ORIGINAL ARTICLE

OPEN ACCESS

Influence of gastrectomy for gastric cancer treatment on faecal microbiome and metabolome profiles

Pande Putu Erawijantari [ID],[1] Sayaka Mizutani [ID],[1,2] Hirotsugu Shiroma [ID],[1]
Satoshi Shiba [ID],[3] Takeshi Nakajima,[4] Taku Sakamoto,[4] Yutaka Saito,[4]
Shinji Fukuda [ID],[5,6,7] Shinichi Yachida [ID],[3,8] Takuji Yamada [ID][1]

ABSTRACT

Objective Recent evidence points to the gut microbiome's involvement in postoperative outcomes, including after gastrectomy. Here, we investigated the influence of gastrectomy for gastric cancer on the gut microbiome and metabolome, and how it related to postgastrectomy conditions.

Design We performed shotgun metagenomics sequencing and capillary electrophoresis time-of-flight mass spectrometry-based metabolomics analyses on faecal samples collected from participants with a history of gastrectomy for gastric cancer (n=50) and compared them with control participants (n=56).

Results The gut microbiota in the gastrectomy group showed higher species diversity and richness (p<0.05), together with greater abundance of aerobes, facultative anaerobes and oral microbes. Moreover, bile acids such as genotoxic deoxycholic acid and branched-chain amino acids were differentially abundant between the two groups (linear discriminant analysis (LDA) effect size (LEfSe): p<0.05, q<0.1, LDA>2.0), as were also Kyoto Encyclopedia of Genes and Genomes modules involved in nutrient transport and organic compounds biosynthesis (LEfSe: p<0.05, q<0.1, LDA>2.0).

Conclusion Our results reveal alterations of gut microbiota after gastrectomy, suggesting its association with postoperative comorbidities. The multi-omic approach applied in this study could complement the follow-up of patients after gastrectomy.

INTRODUCTION

Recent evidence indicates the involvement of the gut microbiome in disease onset and progression and in postoperative outcome.[1–3] A study on the gut mucosal microbiota following ileocolonic resection for Crohn's disease revealed that microbial community structure was associated with disease recurrence or maintenance of remission.[4] Accordingly, microbiota structure might play a role in clinical outcomes.

Several studies have also highlighted gut microbiota alterations following gastrectomy, which represents the primary treatment for gastric cancer and recently for morbid obesity.[5 6] Even though the two procedures have different aims, they share similar anatomical and technical features. Several obesity studies indicate that changes in microbial community and functional potential correlate with

Significance of this study

What is already known on this subject?
► Gastrectomy is a surgical treatment for gastric cancer and morbid obesity.
► Gastrectomy alters physiological properties, such as oxygen availability, pH, food transit time, intestinal motility and hormonal conditions.
► Case studies in obesity treatment have shown that faecal microbiome and metabolome alterations persist over time after surgery; in particular, these alterations correlate with metabolic improvements in patients who are morbidly obese.
► Case studies in gastrectomy for gastric cancer using 16S rRNA amplicon sequencing have shown alterations of gut microbiome composition in patients with subtotal gastrectomy.
► Increased risk of developing metachronous colorectal cancer has been reported in patients with gastric cancer.

What are the new findings?
► The present shotgun metagenomic approach demonstrates overall microbiome community structure changes such as higher abundance of oral microbes, aerobes and facultative anaerobes in faecal samples, which can be related to reconstruction of the GI tract of patients with gastric cancer.
► Gastrectomy-associated alterations in microbial functions, such as nutrient transport and biosynthesis of organic compounds, might relate to changes in postgastrectomy metabolism.
► Several colorectal cancer-related bacteria displayed a similar pattern in postgastrectomy patients: *Fusobacterium nucleatum*, in particular, was significantly enriched (p<0.05) in the total gastrectomy compared with the control group.
► The present metabolomic analysis shows enrichment of deoxycholic acid and branched-chain amino acids in postgastrectomy patients.

weight loss and persistent long-term effects on the gut microbiome.[7–9] Nevertheless, Aron-Wisnewsky *et al* showed that the gut microbiome gene richness

> **How might it impact on clinical practice in the foreseeable future?**
>
> ► The present study provides novel insights into microbiome/metabolome features underlying postgastrectomy comorbidities in patients with gastric cancer; in particular, previously reported occurrence of metachronous colorectal cancer following gastrectomy might be associated with alterations of the gut microbiome.

was partially restored 1 year after gastrectomy, in spite of weight loss and improved metabolism.[10]

Gastrectomy as a curative resection for gastric cancer aims to obtain complete histopathological clearance and involves radical resection of the primary site, as well as resection of affected lymph nodes and adjacent organs if necessary.[6] Two studies using 16S rRNA sequencing have reported microbiome alterations after gastrectomy in patients with gastric cancer. Tseng *et al* revealed that subtotal gastrectomy altered the diversity, community composition and predicted gene functions of gastric microbiota, which associated closely with the altered gastric environment after surgery.[11] Lin *et al* confirmed these findings by analysing the faecal microbiome over the long term after gastrectomy.[12] They showed that patients with subtotal gastrectomy, and in particular Roux-en-Y gastrojejuno anastomosis, had subsequent lower occurrence of type II diabetes and metabolic syndrome than the controls. However, the possible consequences of microbiota alterations on the patients' condition after gastrectomy, especially total gastrectomy, remain poorly understood.

Increasing evidence indicates a possible link between gut microbiota and postoperative outcome after gastrectomy. For example, patients with gastric cancer could be at an increased risk of developing metachronous cancer including colorectal cancer (CRC) after gastrectomy.[13 14] Because the gut microbiota is known to be associated with CRC,[15–17] it may also affect development of metachronous CRC in postgastrectomy patients. Here, we characterised the faecal microbiome of patients with a history of gastrectomy for gastric cancer and compared them with control participants. To provide a better understanding of microbial metabolism, we complemented these data with metabolome profiles. The study aims to comprehensively characterise the influence of gastrectomy for gastric cancer on the gut microbiome, with the intention of improving nutrition and follow-up examinations in postgastrectomy patients.

MATERIALS AND METHODS
Study participants and faecal sample collection
The samples and clinical information used in this study were obtained under conditions of informed consent and with the approval of the institutional review board of each participating institute. A total of 106 participants undergoing total colonoscopy at the National Cancer Center Hospital, Tokyo, Japan, were enrolled. Fifty participants had previously undergone gastrectomy for gastric cancer and did not show signs of gastric cancer recurrence. Those who showed abnormal findings, including precancerous lesions such as adenomatous polyps or carcinomas, were excluded. The remaining 56 participants who did not show any colorectal findings, nor had any history of gastroenterological surgery were included as controls (table 1 and online supplementary tables S1 and S2).

Faecal samples were collected immediately at the first defaecation after starting the oral administration of bowel cleansing

agents at the hospital on the day of colonoscopy.[15 18] The samples were placed directly on dry ice and subsequently stored at −80°C for metagenomic and metabolomic analyses. The participants' lifestyle data, including dietary habits and medical records, were acquired from questionnaires (475 questions, 25 pages), based on the example used in the Japan Public Health Center study[19] (online supplementary figure S1A).

Shotgun metagenomic sequencing and metabolomic quantification
DNA extraction and shotgun metagenomic sequencing of faecal samples (online supplementary methods) generated, on average, 57,888,536 reads per sample and read quality was assessed (online supplementary figure S1B and supplementary table S3). We quantified the metabolites in faecal samples from 44 gastrectomy and 54 control participants using capillary electrophoresis time-of-flight mass spectrometry[20] (online supplementary methods).

Taxonomic profiling
Metagenomic sequencing results were subjected to taxonomic and functional profiling following read quality filtering (online supplementary figure S1B and supplementary methods). Taxonomic profiling using the mOTU pipeline annotated 606 species,[21] whereas 403 species-level clades were obtained using the MetaPhlAn2 pipeline.[22] We categorised the annotated species as oral microbes or others based on the expanded Human Oral Microbiome Database[23] (online supplementary methods). High-level phenotypes were identified by BugBase[24] (online supplementary methods).

Functional profiling
Functional profiles were generated using our *in-house* pipeline (online supplementary methods) and the Human Microbiome Project Unified Metabolic AnalysisNetwork 2 (HUMAnN2) pipeline.[25] Annotation results were then summarised into Kyoto Encyclopedia of Genes and Genomes (KEGG) orthology (KO) and KEGG modules[26] profiles (online supplementary figure S1B and supplementary methods). We annotated 498 KEGG modules that collectively contained 6,108 KOs from our *in-house* pipeline. For the HUMAnN2 pipeline, we used UniRef90 as reference and acquired 437 KEGG modules containing 5,971 KOs.

Statistical analysis
Low-frequency and less abundant microbial features (species, functional modules and metabolites) were discarded (online supplementary methods). Principal coordinates analysis (PCoA) with Bray-Curtis distance was used to examine the separation of species and metabolites across samples. Permutational multivariate analysis of variance (PERMANOVA) ('adonis' function, vegan package in R) was applied to test microbial composition between groups. Chao1 richness and Shannon-Wiener alpha-diversity index were calculated to estimate microbial diversity between gastrectomy and control groups. Differences in relative abundance of the microbial features were determined by linear discriminant analysis (LDA) effect size (LEfSe)[27] (online supplementary methods).

Associations between microbes, KEGG modules or metabolites and clinical parameters (eg, serum glucose, total cholesterol, body mass index (BMI)) and demographic information (eg, age, gender, smoking habits, alcohol consumption, medical history) were assessed using boosted additive generalised linear models available in the MaAsLin R package[28] (online supplementary methods). Significantly different clinical parameters,

Table 1 Participants' characteristics

Characteristic	Gastrectomy	Control	P value
Number of participants	50	56	–
Number of participants with metabolite profiles	44	54	–
Gastrectomy type			
Total gastrectomy (n)	12	–	–
Subtotal gastrectomy (n)	38	–	–
Gastrectomy reconstruction (n)			
Stomach-stomach anastomosis	1	–	–
Billroth I	2	–	–
Jejunal interposition	6	–	–
Pylorus-preserving gastrectomy	8	–	–
Roux-en-Y	29	–	–
Time since surgery (median (range) in years)	5 (1–20)	–	–
Age	61.6±10.7	64.0±11.2	0.205
Body mass index (kg/m^2)	20.6±2.6	23.2±3.0	1.22×10^{-5}
Smokers, n (%)	3 (6.0)	5 (14.3)	0.479
Gender (F/M)	16/34	23/33	0.420
Alcohol consumption (>60 g/day), n (%)	22 (44.0)	31 (55.4)	0.331
Serum glucose (mg/dL)	101.2±12.5	104.6±15.6	0.147
Total cholesterol (mg/dL)	190.4±25.7	209.0±35.8	0.0145
Medications			
Diabetes, n (%)	2 (4.2)	12 (21.8)	0.00889
High blood pressure, n (%)	13 (26.0)	17 (30.4)	0.743
Cholesterol, n (%)	6 (12.0)	10 (17.9)	0.502
Gout, n (%)	0 (0)	2 (3.6)	0.497
Gastric acid suppression, n (%)	4 (8.0)	13 (23.2)	0.0375
Analgesic, n (%)	1 (2.0)	2 (3.6)	1.000
Anticoagulant, n (%)	4 (8.0)	4 (7.1)	1.000
Other, n (%)	8 (16.0)	17 (30.9)	0.108

Statistical test performed: Mann-Whitney U test for numerical data and Fisher's exact test for categorical data; n, number of participants.
F, female; M, male.

demographic characteristics, medical history and diet were tested by a two-sided Mann-Whitney U (MWU) test for numerical data and Fisher exact test for categorical data.

We estimated genus and species association in each group using SparCC[29] (bootstrap n=5000) on selected significantly different genera and species that overlapped between mOTU and Meta-PhlAn2 annotations. Significant co-occurrence and co-excluding interaction (SparCC correlation scores ρ<−0.2 or >0.2 with p<0.05) were visualised and analysed using *igraph*. To determine inter-omics (species and metabolome) correlations, we performed Procrustes analysis (online supplementary methods). Additionally, Spearman's correlation analysis was performed on significantly different metabolites and genera to assess their putative correlation (online supplementary methods). Finally, we applied metabolic model-based integration of metabolite observations and species abundances (MIMOSA)[30] to compare the predicted and measured metabolome trends (online supplementary methods).

RESULTS

Participants' characteristics and microbial community structure of gastrectomy and control groups

Participants' characteristics did not show significant differences between the gastrectomy (n=50) and control (n=56) groups (table 1 and online supplementary table S1). BMI (two-sided MWU test: p=1.22×10^{-5}) and total cholesterol (two-sided MWU test: p=0.0145) were significantly lower in the gastrectomy group, although average cholesterol content was within the normal range (128–219 mg/dL) in each group. Preoperational BMI data for 47 (94%) gastrectomy participants indicated lower BMI at the time of faecal sampling (paired sample t-test: p=7.00×10^{-5}) than prior to surgery, suggesting that weight loss could be attributed to postoperative malnutrition, for example, as a result of dumping syndrome.[31] Fifteen postgastrectomy patients (30%) experienced dumping syndrome.

We further examined the influence of gastrectomy on overall microbiome composition and metabolite profiles in faecal samples by performing PERMANOVA with Bray-Curtis distance (figure 1A and B and online supplementary table S4). Species composition differed significantly between the two groups (adonis: R^2=0.0369, p=9.99×10^{-4}; R^2=0.0387, p=9.99×10^{-4}, in mOTU and MetaPhlAn2, respectively). Within-group Bray-Curtis distance was significantly lower in the gastrectomy group than in the control group (two-sided MWU: p=5.53×10^{-7}) (figure 1C). Metabolite profiles were also significantly different between groups (figure 1D) (adonis: R^2=0.0647, p=9.99×10^{-4}).

PERMANOVA analysis using different surgery types and reconstructions as predictors (online supplementary results and supplementary table S4) revealed no significantly different microbiome and metabolome distributions across different types of surgery (online supplementary figures S2) except for *Fusobacterium nucleatum*. *F. nucleatum* was highly enriched in total gastrectomy compared with control (LEfSe: p=1.53×10^{-5}, q=0.00205, LDA=2.87; p=4.12×10^{-6}, q=3.34×10^{-4},
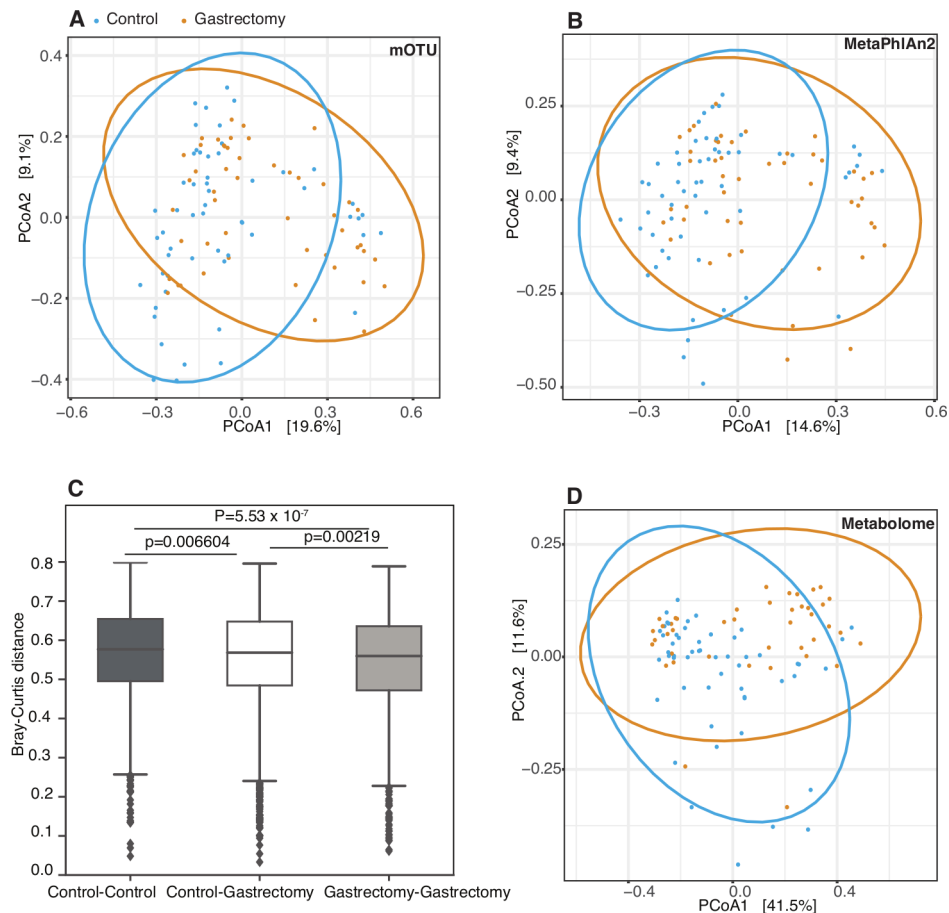
Figure 1 Community structure of the faecal microbiome and metabolome in postgastrectomy and control participants. Principal coordinates analysis (PCoA) with Bray-Curtis distance was performed to assess the community structure of species' relative abundance obtained by mOTU (A) and MetaPhlAn2 (B) in the gastrectomy group (n=50) (orange) and in the control group (n=56) (blue). The PCoA trend was confirmed by significantly lower microbial structure dissimilarity (Bray-Curtis) within groups (p=5.53×10^{-7}) (C). PCoA was performed also on faecal metabolite concentrations in the gastrectomy group (n=44) (orange) and the control group (n=54) (blue) (D).

LDA=1.43, in mOTU and MetaPhlAn2, respectively) and subtotal gastrectomy group (LEfSe: p=5.58×10^{-5}, q=0.0150, LDA=2.34; p=1.45×10^{-5}, q=0.00366, LDA=2.14, in mOTU and MetaPhlAn2, respectively) (online supplementary results and supplementary table S5). Owing to an unbalanced sample size, statistical power may have been insufficient to detect microbiome and metabolome differences across different reconstructions; however, we could clearly observe a different distribution pattern of predominant species and metabolites (online supplementary results and supplementary figures S3 and S4).

To examine the possible confounding effects of clinical parameters (eg, BMI, serum glucose, total cholesterol), demographic data (eg, age, gender) and medical history (eg, medication, diseases), we performed PERMANOVA and MaAsLin. PERMANOVA showed that the participants' grouping explained the variance in species composition and metabolite profiles better than any of the other predictors (online supplementary results and supplementary rable S4). MaAsLin indicated that the association coefficients were not meaningfully affected by adjustment for possible confounders (online supplementary figure S5).

Underlying comorbidities and concurrent medication of the participants may influence gut microbiota. The number of participants using gastric acid secretion inhibitors and diabetes therapeutic drugs varied significantly between the gastrectomy and control groups. Statistical analyses on subsets of the original participants that excluded such users confirmed the lack of any effect of the exclusion on originally reported gastrectomy-enriched microbiome and metabolome signatures (online supplementary results and supplementary tables S6 and S7). The microbial features that were significantly enriched (LEfSe: p<0.05, q<0.1, LDA>2.0) in control individuals who took gastric acid secretion inhibitors or diabetes therapeutic drugs are also reported (online supplementary results and supplementary tables S8 and S9).

