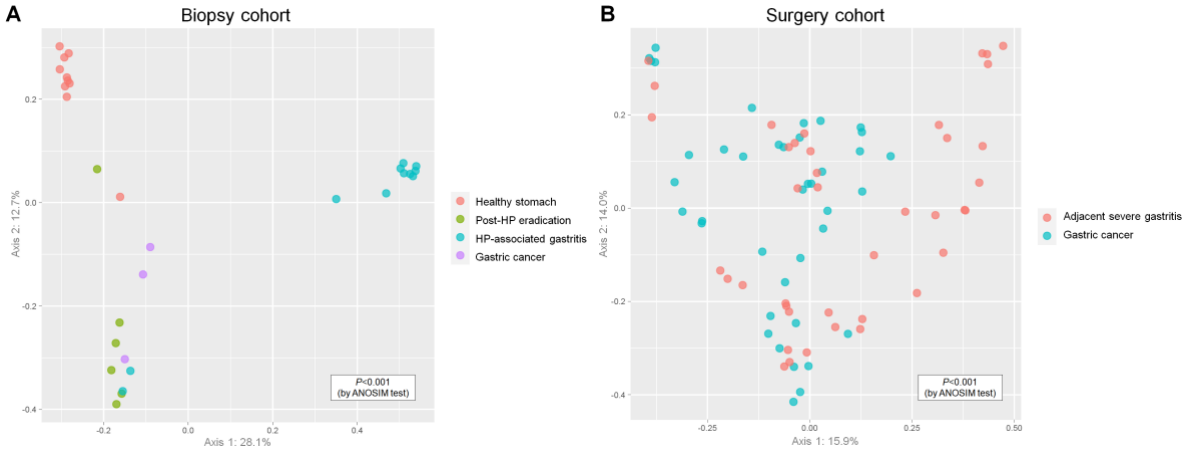

Figure S1. Beta diversity analysis for the microbiome dataset in the biopsy (A) and surgery (B) cohorts, Related to Figure 2.

The Bray-Curtis distance matrix was used for the principal coordinate analysis. The microbiome composition was significantly different between the groups.