Microbial diversity and richness are higher in gastrectomy patients

The gastrectomy group showed higher Chao1 index (mOTU: p=4.64×10^{-4}; MetaPhlAn2: p=1.93×10^{-6}) (figure 2A and B) and Shannon diversity index (mOTU: p=0.0954; MetaPhlAn2: p=0.0104) (figure 2C and D) for species composition than their control counterparts. Higher Shannon diversity indices were conserved across the three major phyla: Firmicutes (mOTU: p=8.98×10^{-7}; MetaPhlAn2: p=2.91×10^{-4}), Actinobacteria (mOTU: p=7.69×10^{-4}; MetaPhlAn2: p=0.0419) and Bacteroidetes (mOTU: p=0.225; MetaPhlAn2: p=0.0283) (figure 2E and F). Our results were in accordance with previous studies, in that species diversity was higher in patients with postgastrectomy gastric cancer[12] and obesity[9 32 33] than in non-surgical participants.
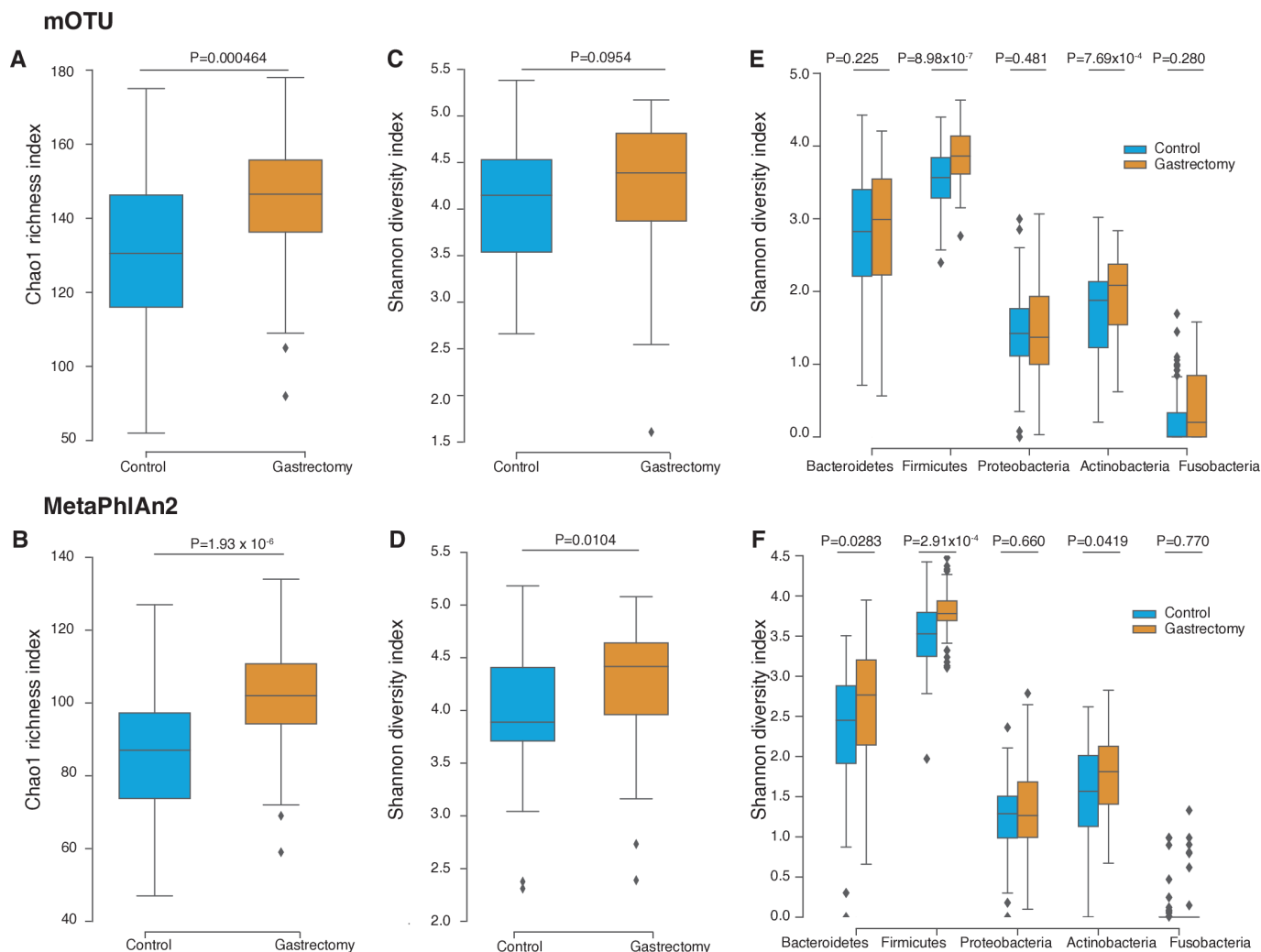
Figure 2 Microbiome diversity in gastrectomy and control groups. Species richness was measured using the Chao1 index calculated from the species annotated by mOTU (A) and MetaPhlAn2 (B). Species alpha-diversity was measured using the Shannon-Wiener index based on mOTU (C) and MetaPhlAn2 (D) annotation. Species alpha-diversity was measured for each major phylum using mOTU (E) and MetaPhlAn2 (F) annotation.

Postgastrectomy patients exhibit a distinct faecal microbiota

Several taxa, mostly Bacilli, were significantly enriched in the gastrectomy group compared with controls (LEfSe: $p<0.05$, $q<0.1$, LDA>2.0) (figure 3A and B). Bacilli were previously reported to be enriched in the gut[12] and gastric microbiota of patients after gastrectomy for gastric cancer.[11] In obesity cases, Bacilli can be used to discriminate between postgastrectomy patients and non-surgical controls.[32]

The mOTU annotation identified 89 species (online supplementary table S10), whereas MetaPhlAn2 annotation identified 76 species (online supplementary table S11) that differed significantly between gastrectomy and control groups (LEfSe: $p<0.05$, $q<0.1$, LDA>2.0), whereas 39 species overlapped between the two methods. The predominant species in postgastrectomy patients were *Streptococcus* (six species), followed by *Prevotella* (four species) and *Veillonella* and *Lactobacillus* (three species, each). Based on MaAsLin analysis of species composition with clinical parameters, enrichment of *Roseburia hominis* and *Ruminococcus gnavus* might be affected by BMI (online supplementary results and supplementary tables S10 and S11). None of the differentially enriched species exhibited significant associations with serum glucose or total cholesterol (MaAsLin: $p>0.05$, $q>0.1$).

Oral microbes (eg, *Streptococcus*, *Veillonella*, *Prevotella*) were relatively more abundant (two-sided MWU test: $p<0.01$) in the gastrectomy group (figure 3C and D), as did also aerobes and facultative anaerobes based on phenotype prediction (figure 3E and F, MWU: $p=6.43\times10^{-7}$, $p=0.0337$, respectively).

Gastrectomy alters microbial functional characteristics

To examine the functional consequences of microbial community changes, we annotated 73 KEGG modules (based on our *in-house* pipeline) and 119 KEGG modules (based on the HUMAnN2 pipeline) that were differentially abundant (LEfSe: $p<0.05$, $q<0.1$, LDA>2.0) between the gastrectomy and control groups (online supplementary table S12 and supplementary figure S6). Among the KEGG modules identified as differentially abundant by both pipelines were those for membrane transport, biosynthesis of organic compounds, multidrug resistance, two-component regulatory system (TCS) and others (figure 4A). In the membrane transport system, phosphate and several amino acid transporters were particularly enriched in the gastrectomy group, followed by manganese/iron and vitamin B_{12} transporters. Our results were in agreement with those reported in a previous obesity study.[9] On the contrary, modules involved
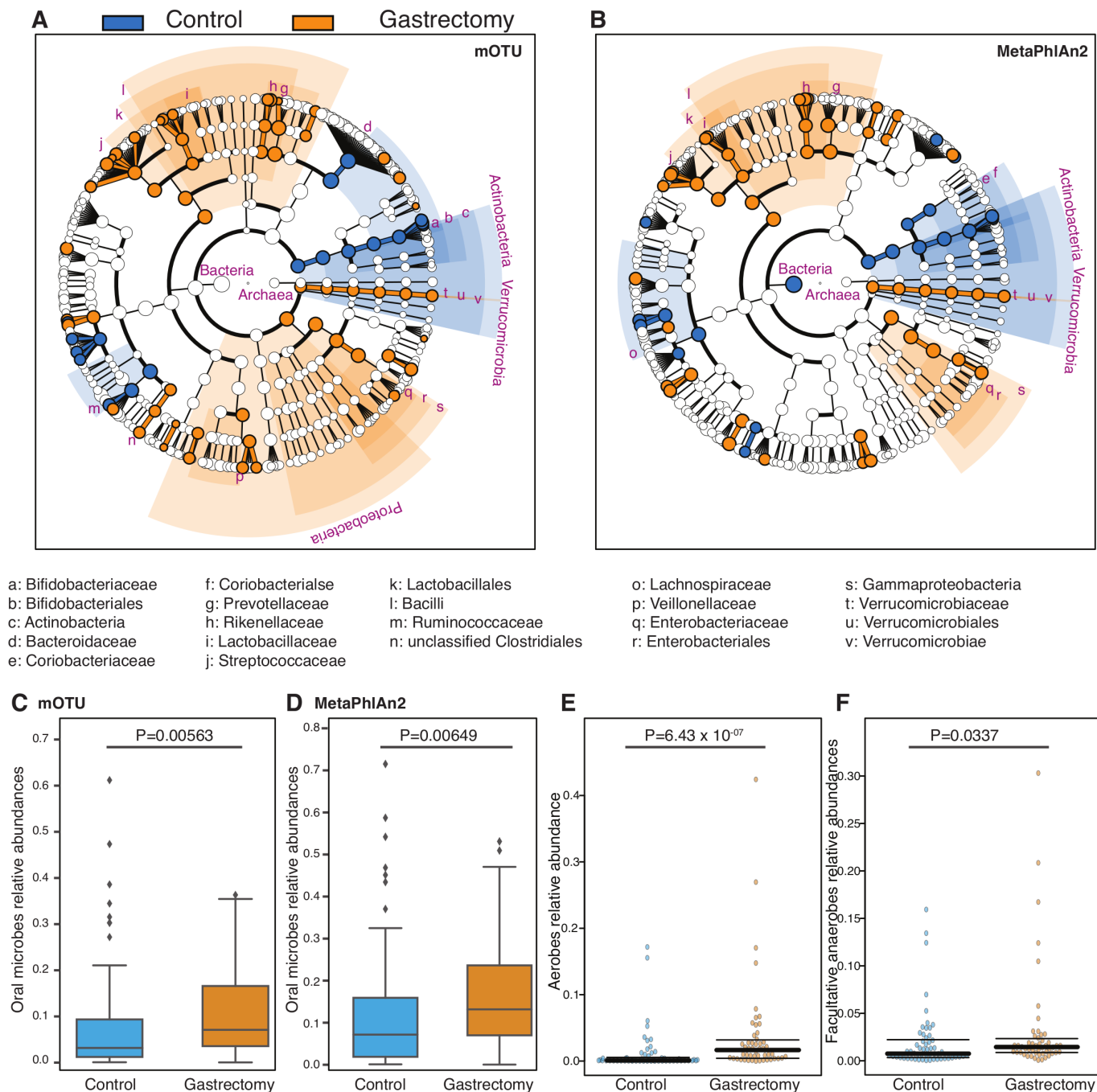
Figure 3 Differential enrichment of microbes in gastrectomy and control groups. Cladogram of species annotated by mOTU (A) and MetaPhlAn2 (B). Each dot represents a taxonomic hierarchy. Dots are marked for significant (LEfSe: p<0.05, q<0.1) enrichment either in the gastrectomy group (n=50) (orange) or in the control group (n=56) (blue). Taxa that reached a linear discriminant analysis score (log10) >3.0 are highlighted and labelled accordingly. The summed relative abundances of oral microbes were compared between the gastrectomy (n=50) and control (n=56) groups based on the species annotated by mOTU (C) and MetaPhlAn (D). The summed relative abundances of aerobes (E) and facultative anaerobes (F) were compared between the two groups.

in raffinose/stachyose/melibiose transport, along with isoleucine and cobalamin (vitamin B_{12}) biosynthesis, were enriched in the controls. We also observed significantly higher abundance of modules related to TCS and multidrug resistance in the gastrectomy group. KEGG modules' association with the participants' demographic characteristics revealed that control or gastrectomy group classification had a predominant impact on module composition (MaAsLin: p<0.05, q<0.1, online supplementary table S12).

To compare module composition between gastrectomy and control groups, we estimated species richness (Chao1) and alpha-diversity (Shannon-Wiener) of each module contributor previously confirmed by our *in-house* and HUMAnN2 pipelines based on the HUMAnN2 output (online supplementary table S13). The three most abundant modules in the control group, raffinose/stachyose/melibiose transport system (M00196), isoleucine biosynthesis (M00535) and cobalamin biosynthesis (M00122), had significantly higher species diversity and
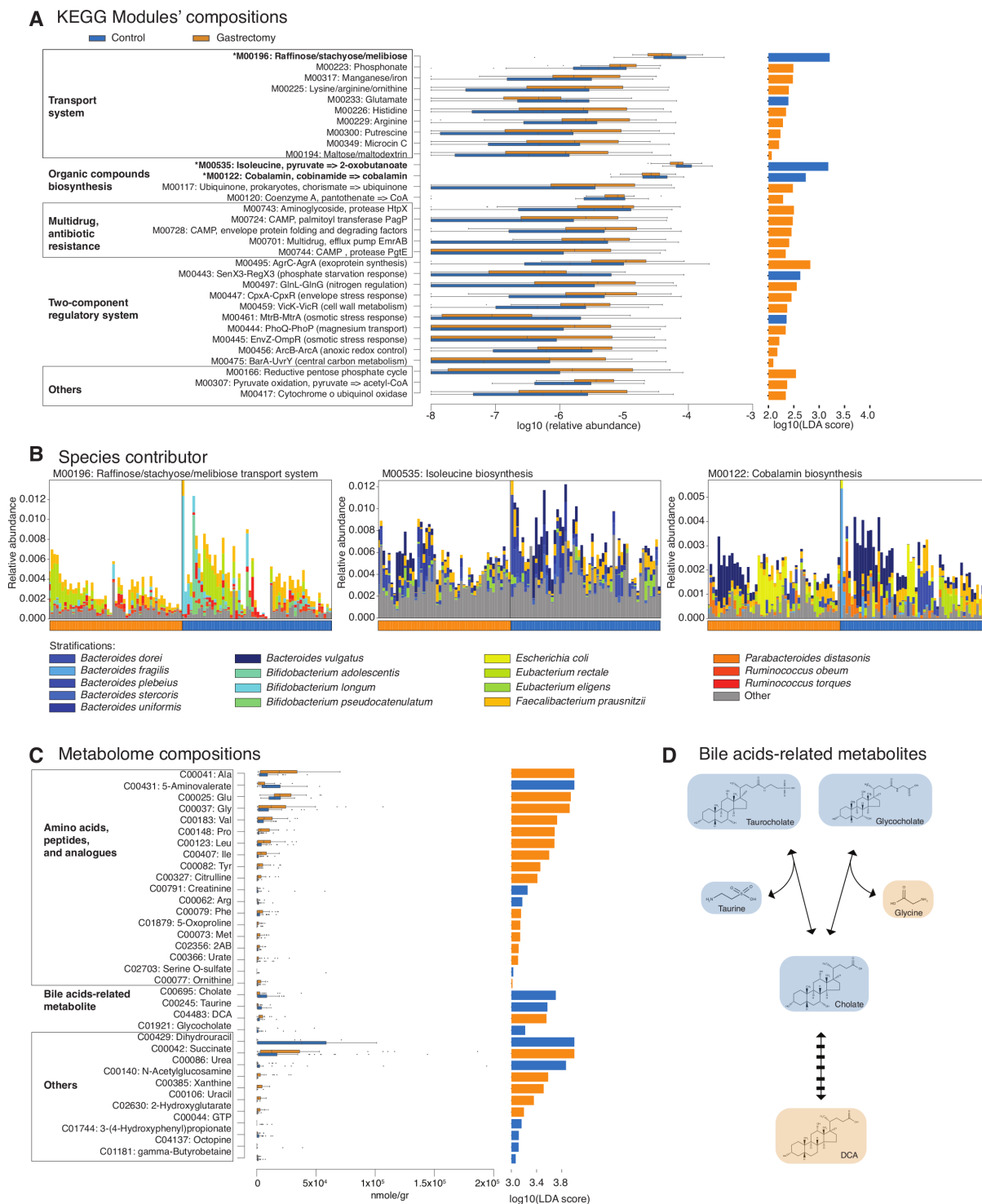
Figure 4 Different trends of functional modules and metabolites from the faecal microbiomes of gastrectomy and control groups. Relative abundance and linear discriminant analysis (LDA) score (log10) of Kyoto Encyclopedia of Genes and Genomes (KEGG) modules annotated by the Human Microbiome Project Unified Metabolic Analysis Network 2 (HUMAnN2) and overlapping with those annotated by our *in-house* pipeline (linear discriminant analysis effect size (LEfSe): p<0.05, q<0.1, LDA>2.0) (A). Richness (Chao1) and alpha-diversity (Shannon-Wiener) of contributor species were estimated. Three modules (marked by an asterisk (*) in A) contributed by significantly more diverse and richer microbes (two-sided Mann-Whitney U test (MWU): p<0.001) in the gastrectomy group, in spite of their enrichment in the control group (B). The modules were M00196, M00535 and M00122 (two-sided MWU: p=5.31×10^{-4}, 3.17×10^{-5}, 7.14×10^{-6}, for richness, respectively, and p=1.12×10^{-4}, 2.22×10^{-4}, 6.10×10^{-4}, for alpha-diversity, respectively). KEGG modules' relative abundance is represented by the top value of each stack of bars. Samples were subsequently sorted according to the dominant contributor to a module and then grouped as either gastrectomy or control (sample order differs between panels). Different trends in metabolites were also observed between the two groups (LEfSe: p<0.05, q<0.1, LDA>3.0) (C). Bile acid reaction consisted of deconjugation of conjugated bile acids (glycocholate and taurocholate) into their primary form (cholate) and amino acids (glycine and taurine) followed by 7-α/β-dehydroxylation to form secondary bile acids (deoxycholic acid (DCA)). The colours highlight enrichment in the control (blue) and gastrectomy (orange) groups (D).

richness in the gastrectomy group (figure 4B and online supplementary table S13). Whereas the dominant species were identical among the two groups, their proportions were different. For example, *Bacteroides vulgatus* contributed predominantly to cobalamin biosynthesis in both groups; however, *Escherichia coli* was a major module contributor in the gastrectomy group (figure 4B). We also detected three different microbes' groups that contributed to the enrichment of three different KEGG modules involved in manganese/iron/zinc/copper transport (online supplementary figure S7). Overall, these findings indicate a microbial community-level shift followed by an alteration in functional potential.

Metabolome profiles change in postgastrectomy patients

We observed different enrichment of metabolites between control (n=54) and gastrectomy (n=44) participants (figure 4C) (52 vs 46 metabolites, respectively; LEfSe: p<0.05, q<0.1, LDA>2.0, online supplementary table S14). Specifically, primary (cholate) and conjugated forms of bile acids (taurocholate and glycocholate) were more abundant in the control group, whereas the secondary form (deoxycholic acid (DCA)) was significantly more abundant in the gastrectomy group (figure 4C and D). Twelve amino acids, including all branched-chain amino acids and two aromatic amino acids (Ile, Leu, Val, Tyr, Phe), were significantly enriched in the gastrectomy group. Metabolites' association with the participants' demographic parameters by MaAsLin showed that cholate enrichment in the control group might be affected by BMI, whereas phenyl-lactate and arginine enrichment might be affected by total cholesterol (online supplementary results and supplementary table S14).

Putative microbes-microbes and microbes-metabolites correlations in each group

Given the notably distinct microbiome composition between gastrectomy and control groups, we compared the topology of genus (figure 5A and B) and species (online supplementary figure S8) co-occurrence and co-excluding networks. In the genus network, the number of edges was higher in the control group (co-occurrence, 25; co-excluding, 15) compared with the gastrectomy group (co-occurrence, 9; co-excluding, 9). Although the edges do not necessarily represent ecological interactions (eg, mutualism, competition), the microbial network can show which organisms form the hub of community composition.[34] Most close relationships ($\rho > 0.6$) in the control group (eg, *Coprobacillus, Eggerthella*) did not appear or showed much lower values ($0.2 < \rho < 0.4$) in the gastrectomy group. In this group, *Veillonella* contributed to most connections (six genera), but co-excluded ($\rho < -0.3$) with four genera including *Coprococcus* ($\rho = -0.569$, $p = 4.00 \times 10^{-4}$) and *Odoribacter* ($\rho = -0.384$, $p = 0.0160$) (online supplementary table S10e). That might reflect the role of *Veillonella* as a network-hub in the gastrectomy group and its importance in microbial community alteration.[34] Results were similar at species level, whereby co-occurrence networks were more abundant in the control (co-occurrence, 31; co-excluding, 8) compared with the gastrectomy group (co-occurrence, 25; co-excluding, 1). Several co-occurrence and co-excluding patterns were shared among genus-level and species-level networks (online supplementary results and supplementary figure S8).

Next, we examined microbiome and metabolite correlations. Procrustes analysis showed an overall significant inter-omics relationship between metabolites and species (mOTU: r=0.279, p=0.003, online supplementary figure S9D; MetaPhlAn2:

r=0.292, p=0.002, online supplementary figure S9E). Bi-clustering of correlations between significantly different genera (LEfSe: p<0.05, q<0.1, LDA>2.0, by mOTU and MetaPhlAn2 annotation) and metabolites (LEfSe: p<0.05, q<0.1, LDA>3.0) revealed distinct clusters (figure 5C). Positive correlation clusters were observed in gastrectomy-enriched metabolites and genera, as well as in control-enriched metabolites and genera. Control-enriched clusters included positive correlations between primary and conjugated bile acids (eg, cholate, taurine, glycocholate) and *Coprobacillus*, *Blautia*, *Eggerthella* and *Bifidobacterium*. In contrast, DCA exhibited a significantly positive correlation with *Alistipes, Odoribacter, Lactobacillus* and *Coprococcus*, which were mutually enriched in gastrectomy patients.

The limited fraction of species contributing to metabolome data (134 out of 326 detected metabolomes) was predicted using MIMOSA based on the KEGG reactions information (online supplementary table S15, supplementary figure S10). In particular, the genus *Roseburia* contributed to the metabolism of bile acids-related metabolites (eg, glycocholate, taurine, cholate) (figure 5C). *Roseburia* was enriched in the gastrectomy group and showed a negative correlation with bile acid metabolites that were enriched in the control group. The prediction indicated a consistent contribution of *Roseburia* to the degradation of those metabolites.

Four (Leu, Ile, Ala, Val) out of nine amino acids formed a cluster and exhibited significant positive correlation with *Atopobium*, *Veillonella* and *Streptococcus*, which were mutually enriched in postgastrectomy patients. *Streptococcus* and *Veillonella* have been reported to ferment amino acids.[35] We also observed the contribution of several gastrectomy-enriched species to amino acids synthesis (eg, *Bacteroides fragilis*) or degradation (eg, *Akkermansia muciniphila*) (online supplementary table S15). Accordingly, a gastrectomy-associated increase in the availability of amino acids might lead to more amino acid fermentators.

DISCUSSION

Gastrectomy followed by reconstruction of the GI tract radically alters oxygen availability, gut pH, food transit time, intestinal motility and hormonal conditions.[36 37] We hereby show that gastrectomy for gastric cancer affects also the faecal microbiome and metabolome. Their links to physiological alterations are summarised as a schematic hypothesis (figure 6). Community PCoA highlighted species-level differences in microbiome composition and metabolite profiles between gastrectomy and control groups. The dissimilarity index within the gastrectomy group was significantly lower than in the control counterpart, demonstrating greater species similarity among postgastrectomy individuals (figure 1).

Higher species richness and diversity in postgastrectomy patients (figure 2) confirms previous reports on both gastric cancer[12] and obesity.[9 32 33] This trend might be explained by a combination of major alterations in the gut environment that could support growth of several microbes. One such change is the presence of more oxygen in the gut after gastrectomy,[38] which may provide a preferable niche for aerobic and facultative anaerobic microbes.[7 9 32] Indeed, relative abundance of aerobes (*Streptococcus* and *Enterococcus*) and facultative anaerobes (*Escherichia*, *Enterobacter* and *Streptococcus*) was higher in postgastrectomy patients compared with control participants (figure 3E and F).

Another possibility might be the migration of oral microbes into the gut. Several microbes frequently detected in the oral
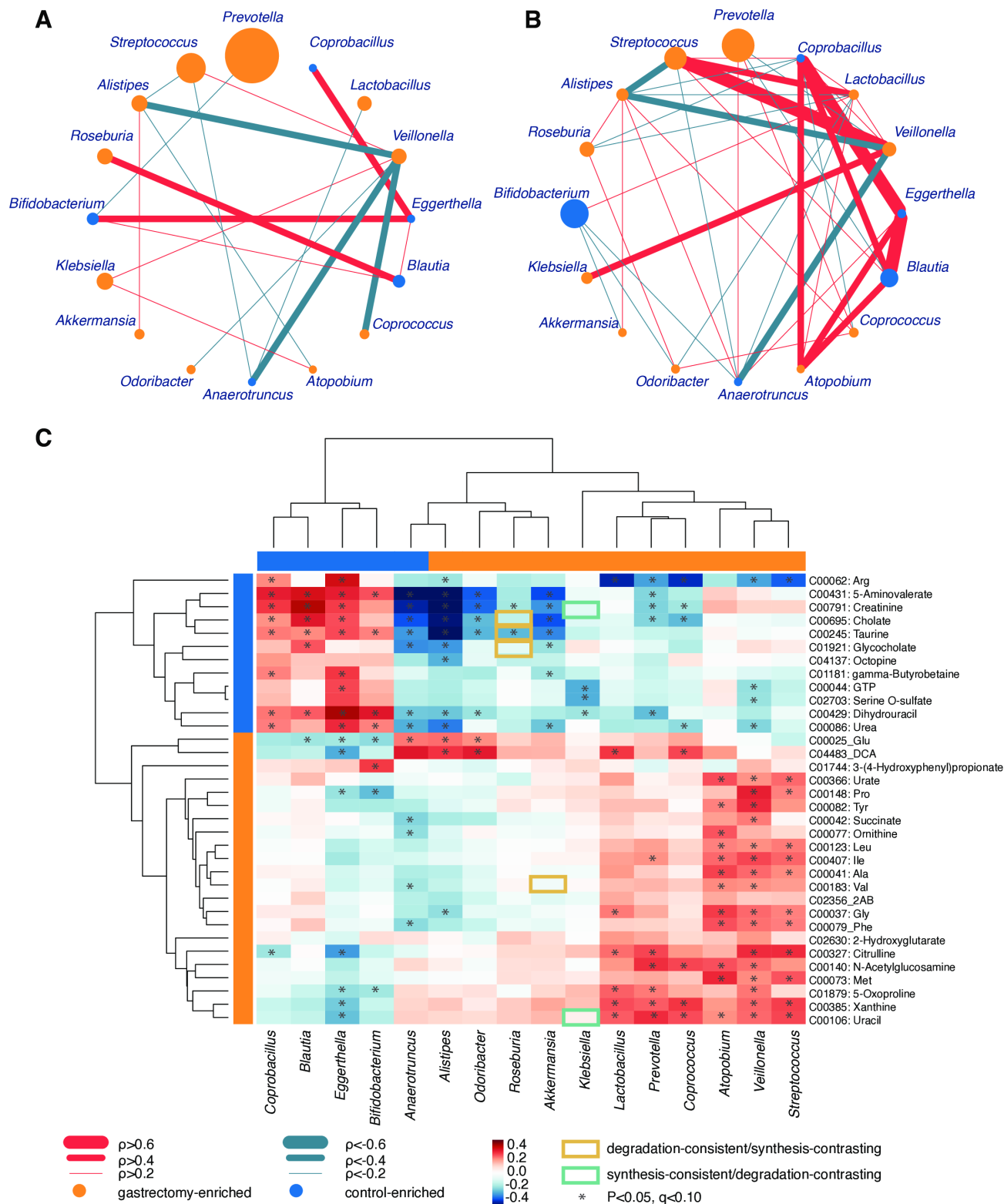
Figure 5 Genus-genus and genus-metabolite correlations. Co-occurrence (red) and co-excluding (green) relationships between genera (SparCC: $-0.2 < \rho < 0.2$, $p < 0.05$) in gastrectomy (n=50) (A) and control (n=56) (B) groups. The edge width corresponds to SparCC correlation coefficients. The nodes' size is scaled based on the genus relative abundance averaged over participants within each group. Nodes' colour represents enrichment of the genus in gastrectomy (orange) and control (blue) participants. Spearman's correlation analysis was performed between differentially abundant genera (linear discriminant analysis effect size (LEfSe): $p < 0.05$, $q < 0.1$, linear discriminant analysis (LDA) >2.0) and metabolites (LEfSe: $p < 0.05$, $q < 0.1$, LDA>3.0) from gastrectomy (n=44) and control (n=54) participants (C). The matrices were derived from Euclidean distance-based bi-clustering of Spearman's RANK correlation matrices. Correlation coefficients in each square represent positive (red) and negative (blue) relationships. Colours are proportional to the absolute value of Spearman's RANK correlations (see legend in the figure). Statistically significant correlations ($p < 0.05$, $q < 0.1$) are marked with asterisks (*). Correlations confirmed by Model-based Integration of Metabolite Observations and Species Abundances are marked by a rectangle (see legend in the figure).
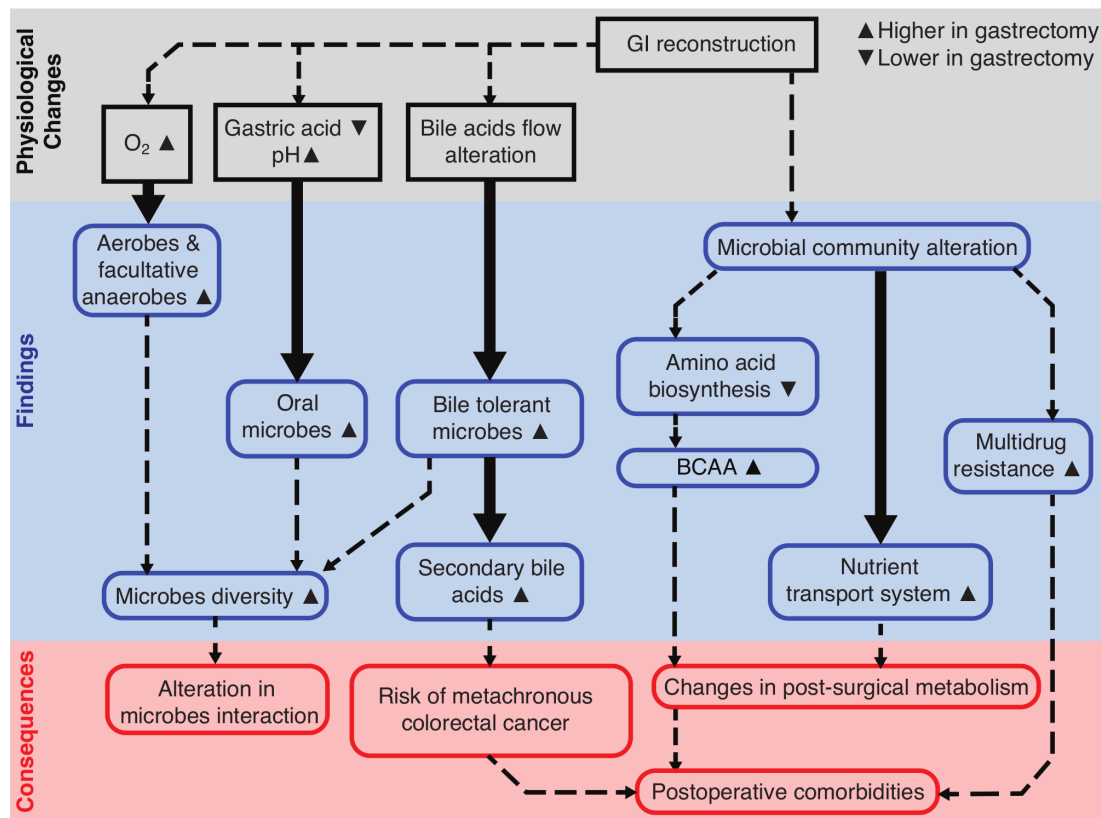
Figure 6 Data integration-derived schematic hypothesis. Schematic hypothesis of gut microbiome and metabolite alterations after gastrectomy. The scheme is divided into three parts denoting physiological changes (grey), findings from our study (blue) and the possible consequences (red). Solid lines show the links confirmed by previous studies. Dashed lines show possible connections that can be inferred from our findings. The higher or lower levels observed in postgastrectomy patients are shown in comparison with those of control participants. BCAA, branched-chain amino acids.

ecosystem[23] were significantly more abundant in postgastrectomy patients (figure 3C and D). They included several species of the genera *Streptococcus*, *Veillonella* and *Prevotella* (online supplementary tables S10 and S11). *Veillonella,* in particular, formed a network-hub in the gastrectomy group, suggesting an important role in microbial community alteration (figure 5A and B, and online supplementary figure S8). Interestingly, *F. nucleatum* was significantly enriched in the total gastrectomy group (n=12), compared with control (n=56) and subtotal gastrectomy (n=38) counterparts. The intestinal tract sustains lower acidity due to reduced gastric acid secretion following gastrectomy.[23 32] These results indicate that gastrectomy promotes higher transport and survival, as well as growth of aerobes, facultative anaerobes and oral microbes in the distal GI tract.[9 32 33]

Previous studies indicated that stomach reconstruction affected microbial functions in the gut.[9–11 32] Metagenome-based functional analysis revealed gastrectomy-associated enrichment of KEGG modules related to nutrient transport system, including manganese/iron/zinc/copper and vitamin B$_{12}$. The obesity study suggested that this feature could reflect the increased availability of those substances to microbes.[9] Here, no significant difference in dietary consumption was observed between the control and gastrectomy groups (online supplementary table S16).

Regarding postoperative metabolism, several studies have reported nutrient intolerance (eg, dumping syndrome) and nutrient deficiency (eg, anaemia) in gastrectomy patients, which could be attributed to malabsorption, impaired food intake and altered transit time, and result in weight loss.[31 39] The observed postoperative reduction in BMI might be associated with these metabolic deficiencies; however, we did not observe

any microbiome or metabolites associated with patients having dumping syndrome (online supplementary results and online supplementary table S17). Furthermore, gastrectomy-associated enrichment of microbial nutrient transport might be associated with the host metabolic functions in two ways. First, postsurgical metabolic deficiencies are at least partially attributed to alterations in gut microbial function (figure 4B and online supplementary figure S7). Second, these changes might be associated with the impaired metabolism of the host. For example, enrichment of the vitamin B$_{12}$ transporter (M00241) in the gastrectomy (online supplementary table S12) and particularly total gastrectomy group (online supplementary table S5) might be related to the malabsorption of vitamin B$_{12}$. This may follow from inadequate secretion of intrinsic factors and reduced gastric acidity in the stomach, which causes a lack of absorption in the terminal ileum following gastrectomy.[40] Consequently, vitamin B$_{12}$ that remain unabsorbed in the colon and lead to more microbes with high vitamin B$_{12}$ intake capacity.[41] In the present study, we observed that the vitamin B$_{12}$ biosynthesis module (M00122) was enriched in the control group. Further comprehensive analysis including the effect of preoperative and postoperative supplemental diet could help decipher these mechanisms.

The present faecal metabolomic analysis also showed that DCA, one of the secondary bile acids, was enriched in postgastrectomy patients, while other forms, such as conjugated and primary bile acids, were enriched in the control counterparts (figure 4C and D). This might be explained by alterations in bile flow after gastrectomy stimulating the growth of bile acid-transforming bacteria. These bacteria play a key role in the deconjugation or transformation of bile acid to secondary

bile acids.[32 42] The process is typically mediated by a 7-α/β-dehydroxylation enzyme, expressed by specific members of *Clostridium* and *Eubacterium* in the human large intestine.[42 43] Although not significant, we confirmed that *Clostridium* and *Eubacterium* genera were comparatively more abundant in postgastrectomy patients (online supplementary tables S10 and S11). DCA is a microbiome-produced carcinogen in liver cancer[44] and CRC.[15 16 45] Obesity studies have raised concerns regarding microbiome changes and secondary bile acids increment following gastrectomy, as well as their link to CRC occurrence.[46 47] Patients with gastric cancer were also reported to be at an increased risk of developing metachronous CRC.[13 14]

Although the mechanism for metachronous CRC onset after gastrectomy may not be similar to sporadic cancer, we observed higher abundances of several CRC-enriched microbes (eg, *F. nucleatum* and *Atopobium parvulum*) in the gastrectomy group. *F. nucleatum* has been suggested to mediate the early steps of carcinogenesis,[15] through FadA adhesion to the epithelium, activation of β-catenin signalling[48] and infiltration of myeloid cells into the tumour microenvironment.[49] Thus, enrichment of *F. nucleatum* in the total gastrectomy group is worth noting. In addition, *A. parvulum* has been associated with multiple polypoid adenomas and intramucosal carcinoma.[15 50] Enrichment of these species suggests a form of dysbiosis, which can lead to CRC development after gastrectomy. In contrast, *Parvimonas micra* and *Peptostreptococcus stomatis*, which are also associated with CRC,[51] were not significantly enriched in postgastrectomy patients (online supplementary tables S10 and S11). This may be partly explained by their relative abundances being high only in the presence of carcinomas.[15]

Of note is also our observation on higher branched-chain amino acids levels and species richness after gastrectomy. Branched-chain amino acids are elevated in plasma and tissue of gastric cancer[52] and CRC.[53] Additionally, increased species richness in CRC has been partly attributed to a greater abundance of oral microbes.[17] Our findings question the supposed association between decreased species richness and intestinal dysbiosis.[54] Altogether, they indicate that although the mechanism of metachronous CRC in postgastrectomy patients may not be exactly the same as that of non-gastrectomy CRC, the similar microbiome contribution can be speculated in the cancer development.

We acknowledge that our study presents also some limitations. Even though our analysis was data-driven, it requires further validation. Future longitudinal or prospective studies with comprehensive medical records including health-related quality of life assessment (eg, based on Gastroenterology Quality of Life Index or Life After Gastric Surgery),[55] intestinal inflammation under colonoscopy, intestinal permeability and nutrition assessment may provide further evidence to support our hypothesis. Additionally, most of our postgastrectomy patients underwent Roux-en-Y reconstruction, so our findings might be representative of this specific reconstruction (online supplementary results and online supplementary table S18).

Our schematic hypothesis (figure 6) could be used as a framework for proof-of-concept studies that link the impact of microbiome alterations with patients' outcome after gastrectomy. Our findings are in many ways consistent with results from other gastrectomy studies. We also identified several microbial functions and metabolites that might correlate with postsurgical metabolism. To determine if the microbiome is indeed involved in CRC occurrence following gastrectomy, further follow-up prospective studies are required. Our analysis also underlines the importance for gastrectomy patients to undergo intensified surveillance such as colonoscopy for early detection of possible metachronous CRC occurrence. To our knowledge, this is the first report on microbiome alterations using metagenomics and metabolomics data analysis after gastrectomy, especially total gastrectomy, for gastric cancer treatment. The present findings may be used to complement a non-invasive method for postsurgical prognosis assessment.

Author affiliations
¹School of Life Science and Technology, Tokyo Institute of Technology, Meguro-ku, Tokyo, Japan
²Research Fellow of Japan Society for the Promotion of Science, Tokyo, Japan
³Division of Cancer Genomics, National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan
⁴Endoscopy Division, National Cancer Center Hospital, Tokyo, Tokyo, Japan
⁵Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan
⁶Intestinal Microbiota Project, Kanagawa Institute of Industrial Science and Technology, Ebina, Kanagawa, Japan
⁷Transborder Medical Research Center, University of Tsukuba, Ibaraki, Japan
⁸Department of Cancer Genome Informatics, Graduate School of Medicine/Faculty of Medicine, Osaka University, Suita, Osaka, Japan

properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Pande Putu Erawijantari http://orcid.org/0000-0002-5173-6347
Sayaka Mizutani http://orcid.org/0000-0003-4168-6195
Hirotsugu Shiroma https://orcid.org/0000-0001-5507-4566
Satoshi Shiba http://orcid.org/0000-0001-5159-5719
Shinji Fukuda http://orcid.org/0000-0001-5161-9880
Shinichi Yachida https://orcid.org/0000-0001-5507-4566
Takuji Yamada http://orcid.org/0000-0002-9622-1849

REFERENCES

1 Guyton K, Alverdy JC. The gut microbiota and gastrointestinal surgery. *Nat Rev Gastroenterol Hepatol* 2017;14:43–54.

2 Lederer A-K, Pisarski P, Kousoulas L, *et al*. Postoperative changes of the microbiome: are surgical complications related to the gut flora? A systematic review. *BMC Surg* 2017;17:125.

3 Gutiérrez-Repiso C, Moreno-Indias I, de Hollanda A, *et al*. Gut microbiota specific signatures are related to the successful rate of bariatric surgery. *Am J Transl Res* 2019;11:942–52.

4 Mondot S, Lepage P, Seksik P, *et al*. Structural robustness of the gut mucosal microbiota is associated with Crohn's disease remission after surgery. *Gut* 2016;65:954–62.

5 Bennett JMH, Mehta S, Rhodes M. Surgery for morbid obesity. *Postgrad Med J* 2007;83:8–15.

6 Penna M, Allum W. New treatments for gastric cancer: are they changing clinical practice? *Clinical Practice* 2013;10:649–59.

7 Zhang H, DiBaise JK, Zuccolo A, *et al*. Human gut microbiota in obesity and after gastric bypass. *Proc Natl Acad Sci U S A* 2009;106:2365–70.