HP, *Helicobacter pylori*; ANOSIM, analysis of similarities

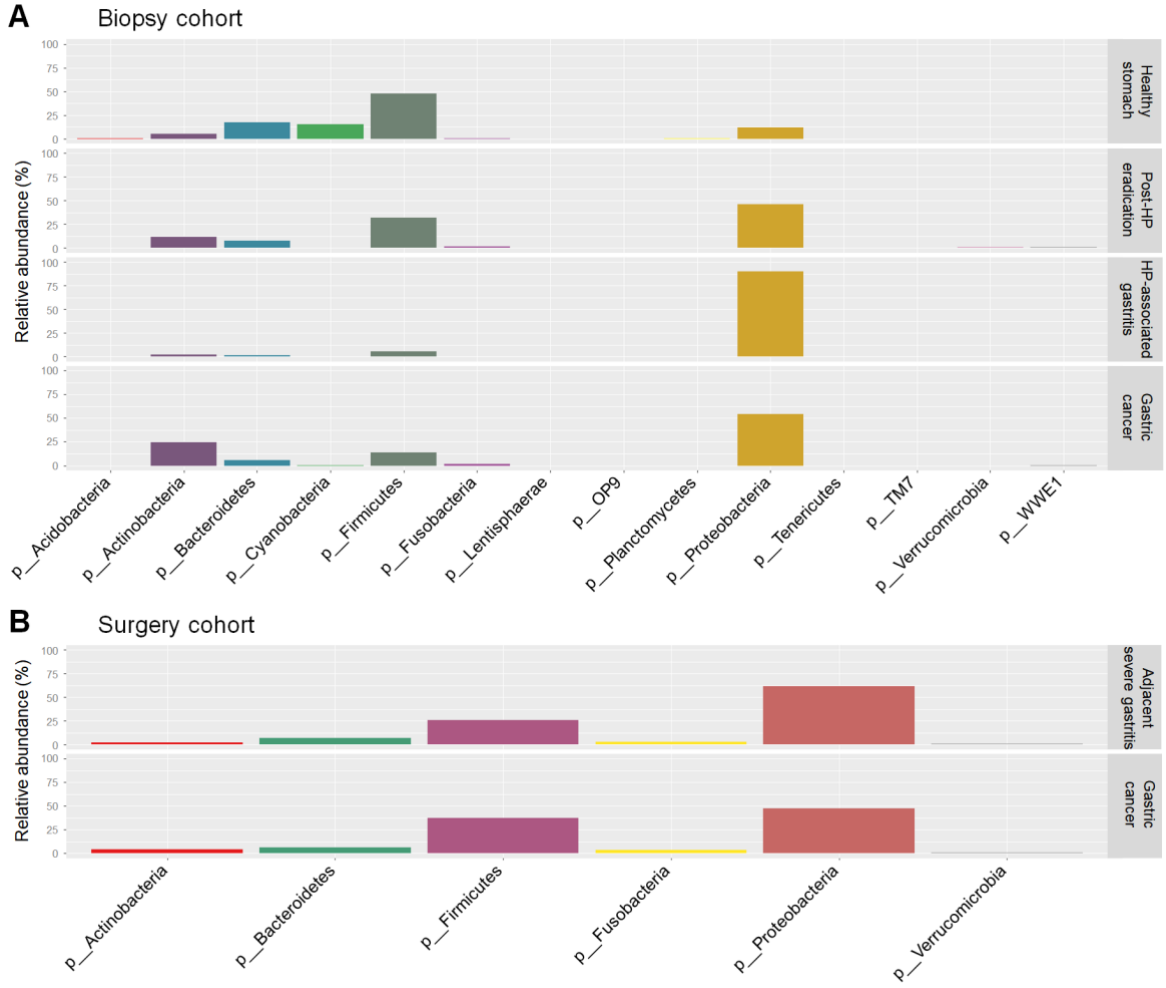

Figure S2. The relative abundance of bacterial taxa at the phylum level in the biopsy (A) and surgery (B) cohorts, Related to Figure 3.

In the healthy stomach group of the biopsy cohort, Firmicutes was the most abundant phylum. In the other groups (post-HP eradication, HP-associated gastritis, and gastric cancer groups of the biopsy cohort and adjacent severe gastritis and gastric cancer groups of the surgery cohort), the most abundant phylum was Proteobacteria, followed by Firmicutes.

HP, *Helicobacter pylori*

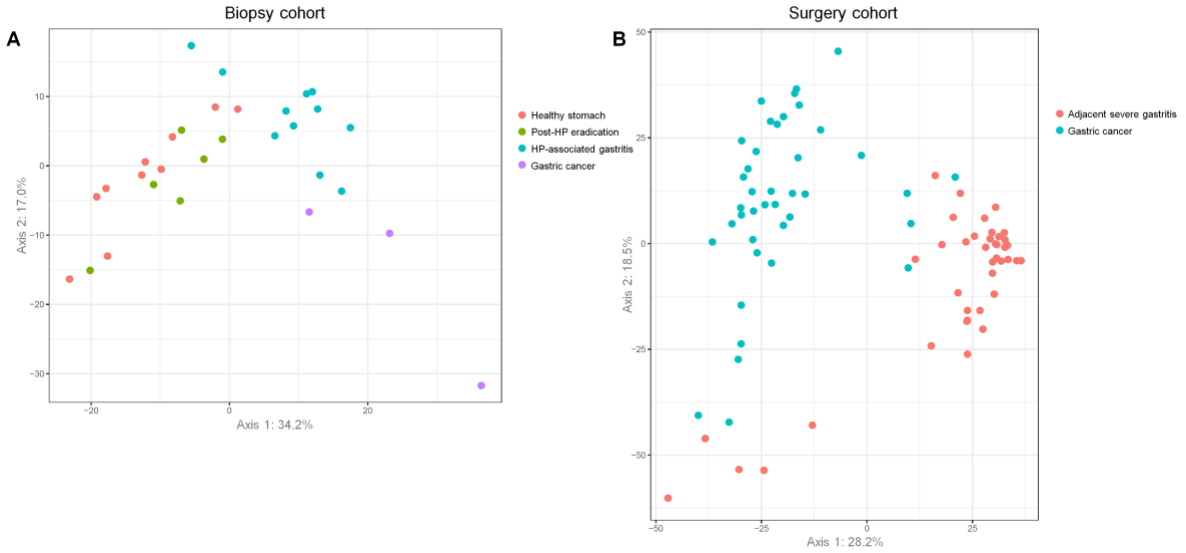

Figure S3. Principal component analysis for gene expression analysis in the biopsy (A) and surgery (B) cohorts, Related to Table 1.

In the biopsy cohort, gene expression patterns differed among the healthy stomach, HP-associated gastritis, and gastric cancer groups; however, a relatively similar gene expression pattern was observed between the healthy stomach and post-HP eradication group. In the surgery cohort, the gastric cancer group showed a different gene expression pattern compared to the adjacent severe gastritis group.

HP, *Helicobacter pylori*

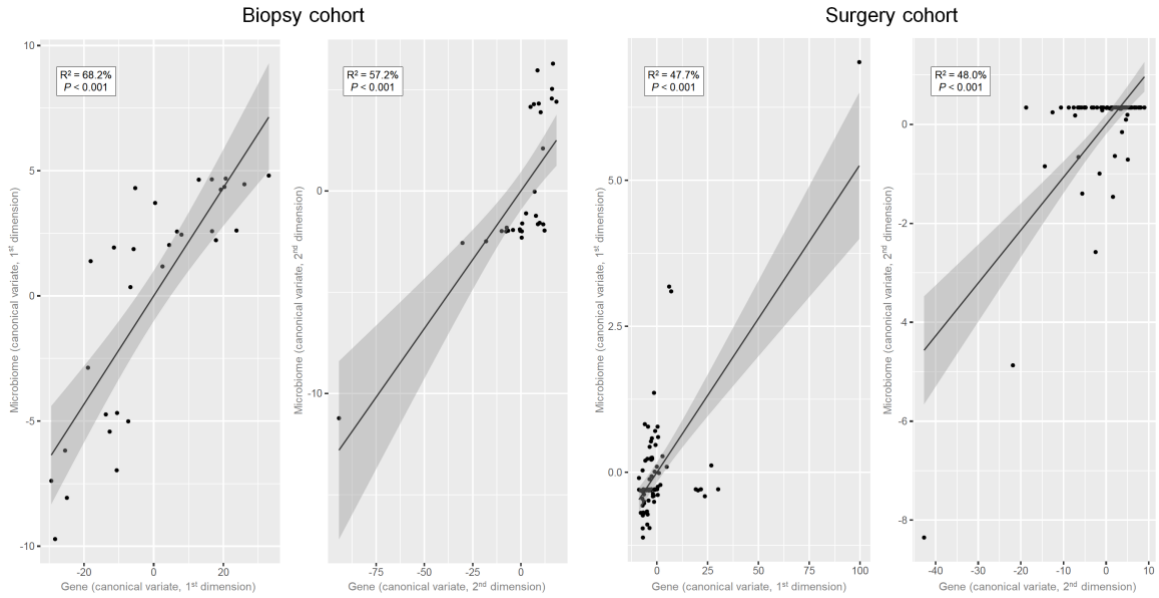

Figure S4. Scatter plots for the first and second canonical variates composed of microbiome and gene features, Related to Figures 4 and 5.

Black points represent study samples (N = 30 [biopsy cohort], N = 80 [surgery cohort]).

The solid black line and gray zone represents the linear regression line and 95% confidence interval, respectively.

Figure S5. Principal component analysis for immune cell-type enrichment analysis in the biopsy (A) and surgery (B) cohorts, Related to Figure 6.

In the biopsy cohort, the healthy stomach group showed a different immune cell-type enrichment pattern compared to the non-healthy stomach group (i.e., post-HP eradication, HP-associated gastritis, and gastric cancer). In the surgery cohort, samples in the gastric cancer group were separated from those in the adjacent severe gastritis group.

HP, *Helicobacter pylori*

Figure S6. Relative abundance of different immune cell types infiltrated in gastric tissues, Related to Figure 6.

The x-axis indicates each sample in the biopsy and surgery cohorts. The y-axis indicates the relative abundance of infiltrated immune cells.