8 Graessler J, Qin Y, Zhong H, *et al*. Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters. *Pharmacogenomics J* 2013;13:514–22.

9 Palleja A, Kashani A, Allin KH, *et al*. Roux-En-Y gastric bypass surgery of morbidly obese patients induces swift and persistent changes of the individual gut microbiota. *Genome Med* 2016;8:67.

10 Aron-Wisnewsky J, Prifti E, Belda E, *et al*. Major microbiota dysbiosis in severe obesity: fate after bariatric surgery. *Gut* (Published Online First: 13 June 2018).

11 Tseng C-H, Lin J-T, Ho HJ, *et al*. Gastric microbiota and predicted gene functions are altered after subtotal gastrectomy in patients with gastric cancer. *Sci Rep* 2016;6.

12 Lin X-H, Huang K-H, Chuang W-H, *et al*. The long term effect of metabolic profile and microbiota status in early gastric cancer patients after subtotal gastrectomy. *PLoS One* 2018;13:e0206930.

13 Eom BW, Lee H-J, Yoo M-W, *et al*. Synchronous and metachronous cancers in patients with gastric cancer. *J Surg Oncol* 2008;98:106–10.

14 Ikeda Y, Saku M, Kawanaka H, *et al*. Features of second primary cancer in patients with gastric cancer. *Oncology* 2003;65:113–7.

15 Yachida S, Mizutani S, Shiroma H, *et al*. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat Med* 2019;25:968–76.

16 Wirbel J, Pyl PT, Kartal E, *et al*. Meta-Analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat Med* 2019;25:679–89.

17 Thomas AM, Manghi P, Asnicar F, *et al*. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat Med* 2019;25:667–78.

18 Nishimoto Y, Mizutani S, Nakajima T, *et al*. High stability of faecal microbiome composition in guanidine thiocyanate solution at room temperature and robustness during colonoscopy. *Gut* 2016;65:1574–5.

19 Shin S, Saito E, Sawada N, *et al*. Dietary patterns and colorectal cancer risk in middle-aged adults: A large population-based prospective cohort study. *Clinical Nutrition* 2018;37:1019–26.

20 Soga T, Ohashi Y, Ueno Y, *et al*. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J Proteome Res* 2003;2:488–94.

21 Sunagawa S, Mende DR, Zeller G, *et al*. Metagenomic species profiling using universal phylogenetic marker genes. *Nat Methods* 2013;10:1196–9.

22 Truong DT, Franzosa EA, Tickle TL, *et al*. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods* 2015;12:902–3.

23 Chen T, Yu W-H, Izard J, *et al*. The human oral microbiome database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* 2010;2010:baq013.

24 Langille MGI, Zaneveld J, Caporaso JG, *et al*. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 2013;31:814–21.

25 Franzosa EA, McIver LJ, Rahnavard G, *et al*. Species-Level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 2018;15:962–8.

26 Kanehisa M, Goto S. Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;28:27–30.

27 Segata N, Izard J, Waldron L, *et al*. Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60.

28 Morgan XC, Tickle TL, Sokol H, *et al*. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;13:R79.

29 Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput Biol* 2012;8:e1002687.

30 Noecker C, Eng A, Srinivasan S, *et al*. Metabolic model-based integration of microbiome taxonomic and metabolomic profiles elucidates mechanistic links between ecological and metabolic variation. *mSystems* 2016;1.

31 Radigan AE. Post-Gastrectomy: managing the nutrition fallout. *Pract Gastroenterol* 2004;28:63–79.

32 Ilhan ZE, DiBaise JK, Isern NG, *et al*. Distinctive microbiomes and metabolites linked with weight loss after gastric bypass, but not gastric banding. *Isme J* 2017;11:2047–58.

33 Kong L-C, Tap J, Aron-Wisnewsky J, *et al*. Gut microbiota after gastric bypass in human obesity: increased richness and associations of bacterial genera with adipose tissue genes. *Am J Clin Nutr* 2013;98:16–24.

34 Röttjers L, Faust K. From hairballs to hypotheses—biological insights from microbial networks. *FEMS Microbiol Rev* 2018;42:761–80.

35 Dai Z-L, Wu G, Zhu W-Y. Amino acid metabolism in intestinal bacteria: links between gut ecology and host health. *Front Biosci* 2011;16:1768–86.

36 Aron-Wisnewsky J, Doré J, Clement K. The importance of the gut microbiota after bariatric surgery. *Nat Rev Gastroenterol Hepatol* 2012;9:590–8.

37 Quercia I, Dutia R, Kotler DP, *et al*. Gastrointestinal changes after bariatric surgery. *Diabetes Metab* 2014;40:87–94.

38 Celiker H. A new proposed mechanism of action for gastric bypass surgery: air hypothesis. *Med Hypotheses* 2017;107:81–9.

39 Lim J, Yoo M-W, Kang SY, *et al*. Long-Term changes in the metabolic and nutritional parameters after gastrectomy in early gastric cancer patients with overweight. *Asian J Surg* 2019;42:386–93.

40 Jun J-H, Yoo JE, Lee JA, *et al*. Anemia after gastrectomy in long-term survivors of gastric cancer: a retrospective cohort study. *International Journal of Surgery* 2016;28:162–8.

41 Degnan PH, Taga ME, Goodman AL. Vitamin B 12 as a modulator of gut microbial ecology. *Cell Metab* 2014;20:769–78.

42 Liu H, Hu C, Zhang X, *et al*. Role of gut microbiota, bile acids and their cross-talk in the effects of bariatric surgery on obesity and type 2 diabetes. *J Diabetes Investig* 2018;9:13–20.

43 Ridlon JM, Kang D-J, Hylemon PB. Bile salt biotransformations by human intestinal bacteria. *J Lipid Res* 2006;47:241–59.

44 Yoshimoto S, Loo TM, Atarashi K, *et al*. Obesity-Induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* 2013;499:97–101.

45 Ajouz H, Mukherji D, Shamseddine A. Secondary bile acids: an underrecognized cause of colon cancer. *World J Surg Oncol* 2014;12:164.

46 Hull MA, Markar SR, Morris EJA. Cancer risk after bariatric surgery — is colorectal cancer a special case? *Nat Rev Gastroenterol Hepatol* 2018;15:653–4.

47 Mackenzie H, Markar SR, Askari A, *et al*. Obesity surgery and risk of cancer. *Br J Surg* 2018;105:1650–7.

48 Rubinstein MR, Wang X, Liu W, *et al*. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-Cadherin/β-Catenin signaling via its FadA adhesin. *Cell Host Microbe* 2013;14:195–206.

49 Kostic AD, Chun E, Robertson L, *et al*. Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* 2013;14:207–15.

50 Kasai C, Sugimoto K, Moritani I, *et al*. Comparison of human gut microbiota in control subjects and patients with colorectal carcinoma in adenoma: terminal restriction fragment length polymorphism and next-generation sequencing analyses. *Oncol Rep* 2016;35:325–33.

51 Yu J, Feng Q, Wong SH, *et al*. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* 2017;66:70–8.

52 Chan AW*et al*. Potential role of metabolomics in diagnosis and surveillance of gastric cancer. *WJG* 2014;20:12874–82.

53 Denkert C, Budczies J, Weichert W, *et al*. Metabolite profiling of human colon carcinoma – deregulation of TCA cycle and amino acid turnover. *Mol Cancer* 2008;7:72.

54 Le Chatelier E, Nielsen T, Qin J, *et al*. Richness of human gut microbiome correlates with metabolic markers. *Nature* 2013;500:541–6.

55 Shan B, Shan L, Morris D, *et al*. Systematic review on quality of life outcomes after gastrectomy for gastric carcinoma. *J Gastrointest Oncol* 2015;6:544–60.

A

Participants (n=106)

Questionnaire

Fecal samples

CE-TOFMS
- Control
  (n=54)
- Gastrectomy
  participants (n=44)

Whole shotgun
metagenomic sequencing
- Control (n=56)
- Gastrectomy participants
  (n=50)

Metabolite profile

Reads quality filtering

Taxonomic profiling

Functional profiling

B

Raw reads

1.Reads quality filtering

Elimination of reads which include N

Elimination of Phix sequences

Trimming of adapter sequences

Length filtering (> 50 bp)

Quality filtering (average > 25)

Elimination of human genome sequence

Elimination of unpaired reads

High-quality reads

mOTU
catalog

RefMG
catalog

Calculate and summarize
gene abundances

Taxonomic profile

mOTU pipeline

Map to MetaPhlAn2
reference

Taxonomic profile

MetaPhlAn2 pipeline

HUMAnN2 pipeline
(based on UniRef90)

UniRef90 gene profile

Annotation with KEGG
Orthology (KO)

KO profile

KO module

HUMAnN2 pipeline

Assembly

Scaffolds

Mapping

Gene prediction

Coverage of scaffolds

Predicted genes

Calculation of
gene abundance

Alignment with
KEGG gene

Coverage of genes

KEGG gene profile

Annotation with KEGG
Orthology (KO)

KO profile

KO module

In-house pipeline

1   **SUPPLEMENTARY METHODS**

2

3   **DNA extraction, shotgun metagenomic sequencing, and quality control**

4   We extracted genomic DNA from the frozen fecal samples by the bead-beating method

5   using the GNOME® DNA Isolation Kit (MP Biomedicals, Santa Ana, CA, USA). The quality

6   was assessed on a 4200 TapeStation (Aglient Technologies, Santa Clara, CA, USA). The

7   extracted DNAs were then subjected to shotgun metagenomic sequencing on an Illumina

8   HiSeq2500 platform with 150-bp paired-end read lengths and a targeted 5 Gbp

9   sequencing depth. The sequencing libraries were generated with the Nextera XT DNA

10   Library Prep Kit (Illumina, San Diego, CA, USA). The quality of the libraries was analyzed

11   by the 4200 TapeStation.

12

13   **Metabolomics quantification of fecal samples**

14   Fecal metabolites were extracted from ten milligrams of fresh thawed freeze-dried fecal

15   samples suspended in 400 μL of 50% methanol in Milli-Q containing internal standards

16   (20 μM each of methionine sulfone and D-camphor-10-sulfonic acid (CSA)[1]. The 3-mm

17   zirconia beads (BioSpec Products, Bartlesville, OK, USA) and 100 mg of 0.1-mm

18   zirconia/silica beads (BioSpec Products, Bartlesville, OK, USA) were added into the

19   mixture then subjected to vigorous shaking using Micro Smash (TOMY, Nerima, Tokyo,

20   Japan). The suspensions were centrifuged for 10 minutes at 1500 rpm, 10 times. The

21   supernatant was transferred to a 5-kDa-cutoff filter column (Ultrafree MC-PLHCC

22   250/pk) for Metabolome Analysis (Human Metabolome Technologies, Tsuruoka,

23   Yamagata, Japan). The flow-through was dried under vacuum and the residue then was

24   dissolved in 40 μL of Milli-Q water containing reference compounds (200 μM each of 3-

25   aminopyrrolidine and trimesate). The levels of extracted metabolites were measured in

26   both positive and negative modes by Capillary Electrophoresis Time-of-Flight Mass

27   Spectrometry (CE-TOFMS) as previously described[2]. The CE-TOFMS experiments were

28   carried out using an Aglient CE Capillary Electrophoresis System (Agilent Technologies,

29   Santa Clara, CA, USA).

30   The raw data were processed for metabolite quantification using automatic

31   integration software MasterHands (ver. 2.16.0.15)[3]. We generated annotations tables

32   based on the measurement of standard compounds and aligned with the datasets

33   according to the similar *m/z* values and normalized migration time. The peak areas were

1

34   then normalized against those of the internal standards methionine sulfone and CSA for

35   cationic and anionic metabolites, respectively. Concentrations of each metabolite were

36   calculated based on their relative peak areas and the concentrations of the standard

37   compounds.

38   For the analysis, concentrations below the detection limit were substituted with

39   zero, and metabolites for which levels were below the detection limit in all of the samples

40   were excluded. The metabolite concentration (nmole) was normalized using the fecal

41   weight to obtain the amount of metabolite in each gram of sample (nmole/g)[1,2].

42

43   **Reads quality filtering**

44   Shotgun metagenomics sequencing was performed on an Illumina HiSeq2500 platform

45   with 150-bp read length to a targeted 5 Gbp sequencing depth. The sequence quality

46   filtering process involved several steps and started by removing reads that contained

47   ambiguous bases (**Supplementary Figure S1B**). The elimination of reads with *PhiX* DNA

48   contamination was then performed, followed by trimming the sequencing adapter and

49   3'-end low quality reads using cutadapt (version 1.9.1)[4]. The adapter sequences were

50   removed   by   cutadapt   using   the   option   of   "-a

51   CTGTCTCTTATACACATCTCCGAGCCCACGAGAC -o 33" for the forward primer sequence

52   and "-a CTGTCTCTTATACACATCTGACGCTGCCGACGA -o 32" for the reverse primer

53   sequence.

54   We continued the filtering process by applying a read length and average quality

55   cut-off. The reads of less than 50 bp in length and an average quality score of less than 25

56   were removed. Furthermore, contaminating human DNA sequences were removed by

57   screening   against   the   human   genome

58   (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/001/405/GCA_000001405.15_GRCh

59   38/GCA000001405.15_GRCh38_assembly_structure/Primary_Assembly/assembled_chr

60   omosomes/FASTA/). We used Bowtie2 (version 2.2.9)[5] to eliminate reads with *PhiX*

61   and human genome contamination. Only paired-end reads were selected as high-quality

62   reads for further analysis.

63

64   **Taxonomic profiling by the mOTU and MetaPhlAn2 pipeline**

65   Taxonomic profiling was performed using the metagenomics operational taxonomic

66   units (mOTUs)[6] and MetaPhlAn2 pipeline[7] that will be explained below.

2

67 ### *mOTU pipeline*

68 There were two processes involved in the abundance profiling of samples. First,
69 the high-quality reads were subjected to mapping against the mOTU.v1.padded database
70 using sequence identity and alignment cutoffs of 97% and 45 bp, respectively. The
71 mOTU.v1.padded database contains the marker gene (MG) sequences that have been
72 extracted from 3,496 reference genomes and 263 published human gut metagenomic
73 samples. The counts of respective mOTUs were increased by one for each read that
74 mapped to one or more MGs that belonged to the same mOTU. For each read that mapped
75 to MGs from *n* different mOTUs with the same alignment scores, the count of respective
76 mOTUs was increased by the fraction of unique mappers of these mOTUs. The reads
77 counts were then normalized to the gene length, scaled by the average gene length, and
78 rounded down to get the mOTU abundances.

79 Second, the taxa were annotated based on the RefMG.v1.padded database. The
80 RefMG.v1.padded database is a subset of the mOTU.v1.padded database, and contains
81 only MGs from 3,496 NCBI reference genomes. Taxa were merged if their NCBI species
82 annotations were the same. The final taxonomic abundance profiles were then stored in
83 the form of tab-delimited files with the sample IDs as the header and the NCBI species ID
84 as the rows. The relative abundances were calculated by dividing the relative abundances
85 by the total sum of relative abundances so that all taxa relative abundances totaled 1
86 (**Supplementary Figure S1B**). Species with average relative abundance exceed 0.001%
87 and appeared in at least 5% of samples number (five samples) were retrieved for
88 downstream analysis.

89

90 ### *MetaPhlAn2 pipeline*

91 In addition to mOTU annotation, we performed taxonomic annotation using MetaPhlAn2
92 pipeline version 2.7.0 with default parameters[7]. High-quality reads were mapped to
93 unique clade-specific marker genes identified from ~17,000 reference genomes
94 (~13,500 bacterial and archaeal, ~3,500 viral, and ~110 eukaryotic). The output listed
95 the relative abundance for detected species level. We only used the bacterial with average
96 relative abundance that were exceed 0.1% in at least 5% of samples number (five
97 samples) for our downstream analysis.

98

99

3

Functional annotation of fecal metagenomes by *in-house* and HUMAnN2 pipeline

We performed our *in-house* and HUMAnN2[8] for functional annotations that will be explained below. We employed the pipeline to generate the KEGG (Kyoto Encyclopedia of Genes and Genome orthology) Orthology (KO) profile[9]. KO abundances were then calculated for KEGG modules, a collection of manually defined functional units, using omixer-rpm (default parameter)[10] that will select the modules that pass the defined coverage (number of observed steps/number of defined steps) cutoff, The abundance were derived for each modules by selecting the combination of KO that maximize the KEGG modules abundance. The KEGG modules with average relative abundance of 0.0001% and appeared in at least 5% of samples number (five samples) were selected for the downstream analysis.

In-house *pipeline*

Our *in-house* pipeline for functional annotation was written in Python 2.7 integrated with several tools that are publicly available. The high-quality reads were subjected to the assembly process using the IDBA-UD assembler (version 1.1.1)[11]. To obtain the scaffold for gene prediction, the parameter --mink 20 -maxk 120 --step 10 was used. This parameter meant that the *k* values ranged from 20 to 120, with an increase in the *k*-mer every 10 iterations.

Genes were predicted by MetaGeneMark (version 3.26) using the bacterial, archaeal, and plant plastid genetic codes as the protein translation options and the parameters for metagenome gene prediction (-g 11 –m MetaGeneMark_v1.mod)[12]. The predicted genes with more than 50 bp of amino acid sequences were selected for gene annotation. We mapped our predicted genes to the Kyoto Encyclopedia of Genes and Genomes (KEGG) GENES database (as of 2017)[9] using DIAMOND[13] (version 0.9.10) (cut-offs: sequence identity >40, bit score >70, coverage >80). The top hit with the highest score was selected from the annotated genes, since one predicted gene can be annotated to more than one gene. The top hit genes were then filtered by applying the cut-off values of 40% identity and a score of 70. The final top hit of annotated gene abundance was calculated by the number of hits weighing process. The calculation of predicted gene abundance was started by determining the position of the high-quality reads within the scaffold. The high-quality reads were mapped back to the scaffold using Bowtie2 (version 2.2.9)[5]. The number of bases in the reads that mapped in the positions of the predicted genes were then summed to calculate the annotated gene abundances. The number of

4

133 bases was then divided by the length of the predicted genes to determine the gene

134 abundances. In other words, the predicted gene abundance was defined by the number

135 of reads that mapped to the predicted gene. After the abundance of the annotated gene

136 was calculated, total sum scaling normalization (TSS) was performed to account for

137 uneven sequencing depth across samples. The abundance value was divided by the total

138 abundance per sample to generate its relative abundance. The KEGG were then annotated

139 into the KEGG orthologous (KO) group profile[14].

140 ***HUMAnN2 pipeline***

141 Functional profiling was performed using HUMAnN2 v0.11.1 in UniRef90

142 mode[8]. High-quality reads were initially mapped to the pangenomes species identified

143 during the taxonomic profiling by MetaPhlAn2 using Bowtie2. The pangenomes have

144 been pre-annotated to its respective UniRef90 families. Subsequently the unmapped

145 reads were mapped to UniRef90 by translated search using DIAMOND. The gene-level

146 outputs were produced in reads per kilobases units and stratified according to

147 known/unclassified community contribution to relative abundance unit. The gene

148 relative abundance was regrouped into its respective KO using the UniRef90-KEGG linked

149 file.

150

151 **Microbial community structure analysis**

152 To visualize the microbial community structure between the gastrectomy and the control

153 group, we calculated between-sample diversity score (Bray-Curtis distance) using the

154 relative abundances of taxa and the concentration of metabolites. We visualized the

155 separation using the Principal coordinates analysis (PCoA). The distance calculation and

156 PCoA analysis were performed by R package phyloseq[15]. Subsequently, we performed

157 permutational multivariate analysis of variance (PERMANOVA) ('adonis' function, vegan

158 package, R) on the matrices between-sample diversity score to test the differences of

159 microbial community between group. We also performed PERMANOVA using the other

160 predictor which are the medical history that were extracted from the questionnaire and

161 the clinical parameters (BMI, serum glucose and total cholesterol) from participants

162 records. For the medical history, we marked the participants who had a history of other

163 diseases or history of medications before the sample collections as "Yes". We excluded

164 the samples who did not have the data for particular parameter (label as NA) in the

165 PERMANOVA analysis. Our gastrectomy participants have undergone different types of

5

166   surgery (total and subtotal gastrectomy) and different gastrectomy reconstructions.