CIBERSORT, Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts

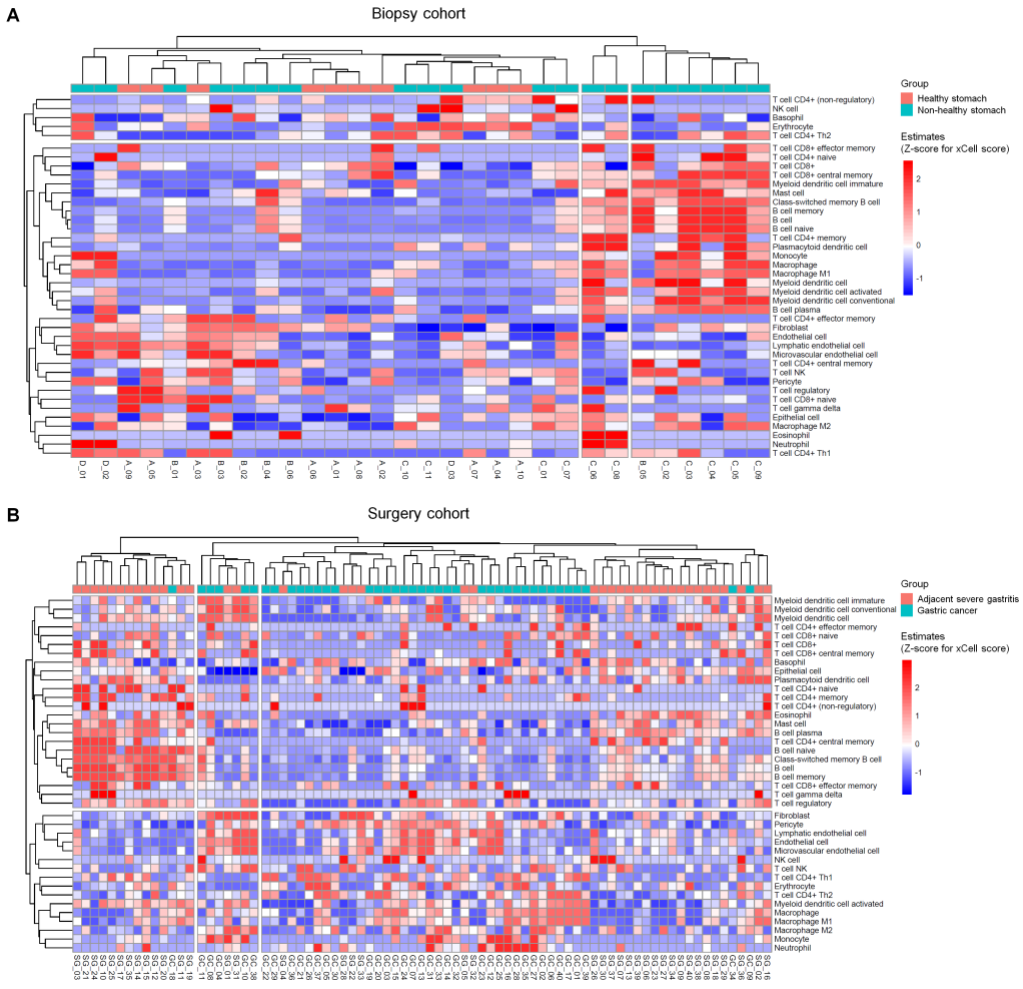

Figure S7. Hierarchical clustering between study samples and immune cell types in the biopsy (A) and surgery (B) cohorts, Related to Figure 6.

In the biopsy cohort, B cell infiltration was prominent in the non-healthy stomach (i.e., post-HP eradication, HP-associated gastritis, and gastric cancer) compared to the healthy stomach. Most non-healthy gastric tissues were differentially clustered from healthy gastric tissues by mainly B cells. In the surgery cohort, gastric cancer tissues were distinguished from adjacent severe gastritis tissues, which showed abundant B cell infiltration. The x-axis indicates each sample in the biopsy and surgery cohorts. The prefix of each sample label denotes the study group (A_: healthy stomach, B_: post-HP eradication, C_: HP-associated gastritis, D_: gastric cancer [biopsy cohort], GC_: gastric cancer [surgery cohort], and SG_: severe gastritis). The y-axis indicates each immune cell type.

HP, *Helicobacter pylori*

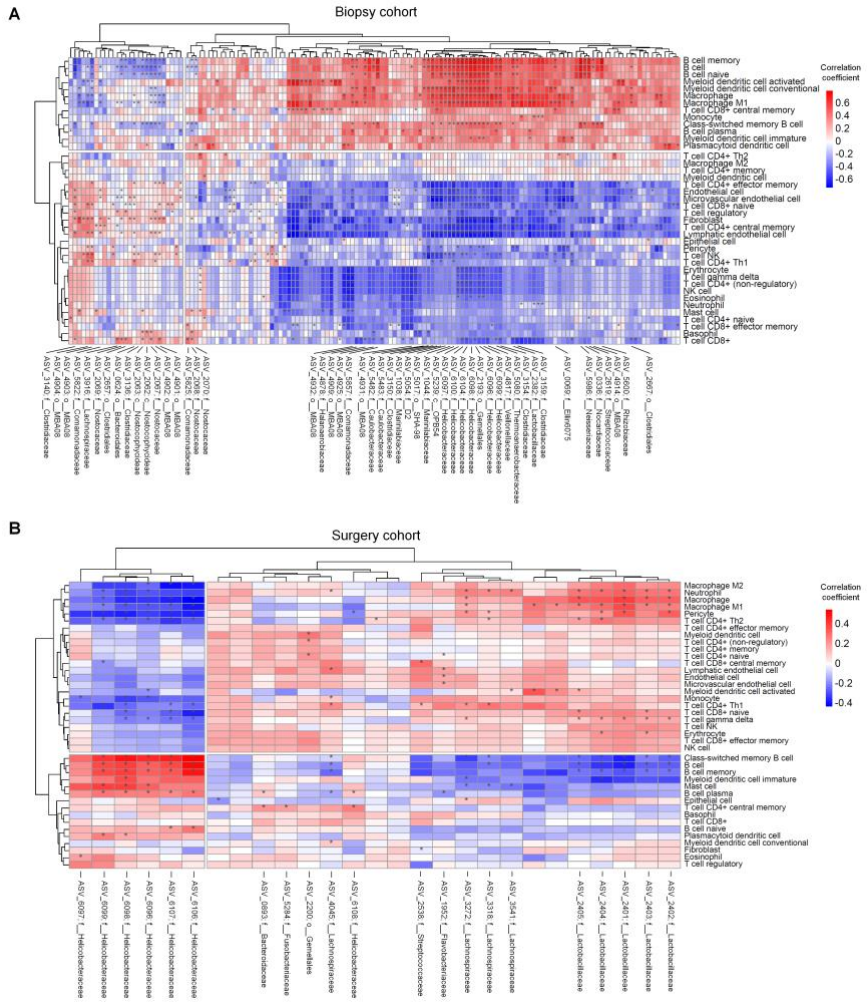

Figure S8. Heatmap for correlations between bacterial ASVs and immune cell types in the biopsy (A) and surgery (B) cohorts, Related to Figures 3 and 6.

In the biopsy cohort, several bacterial taxa, including Helicobacteraceae, Veillonellaceae, and Lactobacillaceae, were correlated with infiltrations of B cells. In contrast, Nostocaceae and Bacteroidales were negatively correlated with B cell infiltration. In comparing gastric cancer and adjacent severe gastritis tissues in the surgery cohort, Helicobacteraceae was positively correlated with B cell infiltration, whereas Lactobacillaceae and Lachnospiraceae were negatively correlated with B cell infiltration. The x-axis indicates bacterial ASVs (ID with taxon). The y-axis indicates immune cell type. The unique identifiers of the ASV IDs are shown in **Table S9**.

*Asterisk denotes significant correlation coefficients (P<0.05 by Spearman-rank correlation test).