167   Thus, we also assess how those factors could explain the variance in the microbiome and

168   metabolome compositions in the gastrectomy subjects. The results were reported in

169   **Supplementary Table S4.**

170   Species richness and the Shannon diversity index were evaluated to estimate the

171   microbial diversity between the gastrectomy patients and controls participants. While

172   richness indicates the number of different species in a community, diversity takes into

173   account both the richness and evenness (relative abundance of species). The analysis was

174   carried out using the skbio.diversity.alpha module from the Python package of scikit-bio

175   (version 0.4.2) using the richness calculation metric of 'chao1' to estimate the Chao1

176   richness. The metric 'shannon' was used for defining the microbial Shannon-Wiener

177   alpha-diversity index. The two-sided Mann–Whitney U (MWU) test and significant

178   differences were stated if the P value was below 0.05.

179

180   **Different microbial features between the control and gastrectomy group**

181   The differences in relative abundance of the features (taxonomy, KO module, and

182   metabolites) were determined by linear discriminant analysis (LDA) effect size (LEfSe)

183   analysis, which emphasizes statistical significance, biological consistency, and effect

184   relevance[16]. LEfSe first identifies features that are statistically different between

185   control and gastrectomy groups using the non-parametric Kruskal–Wallis sum-rank test

186   (P<0.05). We modified the default calculation by controlling the multiple testing using

187   Benjamini–Hochberg (BH) false discovery rate (FDR) correction procedure. The features

188   that pass the threshold (q<0.1) were subsequently tested by a set of pairwise tests among

189   the sub-group using the Wilcoxon rank-sum test (P<0.05) to investigate its biological

190   consistency. LDA coupled with effect size measurement was finally performed to identify

191   bacterial taxa, annotated functions, and metabolome whose sequences were

192   differentially abundant between the controls and gastrectomy participants. In addition

193   to detect significant features, LEfSe also ranks features by effect size, which places

194   features that can explain most of the biological difference at the top[16]. A log-

195   transformed LDA score of 2.0 (default parameter) was used as the threshold for

196   significance in the KEGG module, taxon and metabolites to narrow down the focus. For

197   taxonomic analysis, two types of tables were used for LEfSe analysis. The first table

198   contains hierarchical taxonomic data in which abundances are computed at different

6

199 taxonomic levels, such as kingdom as the highest level and species as the lowest level,

200 and the results are visualized as a cladogram. The second table only contains taxa at the

201 species level and the results are visualized as a histogram.

202 The P-value and FDR-value for each test are shown in **Supplementary Tables**

203 **S10, S11, S12 and S14**. A log2-fold change in a feature was also calculated by dividing

204 the log of the relative abundance in gastrectomy patients to controls participants. A

205 pseudo-count was added to the relative abundances, which was the lowest relative

206 abundance observed for the entire data. This calculation was carried out to indicate in

207 which group the feature was over-represented to confirm the LEfSe calculation.

208

209 **Associations of the detected features to the demographic data**

210 Patient demographic data was acquired from the questionnaire and from demographic

211 measurements (**Supplementary Table S2**) and dietary information (**Supplementary**

212 **Table S16**). First, statistical differences were tested for the demographic data between

213 control and gastrectomy group to test the nature of the possible confounding factors. The

214 two-sided MWU test (scipy.stats.mannwhitneyu version 0.18.1) was performed on

215 numerical data (such as BMI, age, and dietary component information) and Fisher's exact

216 test (FisherExact 1.4.2) was performed on categorical data (such as medical history,

217 gender, smoking status, and alcohol consumption status).

218 In addition, the associations coefficient (variance explained) between the

219 demographic data (such as age, gender), medical history (such as history of diseases and

220 drug usage) and clinical parameters (BMI, serum glucose and total cholesterol) as

221 explanatory variables and the detected microbial features as response were tested by the

222 multivariate associations with linear models (MaAsLin) R package (default parameters

223 except for the fAllvAll=TRUE, dMinAbd=0.0, dMinSamp=0.05 and dSignificanceLevel =

224 0.1 were applied). The relative abundance of species and KEGG modules and metabolite

225 concentrations were transformed by arcsin-square root transformation. The MaAsLin

226 package was assessing the association using boosted additive generalized linear models

227 with the calculation of significance level of multiple testing correction[17]. The significant

228 associations were reported, and the associations were stated as significant if the P value

229 was below 0.05 and FDR (q value) was below 0.1. Additionally, the variance explained

230 calculated using the participant's group as explanatory variables (defined as crude

231 coefficient) were also compared to those adjusted by potential confounder which are

7

232   (BMI, total cholesterol, diabetes medications status , and gastric acids medications, age,

233   and gender) (strForcedPredictors="BMI, total cholesterol, diabetes medications, and

234   gastric acids medications, age,gender" was applied).

235

Oral microbes categorizations and microbial phenotype predictions

237   To examine the effect of gastrointestinal reconstruction, we further categorized the

238   annotated microbiome by performing oral microbes categorizations and phenotype

239   predictions. The categorizations of oral microbes were performed based on the expanded

240   Human Oral Microbiome Database (eHOMD)[18]. The reference is in a tabularized view

241   of all human oral microbial taxa that are defined and curated by eHOMD. The list of

242   microbes retrieved from both mOTU and MetaPhlAn2 pipeline was categorized as oral

243   microbes if it is listed on eHOMD and the non-listed species were categorized as others.

244   The total relative abundance of oral microbes in each participant was calculated and

245   compared between the gastrectomy (n=50) and control (n=56) groups using two-sided

246   MWU test.

247   To reveal the phenotypic properties regarding the oxygen requirement of the

248   species, we used the BugBase tool[19] on the metagenomics data. BugBase determines

249   the proportions for each microbiome sample of Gram positive, Gram negative, biofilm-

250   forming, pathogenic, potential mobile element-containing, oxygen utilizing, and oxidative

251   stress-tolerant microorganisms. For the metagenomics shotgun sequencing data, we

252   selected the OTUs using the IMG database before processing in BugBase. The detected

253   phenotypes were plotted and the statistical P-values from two-sided MWU test between

254   gastrectomy and control groups were also reported.

255

Correlation between genus and species in each group by SparCC

257   We estimated the microbial association (genus and species level) in each group using

258   SparCC (bootstrap n=5000)[20]. First, we selected the significantly different genus and

259   species between the gastrectomy and control groups that were overlapped in the mOTU

260   and MetaPhlAn2 (LEfSe: $P<0.05$, $q<0.1$, LDA>2.0) pipeline annotation. SparCC correlation

261   value between the genus or species counts (mOTU annotated) were then calculated in

262   each group (control and gastrectomy) independently. SparCC has been widely used to

263   estimate the correlation values from compositional data. Significant co-occurrence and

264   co-excluding interaction (SparCC correlation scores $\rho<-0.2$ or $\rho>0.2$ for genus and

8

265   ρ<−0.4 or ρ>0.4 for species, with a P<0.05) were visualized and analyzed using *igraph*.

266   We calculated the degree, betweenness and strength of each node to estimate its

267   importance to the network.

268

269   **Correlation between microbiome and metabolite in each group**

270   We performed Procrustes analysis to determine the congruence of two-dimensional

271   shapes produced by superimpositions of principal component analysis from microbiome

272   and metabolome datasets based on the Euclidean distances of eigenvalues from both

273   matrix (protest functions in vegan R package, n of permutations=1000)[21]. We selected

274   significantly   different   (LEfSe: P<0.05;   q<0.1;   LDA>2   for   comparison   of

275   gastrectomy(n=50) *vs* control (n=56)) metabolites and species that were overlap

276   between mOTU and MetaPhlAn2 annotations for the analysis. First, we generated

277   principal component from each measurement and then measured their inter-omics

278   associations. Procrustes analysis superimpose and scales principal component plots and

279   allows   for   quantification   of   non-random   conformance   between   two   different

280   measurements from similar participants (**Supplementary Figure S9**).

281   Additionally,   to   evaluate   the   possible   microbe–metabolite   associations   a

282   Spearman   rank-based   correlation   matrix   was   employed.   Correlation   analysis   was

283   performed on the relative abundance of 16 statistically-significant genus (LEfSe: P<0.05;

284   q<0.1; LDA score ≥2.0 both confirmed by mOTU and MetaPhlAn2 pipelines, genus relative

285   abundances   annotated   by   mOTU   pipeline   were   used   for   calculation)   and   the

286   concentrations of 34 statistically-significant metabolites (LEfSe: P<0.05; q<0.1; LDA

287   score≥3.0) among 44 gastrectomy patients and 54 controls participants. The correlation

288   analysis was performed in Python 2.7 using spearmanr modules from the scipy.stats

289   (version 0.18.0) package[22]. The P value of the correlation analysis was controlled by

290   the BH-FDR multiple correction test and was stated as the q-value. The correlation was

291   defined as significant if the P-value was below 0.05 and the FDR value was below 0.10.

292   The correlation coefficients were between -1 and 1, for which 0 implied no correlation,

293   while -1 or 1 implied an exact correlation. The cluster heatmap was created from the

294   spearman   correlation   matrices   (without   q   value   thresholding)   using   the

295   seaborn.clustermap based on the Euclidean distance metric. The results were visualized

296   as a heatmap in which the red color implied a positive correlation and the blue color

297   implied   a   negative   correlation[23].   Significant   correlations   (P<0.05;   q<0.10)   were

9

298   indicated by an asterisk (*).

299

300   **Predicting the species contribution on metabolite by Model-based Integration of**

301   **Metabolite Observations and Species Abundances (MIMOSA)**

302   The functional contribution of each species was retrieved from the HUMAnN2

303   annotation. The table of KO relative abundance stratified by species information was

304   utilized to calculate the community metabolic potential (CMP) of each species based on

305   the KEGG reaction information. This method integrates information about gene

306   abundances in terms of KOs and KEGG reaction definition describing the quantitative

307   relationship between genes and metabolites to provide an estimate of the way the species

308   composition may impact each metabolite's abundance. The CMP score was then

309   compared to the measured metabolomic data using Mantel test for each metabolite. The

310   test assesses Spearman correlation between pairwise differences in CMP scores (across

311   all pairs of samples) and the corresponding pairwise differences in measured metabolite

312   concentration[24]. A significant positive pairwise correlation (Mantel test: P≤0.01;

313   q≤0.01) was determined as "consistent" while a negative correlation was determined as

314   "contrasting" trend. We also retrieved the information of key contributor species to the

315   synthesis or degradation process for each metabolite (**Supplementary Figure 10,**

316   **Supplementary Table 10**).

317
318

319 **REFERENCES**

320 1 Ishii C, Nakanishi Y, Murakami S, *et al.* A Metabologenomic Approach Reveals Changes in the
321 Intestinal Environment of Mice Fed on American Diet. *Int J Mol Sci* 2018;**19**.
322 doi:10.3390/ijms19124079

323 2 Hirayama A, Kami K, Sugimoto M, *et al.* Quantitative metabolome profiling of colon and
324 stomach cancer microenvironment by capillary electrophoresis time-of-flight mass
325 spectrometry. *Cancer Res* 2009;**69**:4918–25.

326 3 Sugimoto M, Wong DT, Hirayama A, *et al.* Capillary electrophoresis mass spectrometry-based
327 saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles.
328 *Metabolomics* 2010;**6**:78–95.

329 4 Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
330 *EMBnet.journal* 2011;**17**:10.

331 5 Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*
332 2012;**9**:357–9.

333 6 Sunagawa S, Mende DR, Zeller G, *et al.* Metagenomic species profiling using universal
334 phylogenetic marker genes. *Nat Methods* 2013;**10**:1196–9.

335 7 Truong DT, Franzosa EA, Tickle TL, *et al.* MetaPhlAn2 for enhanced metagenomic taxonomic
336 profiling. *Nat Methods* 2015;**12**:902–3.

337 8 Franzosa EA, McIver LJ, Rahnavard G, *et al.* Species-level functional profiling of metagenomes
338 and metatranscriptomes. *Nat Methods* 2018;**15**:962–8.

339 9 Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*
340 2000;**28**:27–30.

341 10 Kanehisa M, Goto S, Sato Y, *et al.* Data, information, knowledge and principle: back to
342 metabolism in KEGG. *Nucleic Acids Res* 2014;**42**:D199–205.

343 11 Peng Y, Leung HCM, Yiu SM, *et al.* IDBA-UD: a de novo assembler for single-cell and
344 metagenomic sequencing data with highly uneven depth. *Bioinformatics* 2012;**28**:1420–8.

345 12 Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences.
346 *Nucleic Acids Res* 2010;**38**:e132–e132.

347 13 Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat*
348 *Methods* 2015;**12**:59–60.

349 14 Kanehisa M, Goto S, Sato Y, *et al.* Data, information, knowledge and principle: back to
350 metabolism in KEGG. *Nucleic Acids Res* 2014;**42**:D199–205.

351 15 McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and
352 graphics of microbiome census data. *PLoS One* 2013;**8**:e61217.

353 16 Segata N, Izard J, Waldron L, *et al.* Metagenomic biomarker discovery and explanation.
354 *Genome Biol* 2011;**12**:R60.

355 17 Morgan XC, Tickle TL, Sokol H, *et al.* Dysfunction of the intestinal microbiome in
356 inflammatory bowel disease and treatment. *Genome Biol* 2012;**13**:R79.

357 18 Chen T, Yu W-H, Izard J, *et al.* The Human Oral Microbiome Database: a web accessible
358   resource for investigating oral microbe taxonomic and genomic information. *Database*
359   2010;**2010**:baq013.

360 19 Langille MGI, Zaneveld J, Caporaso JG, *et al.* Predictive functional profiling of microbial
361   communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 2013;**31**:814–21.

362 20 Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput
363   Biol* 2012;**8**:e1002687.

364 21 Dixon P. VEGAN, a package of R functions for community ecology. Journal of Vegetation
365   Science. 2003;**14**:927. doi:10.1658/1100-9233(2003)014[0927:vaporf]2.0.co;2

366 22 Zwillinger D, Kokoska S. *CRC Standard Probability and Statistics Tables and Formulae*. 1999.

367 23 Shankar V, Homer D, Rigsbee L, *et al.* The networks of human gut microbe–metabolite
368   associations are different between health and irritable bowel syndrome. *ISME J*
369   2015;**9**:1899–903.

370 24 Noecker C, Eng A, Srinivasan S, *et al.* Metabolic Model-Based Integration of Microbiome
371   Taxonomic and Metabolomic Profiles Elucidates Mechanistic Links between Ecological and
372   Metabolic Variation. *mSystems* 2016;**1**. doi:10.1128/mSystems.00013-15

373

1

1   **SUPPLEMENTARY RESULTS**

2

3   **Associations of clinical parameter with microbiome and metabolome**

4   We have collected available serum glucose and total cholesterol levels from patients'

5   medical records. On average, our participants were under the category of normal BMI

6   (18.5<BMI<25.0). However, 16 participants (control, n=12; gastrectomy, n=4) fall to the

7   overweight category (BMI>25.0)[1]. The control group also tends to have higher BMI

8   (two-sided Mann-Whitney *U* (MWU) test: P=1.224×10^{-5}). Serum glucose level did not

9   differ significantly (two-sided Mann-Whitney *U* (MWU) test: P=0.147) between

10   gastrectomy (n=50) and control (n=42) patients. In 35 subjects (gastrectomy, n=14;

11   control, n=21), serum glucose levels were higher than normal (69–104 mg/dL). Total

12   cholesterol level was significantly higher (two-sided MWU test: P=0.0145) in control

13   (n=40) compared to gastrectomy (n=50) patients. The average amount of cholesterol in

14   each group, however, remained within the normal range (128–219 mg/dL). Twenty-two

15   participants (control, n=15; gastrectomy, n=7) have high cholesterol level (over 219

16   mg/dL).

17   To account for possible confounding effects of those clinical factors, we performed

18   Permutational Multivariate Analysis of Variance (PERMANOVA) on the microbiome and

19   metabolome between samples distance (Bray-curtis) using the clinical parameters (BMI,

20   serum glucose, and total cholesterol level) as the predictor. The microbiome

21   compositions were significantly varied along with the BMI (adonis: R^2=0.0171, P=0.0289;

22   R^2=0.0240, P=9.99×10^{-4} for mOTU and MetaPhlAn2, respectively) but not in metabolome

23   (adonis: R^2=0.0120, P=0.0699). However, the participants grouping explained the

24   variance of microbiome better compare to the BMI (**Supplementary Table S4**).

25   Therefore, the significantly different microbiome compositions along with the BMI, might

26   be potentially explained by the participants' grouping (control and gastrectomy). Overall

27   microbiome and metabolome composition did not vary significantly (adonis: P>0.05) in

28   relation to the serum glucose and total cholesterol level (**Supplementary Table S4**).

29   Additionally, we tested associations of each microbial features (species, functional

30   modules, and metabolite) with the clinical parameters (BMI, serum glucose, and total

31   cholesterol) using the Multivariate associations with Linear Models (MaAsLin) R package.

32   Among the species that were reported to be differentially enriched between control

33   (n=56) and gastrectomy (n=50), we found that *Roseburia hominis*, *Eubacterium eligens*

2

34   that were enriched in gastrectomy (both by mOTU and MetaPhlan2 annotations) were

35   found to be negatively associated with the BMI and *Ruminococcus gnavus* that were

36   enriched in control were positively correlated to BMI (**Supplementary Tables S10 and**

37   **S11**). Similarly, to the PERMANOVA results, the associations values (explained variance)

38   were higher when we used the participants' grouping (control and gastrectomy) as the

39   predictor rather than the BMI (**Supplementary Tables S10 and S11**). We also calculated

40   the associations values between participants status and each species with and without

41   adjustment of BMI (**Supplementary Materials**). We found that the explained variances

42   were increased by 10% or more after adjustment for *Roseburia hominis* and

43   *Ruminococcus gnavus*. Therefore, the BMI might also affect the differential abundance of

44   these species in control and gastrectomy groups (**Supplementary Table S10 and S11**).

45   We did not observe differentially enriched species (**Supplementary Table S10 and S11**),

46   KEGG modules (**Supplementary Table 12**), and metabolites (**Supplementary Table**

47   **S14**) between gastrectomy and control group had significant associations with the serum

48   glucose and total cholesterol (MaAsLin: P>0.05, q>0.1).

49   In the function modules, we did not find overlap between KEGG modules annotated

50   by *in house* and HUMAnN2 pipeline that were in significant association with the clinical

51   parameters. Metabolites' association with the participants' demographic parameters by

52   MaAsLin showed that cholate enrichment in the control group might be affected by BMI,

53   whereas phenyl-lactate and arginine enrichment might be affected by total cholesterol

54   (**Supplementary Results** and **Supplementary Table S14**).

55

56   **Associations of medical history with the microbiome and metabolome**

57   Underlying comorbidities and concurrent medication of the participants may influence

58   gut microbiota. Therefore, we extracted information about the medical history of 23

59   diseases (e.g., hypertension, diabetes, dyslipidemia) and the history of usage of eight

60   drugs (e.g., diabetes medication, gastric acid-suppression medication, cholesterol

61   medication), which were obtained from questionnaires (**Supplementary Tables S2**). No

62   significant (Fisher's exact test: P>0.05) difference was observed in the distribution of

63   individuals with any disease history between gastrectomy and control groups (**Table 1**).

64   However, we found a significantly higher (Fisher's exact test: P<0.05) number of subjects

65   with a history of diabetes medication or gastric acid-suppression medication in the

66   control group. Thus, we performed PERMANOVA to assess how gastric acid-suppression

67 medication or diabetes medication contributed to variations in microbial community

68 data. The microbiome and metabolome compositions were not significantly different

69 between the users of these medications (adonis: P>0.05) (**Supplementary Table S4**). To

70 eliminate the possible confounding effect of gastric acid-suppression medication or

71 diabetes medication, we performed two-step analysis in addition to the associations

72 between microbial profiles (species, functions, metabolite) and drug usage by MaAsLin

73 (**Supplementary Methods**). First, we performed a comparison between participants

74 who took and did not take diabetes medication or gastric acid-suppression medication

75 within the control group to specifically assess the effect of drug usage. Second, we re-

76 evaluated the significantly different microbiome and metabolome between the two

77 groups after excluding users of these drugs and compared them with those before

78 exclusions, whom we referred to as "originally reported".

79

80 ***Gastric acid-suppression medication***

81 There were significant different distributions (Fisher's exact test: P=0.0375) of the user

82 of gastric acid-suppression medication in the gastrectomy (n=4) and control (n=13)

83 groups. PERMANOVA analysis showed that the microbiome and metabolome

84 composition did not vary significantly (P>0.05) between users (n=17) and non-users

85 (n=89) of gastric acid-suppression medication (**Supplementary Table S4**). Association

86 analysis by MaAsLin found that differentially enriched fecal microbial features (species,

87 KEGG modules, metabolites) between control (n=56) and gastrectomy (n=50) patients

88 were not associated with gastric acid-suppression medication (P>0.05, q>0.1)

89 (**Supplementary Tables S10, S11, S12, and S14**).