ASV, amplicon sequence variant; ID, identification

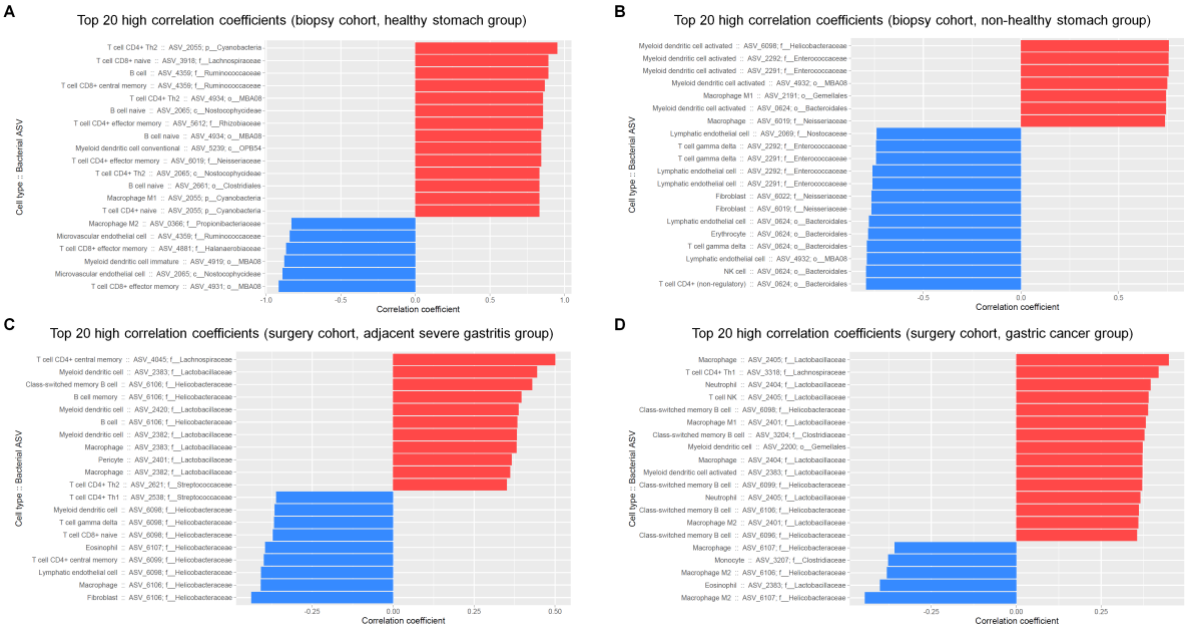

Figure S9. Top 20 high correlation coefficients between bacterial ASVs and cell types according to the clinical phenotype, healthy stomach (A) and non-healthy stomach (B) groups in the biopsy cohort and adjacent severe gastritis (C) and gastric cancer (D) groups in the surgery cohort, Related to Figures 3 and 6.

In the non-healthy stomach group of the biospy cohort, activated myeloid dendritic cells were highly correlated with Helicobacteraceae, Enterococcaceae, and Bacteroidales. Additionally, macrophage was correlated with Gemellales and Neisseriaceae, which were the predominant bacterial taxa in the gastric cancer group. In the gastric cancer group of the surgery cohort, Lactobacillaceae was highly correlated with macrophages and myeloid dendritic cells. Helicobacteraceae was correlated with class-switched memory B cells even in the gastric cancer group of the surgery cohort. The x-axis indicates correlation coefficients between bacterial ASVs and cell types. The y-axis indicates cell types and bacterial ASVs (ID with taxon). The unique identifiers of the ASV IDs are shown in **Table S9**.