90 Furthermore, we compared the microbiome and metabolome compositions between

91 control participants who took (n=13) and did not take (n=43) gastric acid-suppression

92 medication usage by PERMANOVA and LEfSe. PERMANOVA analysis showed that the

93 composition of the mOTU-annotated species, MetaPhlan2-annotated species, and

94 metabolome (R^2=0.0195, P=0.337; R^2=0.0255, P=0.0829; R^2=0.0100, P=0.829,

95 respectively) (**Supplementary Table S4f**) were not significantly varied along with the

96 gastric acid-suppression medication in our control participants. LEfSe results showed

97 eight mOTU-annotated species and five MetaPhlAn2-annotated species that were

98 differentially enriched (LEfSe: P<0.05, q<0.1, LDA>2.0) between control participants who

99 took (n=13) and did not take (n=43) gastric acid-suppression medication

100   (**Supplementary Table S8**). Among them, *Streptococcus mutans* and *Haemophilus*

101   *parainfluenzae* were significantly enriched in the participants with gastric acid-

102   suppression medication both in mOTU and MetaPhlan2 species annotations. Twenty-two

103   and twenty-three KEGG modules annotated by *in-house* pipeline and HUMAnN2 pipeline,

104   respectively, were also significantly enriched (LEfSe: P<0.05, q<0.1, LDA>2.0) in the

105   participants with gastric acid-suppression medication. Among them, fourteen were

106   overlapped between two functional annotations. Interestingly, some of the microbiome

107   features that were enriched in gastric acid-suppression users overlapped with those

108   enriched in gastrectomy individuals (**Supplementary Table S8**). Those included two

109   nutrient transporter (M00229, "arginine transport system"; M00317, "manganese/iron

110   transport system"), two two-component regulatory system (M00447, "CpxA-CpxR

111   (envelope stress response) two-component regulatory system"; M00456, "ArcB-ArcA

112   (anoxic redox control) two-component regulatory system"), Cationic antimicrobial

113   peptide (CAMP) resistance (M00728), and "Pyruvate oxidation" (M00307) that were

114   significantly enriched in the gastrectomy. These findings might indicate some of the

115   microbiome features may be influenced by reduced gastric acid. We did not observe any

116   metabolites that were differentially enriched between the participants with or without

117   history of gastric acid-suppression medication.

118   Additionally, we examined the possible confounding effect of the gastric acid-

119   suppression medication by re-performing LEfSe on the subset of the two groups

120   excluding those with gastric acid-suppression medication (control, n=43; gastrectomy,

121   n=46) to confirm that the originally reported gastrectomy enriched microbiome and

122   metabolome signatures were not affected by the exclusion. In the species level, 31 out of

123   38 differentially enriched species (LEfSe: P<0.05, q<0.1, LDA>2.0) between control

124   (n=56) and gastrectomy (n= 50) groups before exclusion that overlap in annotation

125   based on MetaPhlAn2 and mOTU were at similar enrichment after exclusion

126   (**Supplementary Table S6**). From this analysis, we also observed the different

127   enrichment pattern of *Streptococcus mutans* before and after exclusion. The

128   *Streptococcus mutans* enrichment in the control group before exclusion might be

129   contributed by the participants with gastric acid-suppression medication. In fact, we

130   observed that *Streptococcus mutans* was enriched in the gastric acid-suppression

131   medication in the control group (**Supplementary Table S8**). In the functional modules

132   level, we detected 32 KEGG modules that were differentially enriched (LEfSe: P<0.05,

133   q<0.1, LDA>2.0) between control (n=56) and gastrectomy (n= 50) before exclusion and

134   were overlapped in the uniref90 and KEGG gene-based annotation. Among them, 26

135   KEGG modules were conserved before and after exclusion (**Supplementary Table S6**).

136   Similarly, large portions of significantly different metabolites (86 out of 104) between

137   control and gastrectomy groups were conserved before and after exclusion

138   (**Supplementary Table S6**). Thus, it may reflect that the differences of microbial features

139   were mostly driven by gastrectomy rather than the gastric acid-suppression medication.

140

141   ***Diabetes medication***

142   We compared the microbiome and metabolome compositions between control

143   participants who took (n=12) and did not take (n=43) diabetes medication by

144   PERMANOVA and LEfSe. The compositions of mOTU-annotated species by PERMANOVA

145   were significantly varied along with diabetes medication (adonis: R^2=0.0342, P=0.0220)

146   but it was not significantly different in the MetaPhlan2-annotated species and

147   metabolome (R^2=0.0251, P=0.114; R^2=0.0178, P=0.477, respectively) (**Supplementary**

148   **Table S4f**). LEfSe results showed eight mOTU-annotated species and one MetaPhlAn2-

149   annotated species that were differentially enriched (LEfSe: P<0.05, q<0.1, LDA>2.0)

150   between control participants who took (n=12) and did not take (n=43) diabetes

151   medication (**Supplementary Table S9**). Among them, *Mitsuokella multacida* was

152   significantly enriched (LEfSe: P=0.00159, q=0.0687, LDA=3.23; P=9.93×10^{-5}, q=0.0236,

153   LDA=3.357, in mOTU and MetaPhlAn2 annotation, respectively) in the participants who

154   took diabetes medication. We did not detect significantly different (LEfSe: P<0.05, q<0.1,

155   LDA>2.0) functional modules that were overlapped based on the annotation by our *in-*

156   *house* pipeline and HUMAnN2 and metabolites between these two groups

157   (**Supplementary Table S9**).

158   Additionally, we re-performed LEfSe on the subset of the two groups excluding

159   participants who took diabetes medication (control, n=43; gastrectomy, n=48) to confirm

160   that the originally reported gastrectomy enriched microbiome and metabolome

161   signatures were not affected by the exclusion. In the species level, 35 out of 38

162   differentially enriched species (LEfSe: P<0.05, q<0.1, LDA>2.0) between control (n=56)

163   and gastrectomy (n= 50) group before exclusion that were annotated based on

164   MetaPhlAn2 and mOTU were at similar enrichment after exclusion (**Supplementary**

165   **Table S7**). In the functional modules level, we detected 32 KEGG modules that were

6

166   differentially enriched (LEfSe: P<0.05, q<0.1, LDA>2.0) between control (n=56) and

167   gastrectomy (n= 50) before exclusion and were overlapped in the uniref90 and KEGG

168   gene based annotation. Among them, 21 KEGG modules were conserved before and after

169   exclusion (**Supplementary Table S7**). Similarly, large portions of significantly different

170   metabolites (85 out of 94) between control and gastrectomy groups were conserved

171   before and after exclusion (**Supplementary Table S7**). We also did not find the

172   associations between each of microbial features (species, functional modules, and

173   metabolites) with the diabetes medications (MaAsLin: P>0.05, q>0.1). Thus, it may reflect

174   that the differences of microbial features were driven by gastrectomy rather than

175   diabetes medication.

176

177   **Microbiome, functional potential and metabolome differences in different types of**

178   **gastrectomy**

179   The post-gastrectomy patients in the present study were underwent different types of

180   gastrectomy (total gastrectomy, n=12 and subtotal gastrectomy, n=38) and followed by

181   different types of reconstructions (*Stomach-stomach anastomosis*, n=1; *Billroth I*, n=2;

182   *Jejunal interpositions*, n=6; *Pylorus-preserving gastrectomy,* n=8; *Roux-en-Y*, n=29). The

183   overall profiles analysis by PERMANOVA revealed a tendency towards different species

184   and metabolite composition between different types of and reconstructions (adonis:

185   P<0.05) (**Supplementary Table S4**). We additionally performed the LEfSe analysis to

186   compare the species, functional modules, and metabolites compositions in the different

187   types of gastrectomy. Regarding different types of surgical reconstructions, we were

188   limited by a small number of patients for each reconstruction. Therefore, any statistical

189   analysis may not be powerful enough to detect microbiome and metabolome differences

190   across reconstructions. To partially address this issue, we provided analysis on the

191   microbial features comparison between control (n=56) and patients with Roux-en Y

192   reconstruction (n=29). We also discuss the microbial features of interest and their

193   distribution in different types of surgery and reconstruction

194

195   ***Surgery types***

196   The overall profiles revealed a tendency towards different species and metabolite

197   composition between different types of surgery (total gastrectomy, n=12; subtotal

198   gastrectomy=38) (adonis: R^2=0.0318, P=0.0709; R^2=0.0370, P=0.175 for mOTU-

199   annotated species and metabolome, respectively); however, only MetaPhlAn2-annotated

200   species reached a statistically significant level (adonis: R^2=0.0337, P=0.0360)

201   (**Supplementary Table S4e**). Based on species categorizations, we did not observe

202   significantly different (P>0.05) compositions of oral microbes, aerobes, and facultative

203   anaerobes between total gastrectomy (n=12) and subtotal gastrectomy (n=38) groups

204   (**Supplementary Figure S2**). Species richness (Chao1 index) and diversity (Shannon-

205   Wiener alpha-diversity index) also not significantly varied (P>0.05) between those two

206   groups. However, the total relative abundance of oral species and aerobes tended to differ

207   between both types of gastrectomy (subtotal and total gastrectomy) and control groups

208   (**Supplementary Figure S2**).

209   Most of post-gastrectomy patients (38 of 50) underwent subtotal gastrectomy. Thus,

210   we performed LEfSe pairwise comparison between control (n=56) and each type of

211   surgery to analyze whether the detected microbiome features (species, functional

212   modules, and metabolites) were mainly represented in subtotal gastrectomy. From this

213   analysis, we recovered several microbiome features that mutually enriched in subtotal

214   and total gastrectomies in comparison to the control group (**Supplementary Table S5**).

215   Among 27 species that were significantly enriched (LEfSe: P<0.05, q<0.1, LDA>2.0) in the

216   gastrectomy group (n=50) compared to the control group (n=56) and overlapped in

217   between mOTU and MetaPhlAn2 annotations, seven species were mutually enriched in

218   the total and subtotal gastrectomy compared to the control group (**Supplementary

219   Table S5B and S5C**). In addition, ten species were enriched in the subtotal gastrectomy

220   (**Supplementary Table S5B**) and three species were enriched in the total gastrectomy

221   (**Supplementary Table S5C**) compared to the control group. Similar pattern was

222   observed in the metabolites enrichment in the gastrectomy. The majority of gastrectomy-

223   enriched metabolites were found to be enriched in the subtotal gastrectomy (43 out of

224   46). However, the different patterns were observed in the functional modules. There

225   were 8 out of 26 of the gastrectomy-enriched features were mutually enriched in both

226   types of gastrectomy (total and subtotal gastrectomy) compared to control. The majority

227   of modules which were thirteen functional modules were enriched in the total

228   gastrectomy, while seven functional modules were enriched in the subtotal gastrectomy

229   compared to the control group (**Supplementary Table S5**).

230   In addition, LEfSe analysis of total gastrectomy *versus* subtotal gastrectomy showed

231   that *Fusobacterium nucleatum* was enriched in the former (LEfSe: P=5.58×10^{-5}, q=0.0150,

232   LDA=2.34; P=1.45×10^{-5}, q=0.00366, LDA=2.14, for mOTU and MetaPhlAn2 annotations,

233   respectively) (**Supplementary Table S5**). A comparison between total gastrectomy

234   (n=12) and control (n=56) groups further confirmed enrichment of *F. nucleatum* in total

235   gastrectomy (LEfSe: P=1.53×10^{-5}, q=0.00205, LDA=2.87; P=4.12×10^{-6}, q=3.34×10^{-4},

236   LDA=1.43, in mOTU and MetaPhlAn2 annotations, respectively). Additionally, MaAsLin

237   confirmed positive associations between total gastrectomy and *F. nucleatum* (MaAsLin:

238   P=3.17×10^{-12}, q=4.58×10^{-8}, r=0.0122; P=3.78×10^{-5}, q=0.0143, r=0.00776, in mOTU and

239   MetaPhlAn2 annotations, respectively). Its enrichment might be reflecting its survival in

240   the higher pH environment following total gastrectomy. This is worth noting because *F.*

241   *nucleatum* has long been considered as an opportunistic pathogen which is important

242   during the development of the plaque biofilm and recently associated with

243   gastrointestinal related diseases such as colorectal cancer (CRC)[3]. None of the KEGG

244   modules and metabolites were significantly (LEfSe: P<0.05; q<0.1, LDA>2.0) enriched

245   either in total or subtotal gastrectomy (**Supplementary Table S5**). Therefore, the

246   majority type of gastrectomy might not highly affect the observed gastrectomy-enriched

247   signatures.

248

249   ***Reconstructions Type***

250   The post-gastrectomy patients underwent different types of reconstructions (*Stomach-*

251   *stomach anastomosis*, n=1; *Billroth I*, n=2; *Jejunal interpositions*, n=6, *Pylorus-preserving*

252   *gastrectomy*, n=8; *Roux-en-Y*, n=29). The overall species compositions (mOTU and

253   MetaPhlAn2-annotated species) were significantly different in different types of

254   reconstruction but not in metabolite profiles (adonis: R^2=0.123, P=0.0150; R^2=0.121,

255   P=0.0119; R^2=0.0986, P=0.483, for mOTU-annotated species, MetaPhlAn2-annotated

256   species, and metabolome, respectively, **Supplementary Table S4**). However, we were

257   limited by a small number of patients for each reconstruction types, thus, any statistical

258   analysis may not be powerful enough to detect microbiome and metabolome.

259   To partially address this issue, we show distribution patterns of microbial features

260   of interest (species, functional modules, and metabolites). First, in terms of predominant

261   species in post gastrectomy patients across different reconstructions, we observed two

262   patterns. The species enrichment that might reflect the Roux-en Y reconstructions

263   (**Pattern I**) and those that might be driven by other reconstructions (**Pattern II**). The

264   majority of predominantly enriched species in gastrectomy came into **Pattern I**. For

265 instances, in **Pattern I,** we observed that *Streptococcus anginosus, Streptococcus*
266 *parasanguinis, Streptococcus vestibularis* and *Streptococcus salivarius* were enriched in
267 patients undergoing Roux-en Y reconstruction, when we compared the control (n=56)
268 and Roux-en Y groups (n=29) (**Supplementary Table S18**). The distribution pattern also
269 showed that those species were more abundant among patients undergoing Roux-en Y
270 reconstruction (**Supplementary Figure S3**). Similar patterns were observed in three
271 species of *Veillonella* (**Supplementary Figure S3**). In contrast, three species of
272 *Lactobacillus* (*Lactobacillus gasseri, Lactobacillus oris* and *Lactobacillus salivarius*) were
273 more abundant in Billroth I reconstruction (**Supplementary Figure S3**). When we
274 compared control (n=50) and post-gastrectomy patients undergone Roux-en Y
275 reconstruction (n=29) we did not find that these three species were differentially
276 enriched (**Supplementary Table S18**). Thus, enrichment of those three species might be
277 driven by Billroth I reconstructions.
278   Different distribution pattern of CRC-related species such as *Atopobium parvulum*
279 and *F. nucleatum* were also observed in different reconstructions. *A. parvulum* were
280 observed to be more abundant in the Roux-en Y reconstructions both in total and subtotal
281 gastrectomies, while *F. nucleatum* were enriched in Roux-en Y reconstructions in patients
282 undergoing total gastrectomy (**Supplementary Figure S3**). These results were in
283 accordant with our analyses between total and subtotal gastrectomies (**Supplementary**
284 **Table S5**). We also observed different distribution pattern of metabolites that were
285 associated with CRC (**Supplementary Figure S4**).
286
287 ***Control versus Roux-en Y reconstructions***
288 Majority of post-gastrectomy patients (29 of 50) underwent Roux-en Y reconstruction
289 following the surgery. Thus, we performed LEfSe analysis to analyze whether the
290 detected microbiome features (species, functional modules, and metabolites) were
291 mainly represented in patients undergoing Roux-en Y reconstruction. We did recover
292 high number of microbial features that were overlapped when we compared the control
293 group (n=56) with the gastrectomy group (n=50) in a subset of post-gastrectomy patients
294 who underwent R-Y reconstruction (n=29). For instances, in the species level, 26 out of
295 38 of differentially enriched species (LEfSe: $P<0.05$, $q<0.1$, LDA>2.0) between control
296 (n=56) and gastrectomy (n= 50) groups that were overlapped in annotation based on
297 MetaPhlAn2 and mOTU were at similar enrichment in the comparison between control

10

298   (n=56) and Roux-en Y (n=29) (**Supplementary Table S18**). In the functional modules

299   level, the 32 KEGG modules that were differentially enriched (LEfSe: P<0.05, q<0.1,

300   LDA>2.0) between control (n=56) and gastrectomy (n= 50) and overlap in two

301   annotations pipelines (uniref90 and KEGG gene based-annotation) were also retained in

302   the comparison between control (n=56) and Roux-en Y (n=29) (**Supplementary Table**

303   **S18**). Similarly, large portions of significantly different metabolites (86 out of 104)

304   between control and gastrectomy groups were conserved before and after exclusion

305   (**Supplementary Table S18**). The non-overlap features might possibly come from other

306   reconstructions. The number of subjects in the other surgery reconstructions were

307   relatively low compared to the control to give a statistical power to account for the effects

308   of reconstructions to the microbiome and metabolome.

309

310   **Observed gastrointestinal complications following gastrectomy**

311   Information about any gastrointestinal complications (e.g., diarrhea, dumping syndrome,

312   anemia) was available for 47 out of 50 gastrectomy patients from their medical records

313   (**Supplementary Table S2**). Among 29 patients who had gastrointestinal complications

314   after gastrectomy, 15 subjects experienced dumping syndrome. After gastrectomy

315   patients were divided into those with (n=15) and without (n=32) dumping syndrome,

316   overall microbiome and metabolome profiles revealed a generally different composition,

317   but this was not statistically significant (adonis: $R2=0.0557$, $P=0.0629$; $R2=0.0518$,

318    $P=0.0959$; $R2=0.0476$, $P=0.442$, for mOTU-annotated species, MetaPhlAn2-annotated

319   species, and metabolome, respectively, **Supplementary Table S4e**). In addition, we did

320   not observe any significant difference (LEfSe: P>0.05, q>0.1) regarding the abundance of

321   species, functional modules, and metabolites between patients with and without

322   dumping syndrome (**Supplementary Table S17**). Notably, the diagnosis of dumping

323   syndrome depends mainly on the individual clinician's perspective.

324   Furthermore, we observed a high rate of dumping syndrome in patients with total

325   gastrectomy (8 out of 12 patients, 66.7%) compared to subtotal gastrectomy (7 out of 38,

326   18.4%), which does not deviate from previous studies such as a Japanese large-scale

327   investigation (n=1,153) reported by Mine S *et al.* (79.6% early and 48.7% late dumping

328   syndrome for total gastrectomy)[4]. In general, the frequency of postsurgical dumping

329   syndrome is estimated at 25-50%[5]. Such a wide range might be explained by diagnosis

11

330 for dumping syndrome being highly dependent on the individual clinician's
331 perspective[6].
332
333 **Differences of species-species correlations between control and gastrectomy**
334 **groups**
335   We performed the microbes correlations in the species level, owing to the different
336 characteristics of different species in the same genus. We observed that several co-
337 occurrence and co-excluding patterns that appeared at genus level were present also at
338 species level. The number of edges was higher (ρ>0.4; ρ<-0.4) in the control group (co-
339 occurrence, 31; co-excluding, 8) compared to the gastrectomy group (co-occurrence, 25;
340 co-excluding, 1). *Veillonella*, which formed the hub of the network in the gastrectomy
341 group, was confirmed at species-level network (**Supplementary Figure S8**). Loss of
342 edges between *Veillonella* and *Lactobacillus* as well as *Anaerotruncus* and *Alistipes* in the
343 microbes' network in the gastrectomy group was observed also at species level. In the
344 species-level network, the species in the same genus tended to form a common cluster,
345 such as the cluster of species from *Streptococcus* and *Veillonella* genera. It should be
346 noted, however, that some patterns have been lost in the species-level network during
347 network construction process as only differentially abundant species between the
348 gastrectomy and control groups, which overlapped in mOTU and MetaPhlAn2 pipelines,
349 were used. This was the case, for example, of the genus *Coprobacillus*, which disappeared
350 in the species network.