ASV, amplicon sequence variant; ID, identification

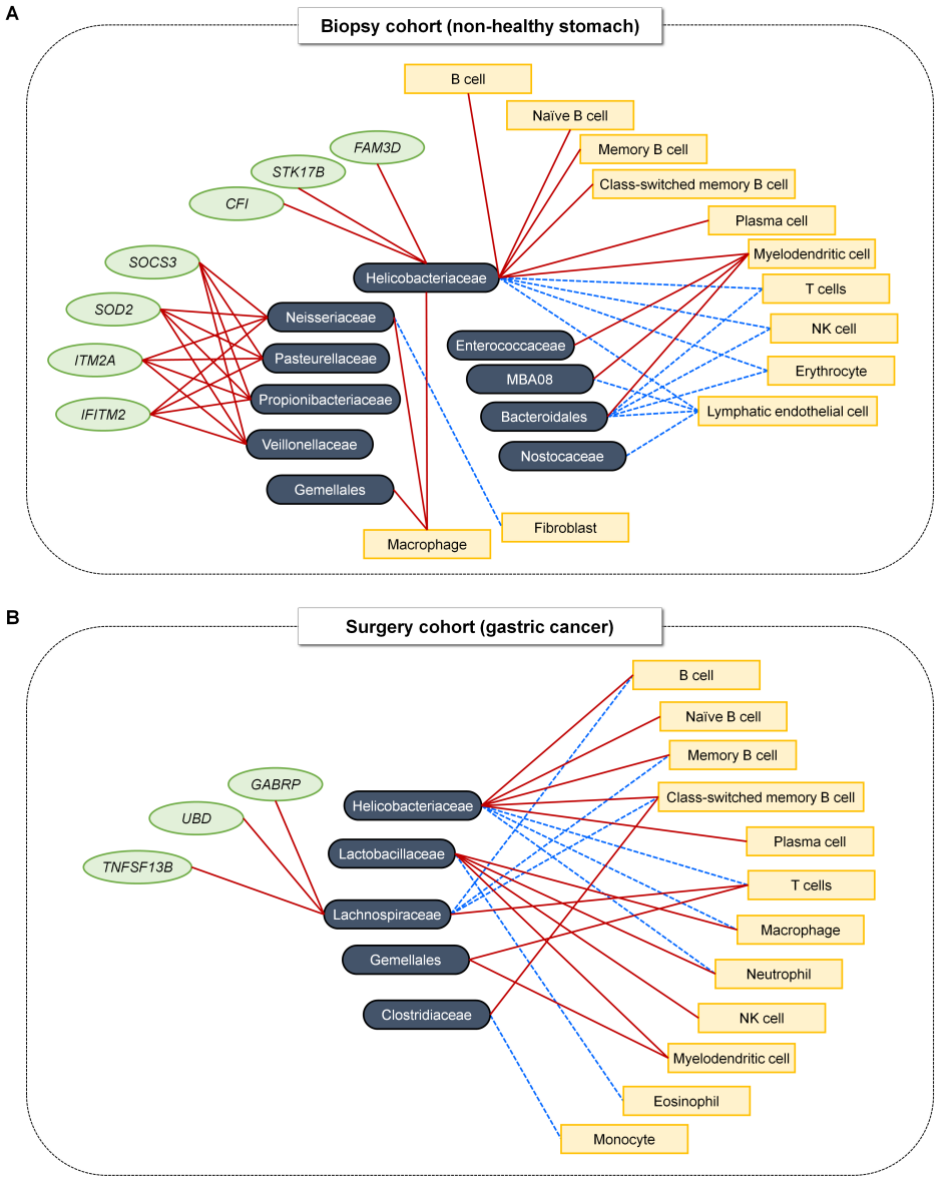

Figure S10. Association network illustrating significant relationships among microbiome, expressed genes, and tumor microenvironment, Related to Figures 4 and 5.

Navy rounded rectangles, green ellipses, and orange rectangles indicate the bacterial taxa, host genes, and infiltrating immune cells in gastric tissues, respectively. Red solid and blue dotted lines represent positive and negative correlations, respectively, found through our correlation analyses. Only representative associations among bacterial taxa, expressed genes, and infiltrated immune cell types are shown.

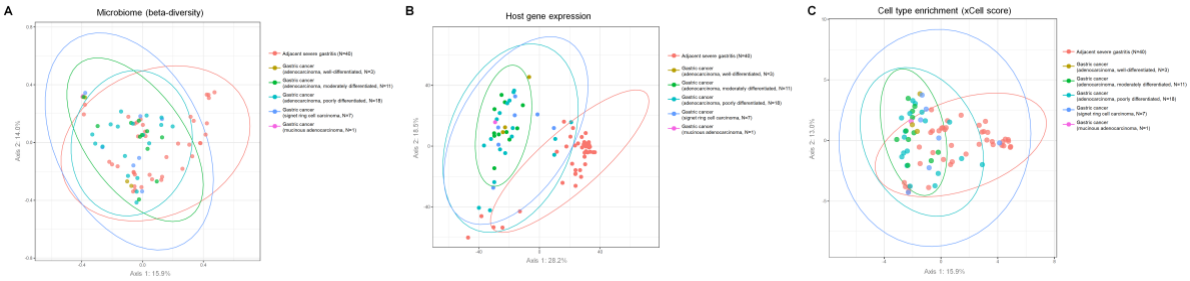

Figure S11. Principal component analysis for microbiome (A), host gene expression (B), and immune cell-type enrichment patterns (C) according to histological subtypes in the surgery cohort, Related to Table 1 and Figures 3 and 6.

Although the adjacent severe gastritis group showed different microbiome, gene expression, and immune cell-type enrichment patterns compared to the gastric cancer group, the histological subtypes of gastric cancer (well-differentiated adenocarcinoma, moderately differentiated adenocarcinoma, poorly differentiated adenocarcinoma, signet ring cell carcinoma, and mucinous adenocarcinoma) did not show observable differences compared to the other subtypes.

Ellipses represent the 95% confidence interval.