12

351 **SUPPLEMENTARY TABLES**

352 **Supplementary Table S1.**

353 Clinical characteristics of participants in post-gastrectomy and control groups

354 **Supplementary Table S2.**

355 Clinical characteristics of each participant based on medical records

356 **Supplementary Table S3.**

357 Quality control and annotation profile of each sample

358 **Supplementary Table S4.**

359 PERMANOVA analysis between microbiome and metabolome based on clinical

360 parameters, demographic data, and medical history

361 **Supplementary Table S5.**

362 Microbiome and metabolome enrichment in pairwise comparisons between different

363 types of gastrectomy and the control group

364 **Supplementary Table S6.**

365 Microbiome and metabolome profiles after exclusion of gastric acid-suppression

366 medication users (control, n=43; gastrectomy, n=46)

367 **Supplementary Table S7.**

368 Microbiome and metabolome profiles after exclusion of diabetes therapeutic medication

369 users (control, n=43; gastrectomy, n=48)

370 **Supplementary Table S8.**

371 Effect of gastric acid-suppression medication on microbiome and metabolome profiles in

372 the control group (user, n=13; non-user, n=43)

373 **Supplementary Table S9.**

374 Effect of diabetes therapeutic drugs on microbiome and metabolome profiles in the

375 control group (user, n=12; non-user, n=43)

376 **Supplementary Table S10.**

377 Significantly different taxa between gastrectomy and control groups and their associated

378 clinical information (annotation with mOTU)

379 **Supplementary Table S11.**

380 Significantly different taxa between gastrectomy and control groups and their associated

381 clinical information (annotation with MetaPhlAn2)

382 **Supplementary Table S12.**

13

383   Significantly different KEGG modules between gastrectomy and control groups and their

384   associated demographic information (annotation by *in house* pipeline and HUMAnN2)

385   **Supplementary Table S13.**

386   Species alpha-diversity and richness of KEGG modules contributor

387   **Supplementary Table S14.**

388   Significantly different metabolites between gastrectomy and control groups and their

389   associated demographic information

390   **Supplementary Table S15.**

391   MIMOSA output showing the species predicted to contribute to each metabolite

392   **Supplementary Table S16.**

393   Consumption of each dietary component in post-gastrectomy and control groups

394   **Supplementary Table S17.**

395   Microbiome and metabolome profiles in participants with (n=15) and without (n=32)

396   dumping syndrome

397   **Supplementary Table S18.**

398   Microbiome and metabolome profiles between control (n=56) and post-gastrectomy

399   patients undergoing Roux-en-Y reconstruction (n=29)

400

401

402

403

404

405

406

407

408

409

410

411

412

14

413 **SUPPLEMENTARY FIGURES**

414 **Supplementary Figure S1. Study participants' overview and analysis workflow**

415 (**A**) Sample collection and general analysis workflow. (**B**) Detailed workflow for our

416 metagenome pipeline including quality filtering, functional annotation, and taxonomic

417 annotation.

418

419 **Supplementary Figure S2. Microbiome and metabolome composition in different**

420 **types of surgery (total gastrectomy *versus* subtotal gastrectomy)**

421 Principal coordinates analysis (PCoA) with Bray-Curtis distance (**A**) was performed to

422 assess the community structure of species' relative abundance obtained by mOTU and

423 MetaPhlAn2, and metabolites in the subtotal gastrectomy group (n=38) (**red**) and in the

424 total gastrectomy group (n=12) (**blue**). Species richness was measured using the Chao1

425 index (**B**) calculated from the species annotated by mOTU and MetaPhlAn2. Species

426 alpha-diversity was measured using the Shannon-Wiener index (**C**) based on mOTU and

427 MetaPhlAn2 annotation. The summed relative abundances of oral microbes (**D**) were

428 compared between the control (n=50), subtotal gastrectomy (n=38), and total

429 gastrectomy (n=12) groups based on species annotated by mOTU and MetaPhlAn2

430 annotation. The summed relative abundances of aerobes (**E**) and facultative anaerobes

431 (**F**) were also compared between the three groups.

432

433 **Supplementary Figure S3. Species distributions in different types of surgery and**

434 **reconstructions**

435 Relative abundances (log10) were plotted as boxplots to shows the distribution of each

436 species of interests. Two distributions patterns were observed. Most species followed

437 Pattern I which reflects the Roux-en-Y reconstructions (A). Several species followed

438 Pattern II and they might be driven by other reconstructions (B). CRC-enriched species

439 that were enriched in the gastrectomy group also showed different distribution patterns

440 across reconstructions (C). Different types of reconstructions are labeled with different

441 colors (see legends in the figures). The distributions were also divided into control

442 (n=50), subtotal gastrectomy (n=38), and total gastrectomy (n=12).

443

444 **Supplementary Figure S4. Metabolites distributions in different types of surgery**

445 **and reconstructions**

15

446   The concentrations of metabolites (nmol/g) were plotted as boxplots to shows the

447   distributions of each predominant metabolite that were associated to colorectal cancer

448   and were enriched in the gastrectomy group (**A**). Different types of reconstructions are

449   labeled with different colors (**see legends in the figures**). The distributions were also

450   divided into control (n=50), subtotal gastrectomy (n=38), and total gastrectomy (n=12).

451

452   **Supplementary Figure S5. Comparison between the variance explained calculated**

453   **in terms of the participants' control or gastrectomy groups (crude coefficient) and**

454   **those adjusted by possible confounder and demographic data (adjusted**

455   **coefficient) as explanatory variable**

456   Crude and adjusted coefficient (variance explained) are shown as scatter plots. The

457   adjusted coefficients were calculated by adjusting for possible confounders (BMI, total

458   cholesterol, status of diabetes medication, gastric acids-suppression medication) in

459   addition to the demographic variables (age and gender) as explanatory in (**A-E**). Each dot

460   represents the response variable (species, KEGG modules, metabolites). The **red** dot

461   represents the features which are significantly (LEfSe: P<0.05, q<0.1, LDA>2.0) different

462   between control (n=56) and gastrectomy (n=50) groups based on species annotated by

463   mOTU (**A**), species annotated by MetaPhlAn2 (**B**), functional modules annotated by the

464   *in-house* pipeline (**C**), functional modules annotated by the HUMAnN2 pipeline (**D**), and

465   metabolites (**E**)

466

467   **Supplementary Figure S6. Differences in KEGG modules between control (n=56) and**

468   **gastrectomy (n=50) groups**

469   (**A**) Relative abundance and LDA score (log10) of KEGG modules annotated by our *in-*

470   *house* pipeline (LEfSe: P<0.05, q<0.1, LDA>2.0). (**B**) Relative abundance and LDA score

471   (log10) of KEGG modules annotated by HUMAnN2 (LEfSe: P<0.05, q<0.1, LDA>2.0).

472

473   **Supplementary Figure S7. Species contribution to KEGG modules**

474   KEGG modules involved in phosphate transport (**A**) and manganese/zinc/iron/ transport

475   (**B**) that were differentially abundant between the gastrectomy (n=50) and control

476   (n=56) groups are annotated by their taxonomic contributor (**see legends in the figure**).

477   The KEGG modules' relative abundances are represented by the top value of each stack

478   of bars. Samples were subsequently sorted according to the dominant contributor to a

16

479   module and then grouped as either gastrectomy or control (sample in order differs

480   between panels).

481

482   **Supplementary Figure S8. Species-species correlations**

483   Co-occurrence (**red**) and co-excluding (**green**) relationships between species (SparCC: -

484   0.4< ρ<0.4, P<0.05) in gastrectomy (n=44) (**A**) and control (n=54) (**B**) groups. The edge

485   width corresponds to the SparCC correlation coefficients. The nodes' size is scaled based

486   on the genus relative abundance averaged over participants within each group. Nodes'

487   color represents enrichment of the genus in gastrectomy (**orange**) and control (**blue**)

488   participants.

489

490   **Supplementary Figure S9. Procrustes analysis between species and metabolites**

491   Principal component analysis (PCA) plots for mOTU-annotated species (**A**), MetaPhlAn2-

492   annotated species (**B**), and metabolites profile (**C**). Procrustes analysis was performed

493   between metabolite profiles and species annotated by mOTU (**D**) and MetaPhlAn2 (**E**).

494

495   **Supplementary Figure S10. Genus contribution to metabolites based on MIMOSA**

496   **analysis**

497   Each table cell in the matrix represents the contribution of a particular genus to a

498   metabolite (**see legends in the figure**). Table cells are colored based on the community

499   metabolic potential (CMP) score calculated from the KEGG reaction data and KO relative

500   abundances stratified by species information. The species were later summarized at the

501   genus level. Each metabolite is given a prediction level, which represents how the

502   observed metabolite values are consistent or contrasting with the predicted metabolite-

503   producing potential. High prediction scores indicate that a metabolite is enriched in the

504   gastrectomy group, and it is predicted to be produced in sufficient amounts by a certain

505   genus. Based on reaction information, MIMOSA predicted metabolite enrichment in one

506   of the groups and compared that enrichment trend to those observed in the actual

507   quantification. A positive value (**green**) shows a consistent trend and a negative value

508   (**orange**) shows a contrasting trend compared to the measured metabolite.

17

References

1   Yoshinaga M, Ichiki T, Ito Y. Prevalence of Overweight and Obesity in Japan. Epidemiology of Obesity in Children and Adolescents. 2011;:153–62. doi:10.1007/978-1-4419-6039-9_9

2   Huang L, Xu A-M, Li T-J, *et al.* Should peri-gastrectomy gastric acidity be our focus among gastric cancer patients? *World J Gastroenterol* 2014;**20**:6981–8.

3   Brennan CA, Garrett WS. Fusobacterium nucleatum — symbiont, opportunist and oncobacterium. Nature Reviews Microbiology. 2019;**17**:156–66. doi:10.1038/s41579-018-0129-6

4   Mine S, Sano T, Tsutsumi K, *et al.* Large-scale investigation into dumping syndrome after gastrectomy for gastric cancer. *J Am Coll Surg* 2010;**211**:628–36.

5   Vavricka SR, Greuter T. Gastroparesis and Dumping Syndrome: Current Concepts and Management. *J Clin Med Res* 2019;**8**. doi:10.3390/jcm8081127

6   Gys B, Plaeke P, Lamme B, *et al.* Heterogeneity in the Definition and Clinical Characteristics of Dumping Syndrome: a Review of the Literature. *Obes Surg* 2019;**29**:1984–9.
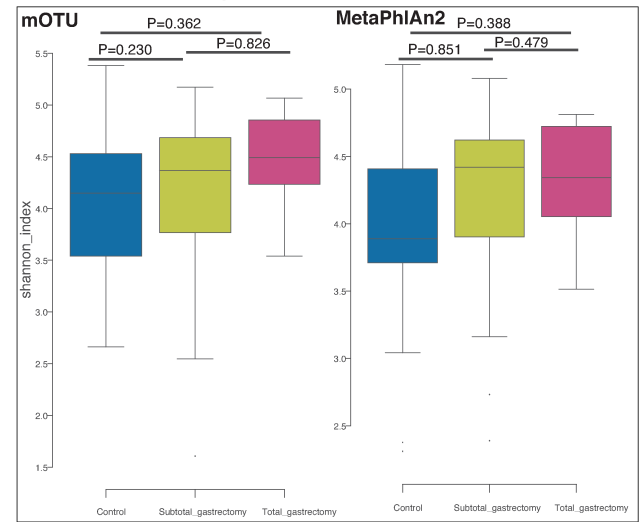
A Principal coordinates analysis (PCoA)
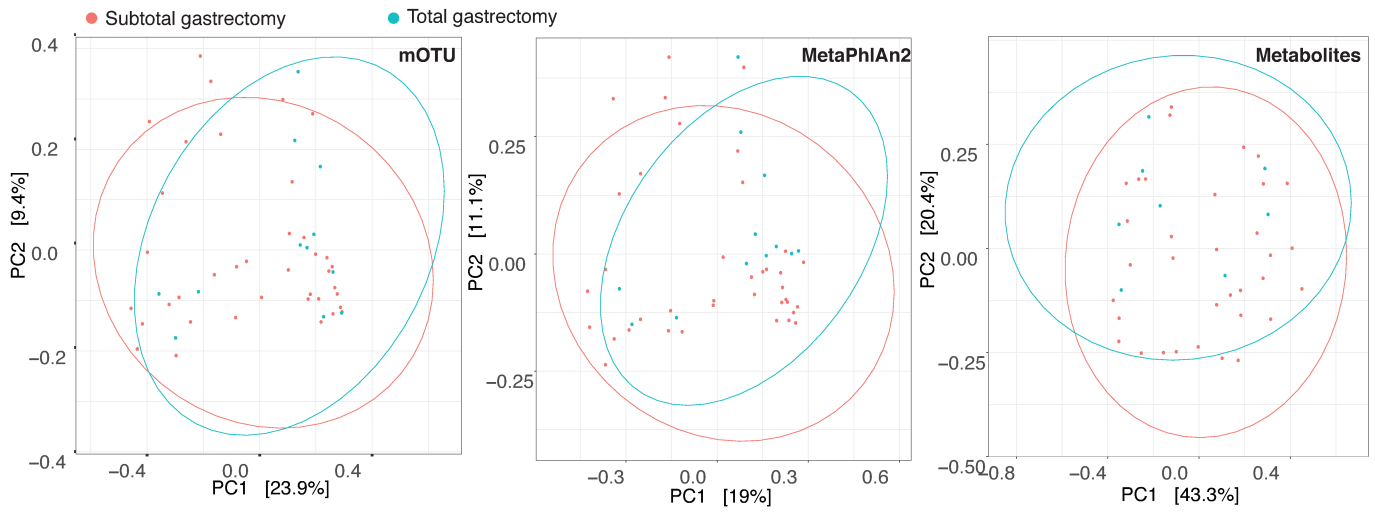
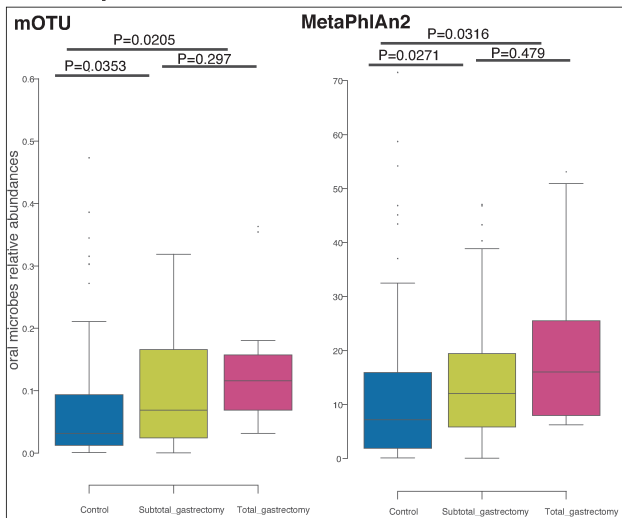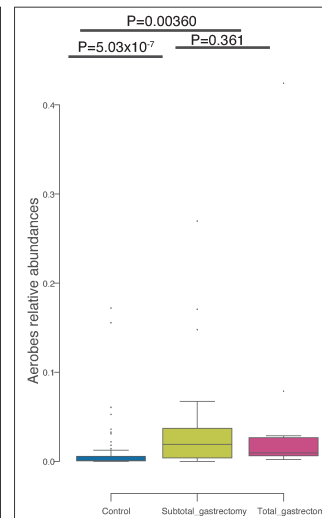## B Species richness (Chao1 index)
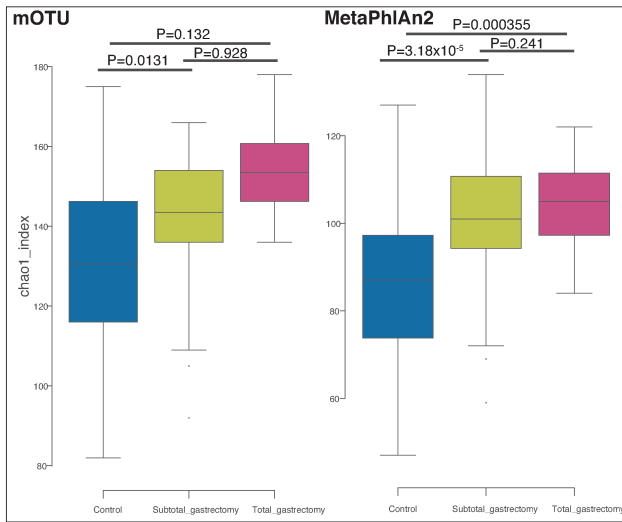
C Species diversity (Shannon-Wiener index)
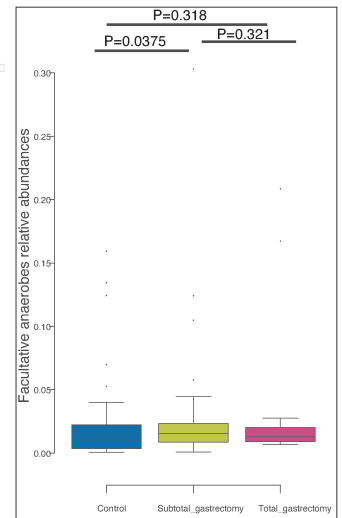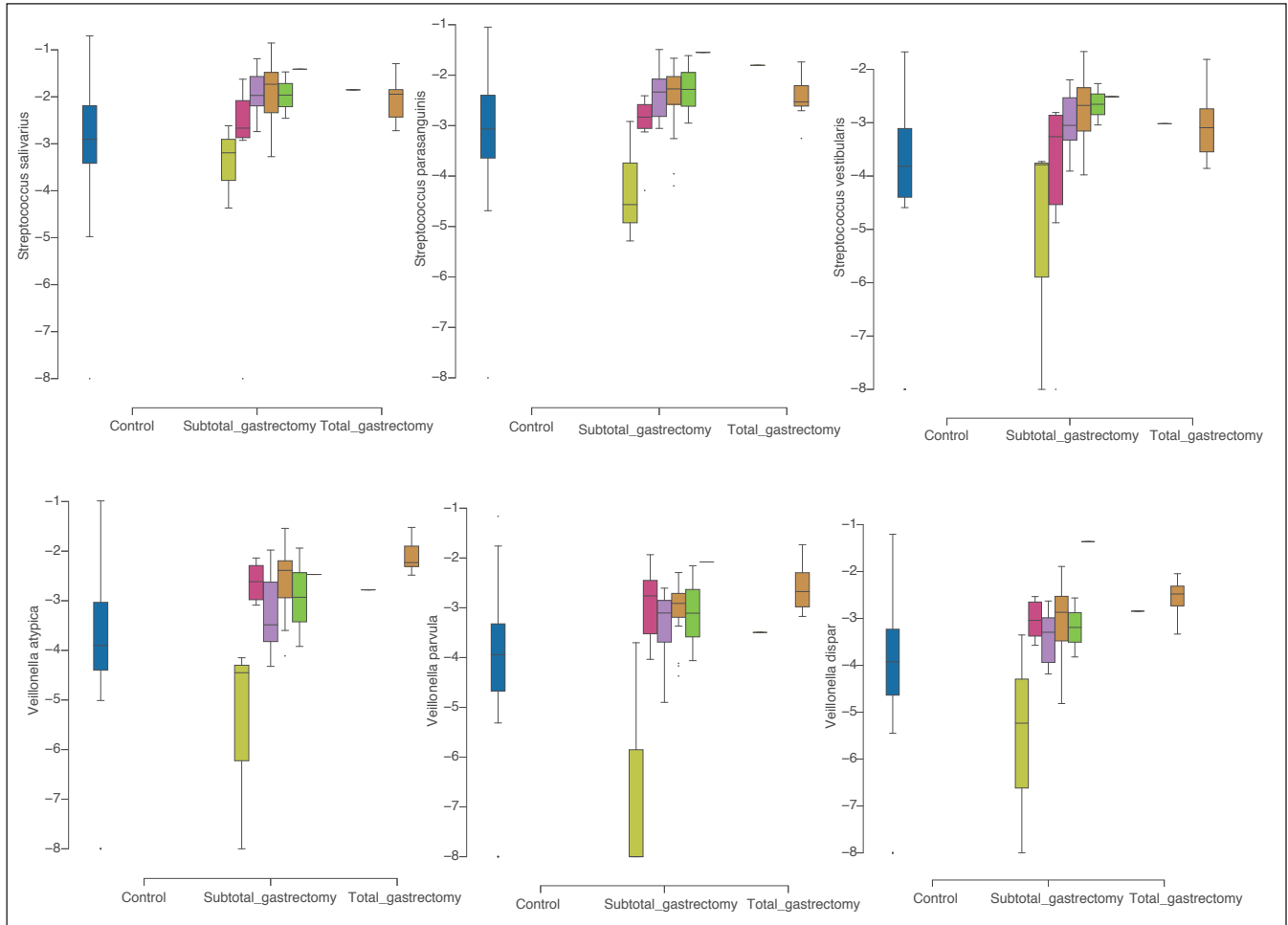

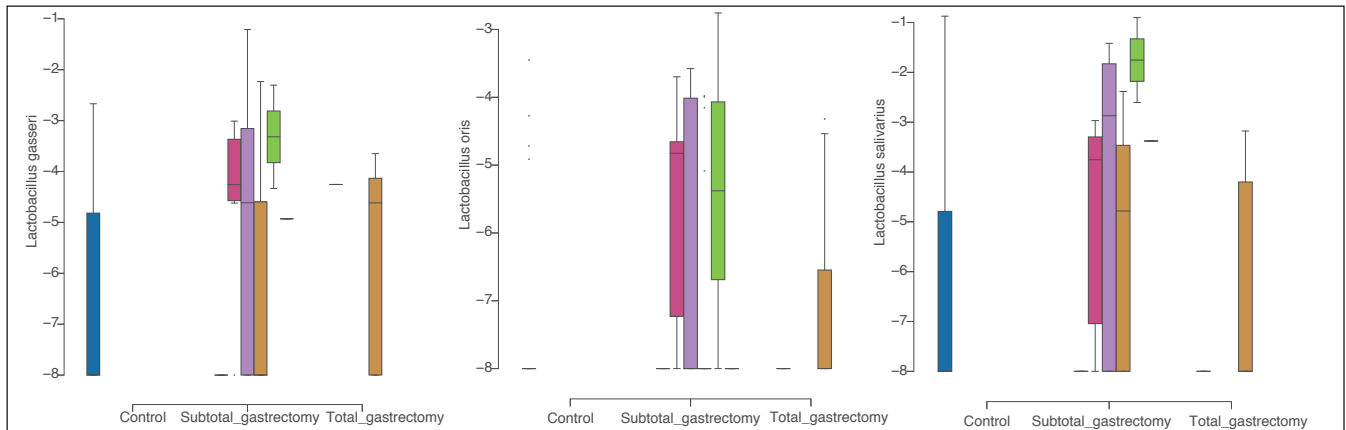## D Oral species

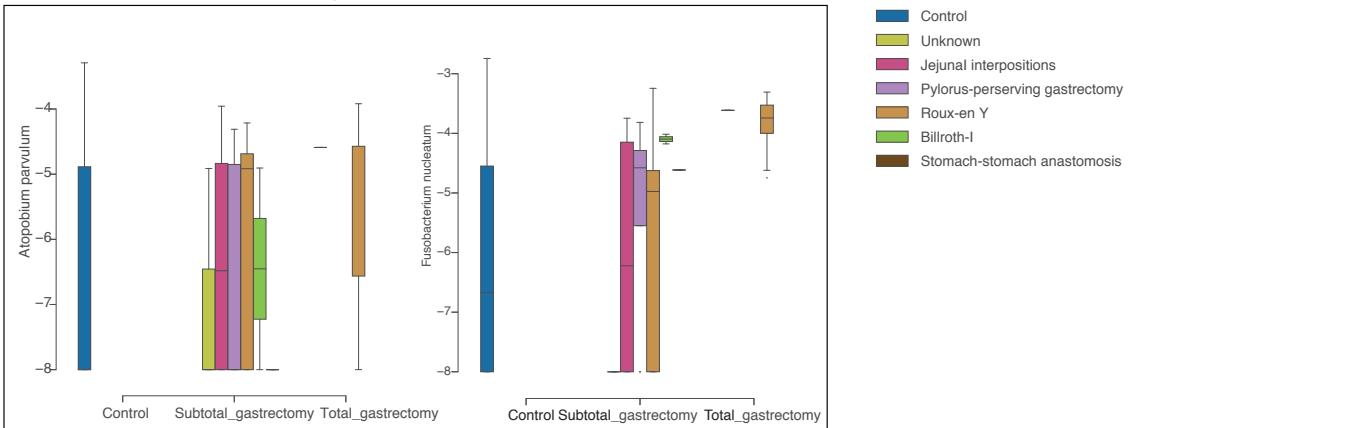## E Aerobes

F Facultative anaerobes

A Pattern I: Species enrichment that might reflect the Roux-en Y reconstructions
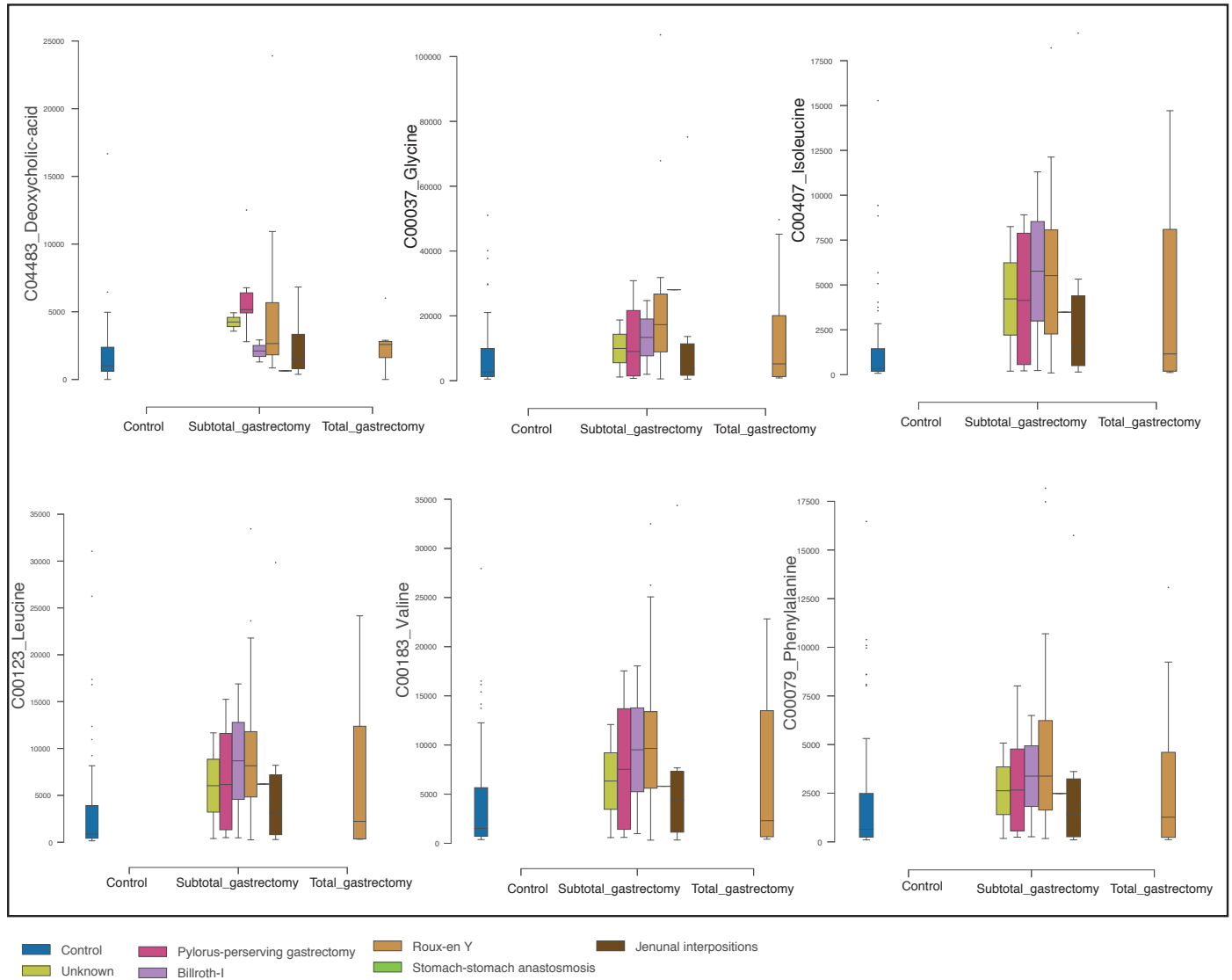

B Pattern II: Species enrichment that might driven by the other reconstructions

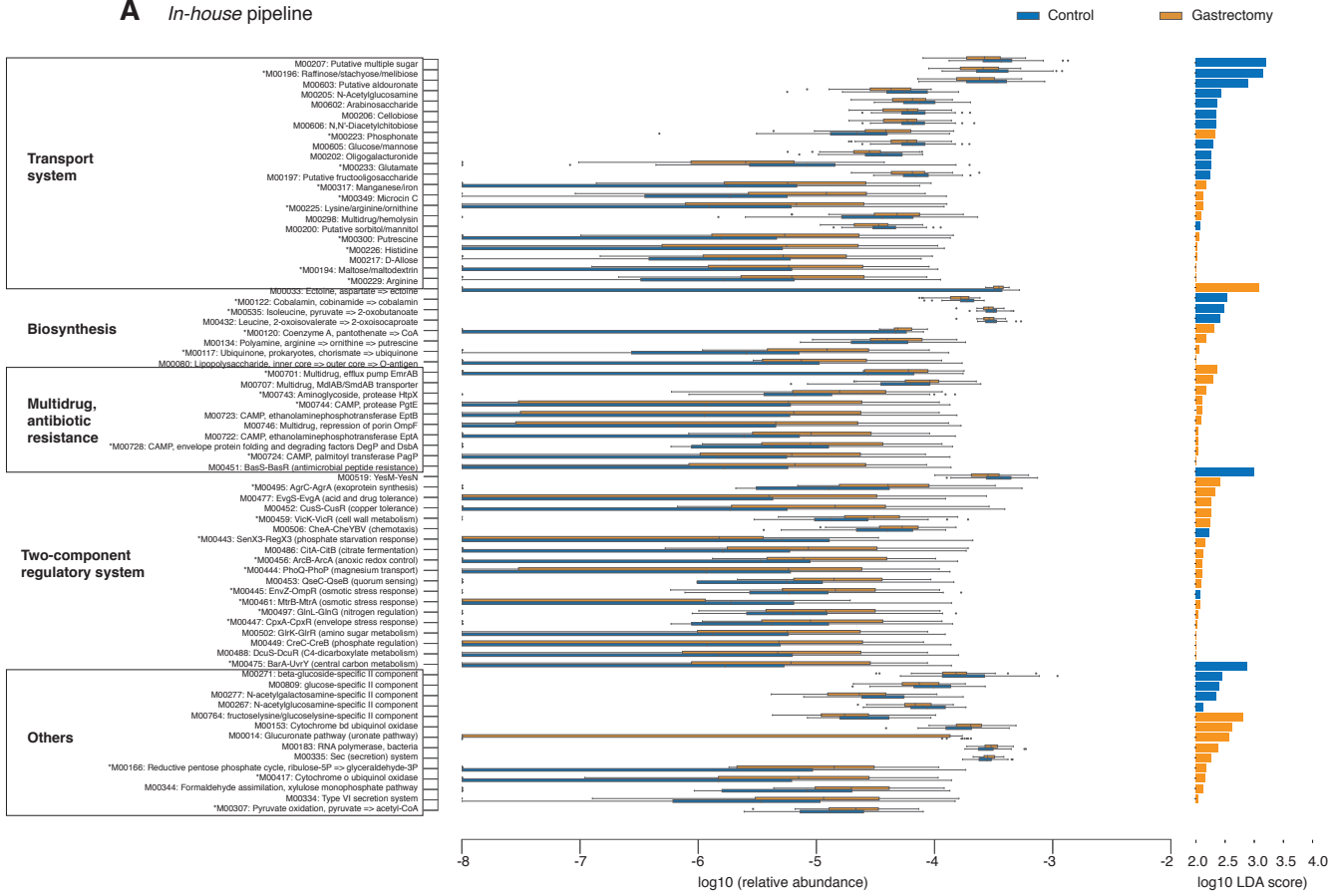C Pattern of colorectal cancer-enriched species in gastrectomy accros different types of reconstructions

Legend:
- Control
- Unknown
- Jejunal interpositions
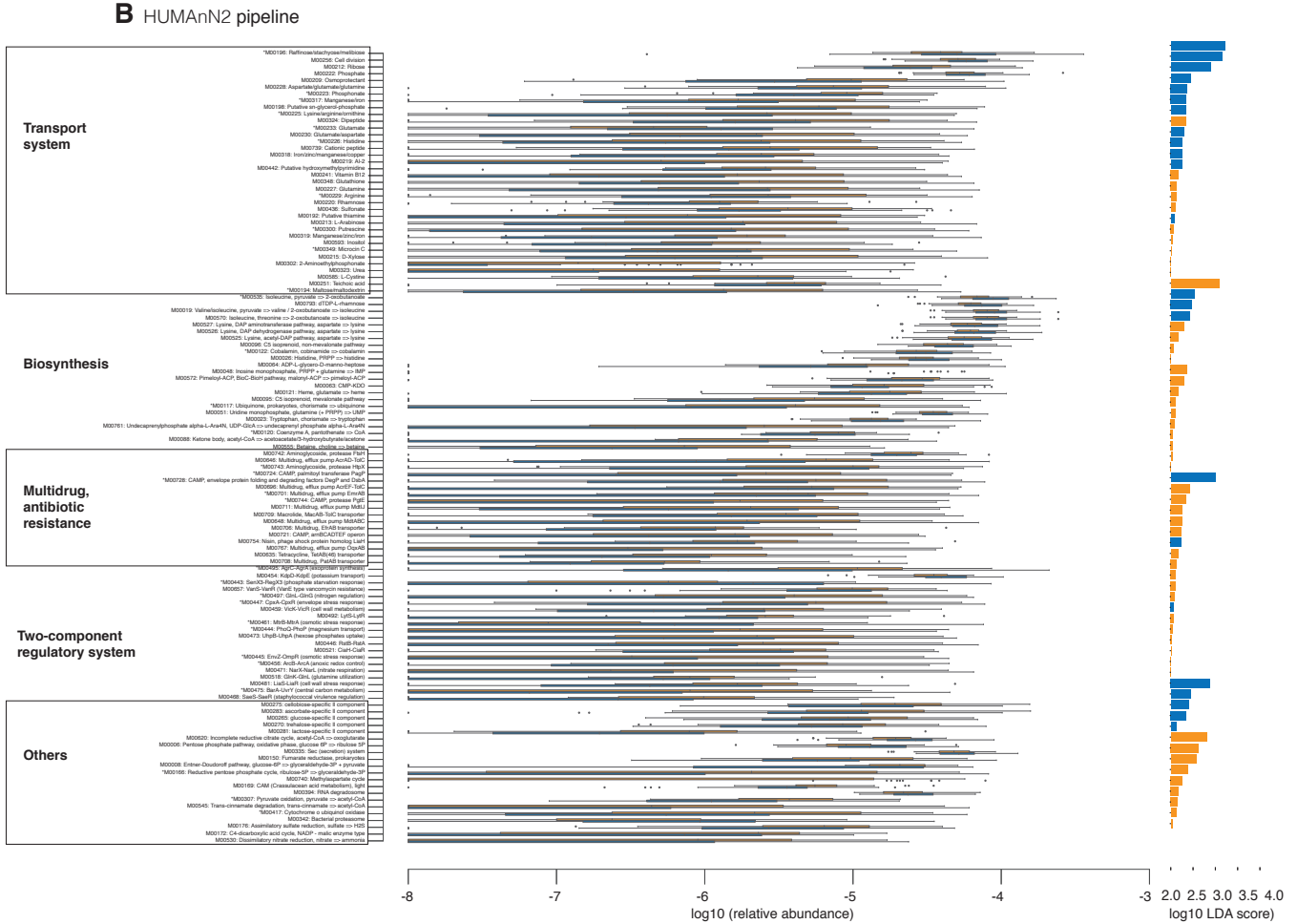- Pylorus-perserving gastrectomy
- Roux-en Y
- Billroth-I
- Stomach-stomach anastomosis

A Pattern of predominant metabolites associated to colorectal cancer across different types of reconstructions

Legend:
- Control
- Unknown
- Pylorus-perserving gastrectomy
- Billroth-I
- Roux-en Y
- Stomach-stomach anastosmosis
- Jenunal interpositions

A mOTU

B MetaPhlAn2

C *In-house* pipeline

D HUMAnN2

E Metabolite
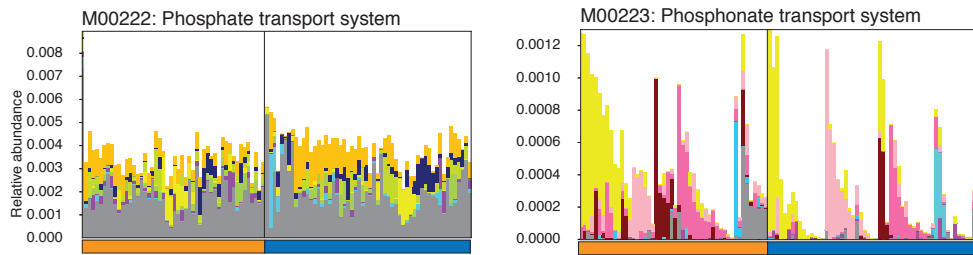
• P≤0.05, FDR≤0.1, LDA score (log10)≥2.0

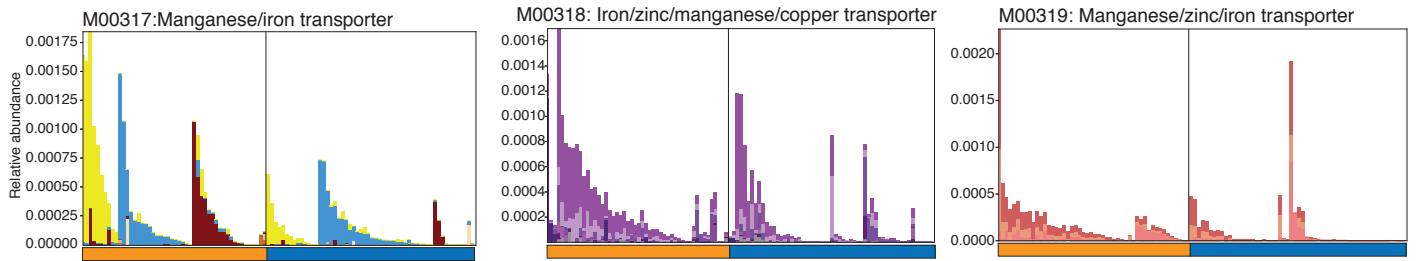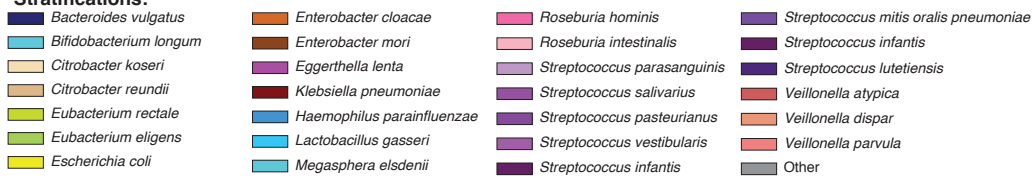**A** *In-house* pipeline

B HUMAnN2 pipeline
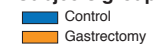
A Species contributor for phosphate transporter

B Species contributor for manganese/zinc/iron transporter

Stratifications:

- *Bacteroides vulgatus*
- *Bifidobacterium longum*
- *Citrobacter koseri*
- *Citrobacter reundii*
- *Eubacterium rectale*
- *Eubacterium eligens*
- *Escherichia coli*
- *Enterobacter cloacae*
- *Enterobacter mori*
- *Eggerthella lenta*
- *Klebsiella pneumoniae*
- *Haemophilus parainfluenzae*
- *Lactobacillus gasseri*
- *Megasphera elsdenii*
- *Roseburia hominis*
- *Roseburia intestinalis*
- *Streptococcus parasanguinis*
- *Streptococcus salivarius*
- *Streptococcus pasteurianus*
- *Streptococcus vestibularis*
- *Streptococcus infantis*
- *Streptococcus mitis oralis pneumoniae*
- *Streptococcus infantis*
- *Streptococcus lutetiensis*
- *Veillonella atypica*
- *Veillonella dispar*
- *Veillonella parvula*
- Other

Subjects group

- Control
- Gastrectomy

A. Gastrectomy

B. Control

ρ>0.8	ρ<-0.8	Gastrectomy-enriched
ρ>0.6	ρ<-0.6	Control-enriched
ρ>0.4	ρ<-0.4	

A. mOTU

B. MetaPhlAn2

C. Metabolites

D r=0.279, P=0.003

E r=0.292, P=0.002

Gastrectomy
Control
Species
Metabolites