SUPPLEMENTARY METHODS

DNA extraction, shotgun metagenomic sequencing, and quality control

We extracted genomic DNA from the frozen fecal samples by the bead-beating method using the GNOME® DNA Isolation Kit (MP Biomedicals, Santa Ana, CA, USA). The quality was assessed on a 4200 TapeStation (Aglient Technologies, Santa Clara, CA, USA). The extracted DNAs were then subjected to shotgun metagenomic sequencing on an Illumina HiSeq2500 platform with 150-bp paired-end read lengths and a targeted 5 Gbp sequencing depth. The sequencing libraries were generated with the Nextera XT DNA Library Prep Kit (Illumina, San Diego, CA, USA). The quality of the libraries was analyzed by the 4200 TapeStation.

Metabolomics quantification of fecal samples

Fecal metabolites were extracted from ten milligrams of fresh thawed freeze-dried fecal samples suspended in 400 μL of 50% methanol in Milli-Q containing internal standards (20 μM each of methionine sulfone and D-camphor-10-sulfonic acid (CSA)[1]. The 3-mm zirconia beads (BioSpec Products, Bartlesville, OK, USA) and 100 mg of 0.1-mm zirconia/silica beads (BioSpec Products, Bartlesville, OK, USA) were added into the mixture then subjected to vigorous shaking using Micro Smash (TOMY, Nerima, Tokyo, Japan). The suspensions were centrifuged for 10 minutes at 1500 rpm, 10 times. The supernatant was transferred to a 5-kDa-cutoff filter column (Ultrafree MC-PLHCC 250/pk) for Metabolome Analysis (Human Metabolome Technologies, Tsuruoka, Yamagata, Japan). The flow-through was dried under vacuum and the residue then was dissolved in 40 μL of Milli-Q water containing reference compounds (200 μM each of 3-aminopyrrolidine and trimesate). The levels of extracted metabolites were measured in both positive and negative modes by Capillary Electrophoresis Time-of-Flight Mass Spectrometry (CE-TOFMS) as previously described[2]. The CE-TOFMS experiments were carried out using an Aglient CE Capillary Electrophoresis System (Agilent Technologies, Santa Clara, CA, USA).

The raw data were processed for metabolite quantification using automatic integration software MasterHands (ver. 2.16.0.15)[3]. We generated annotations tables based on the measurement of standard compounds and aligned with the datasets according to the similar *m/z* values and normalized migration time. The peak areas were

Erawijantari PP, *et al. Gut* 2020;0:1–12. doi: 10.1136/gutjnl-2019-319188

then normalized against those of the internal standards methionine sulfone and CSA for cationic and anionic metabolites, respectively. Concentrations of each metabolite were calculated based on their relative peak areas and the concentrations of the standard compounds.

For the analysis, concentrations below the detection limit were substituted with zero, and metabolites for which levels were below the detection limit in all of the samples were excluded. The metabolite concentration (nmole) was normalized using the fecal weight to obtain the amount of metabolite in each gram of sample (nmole/g)[1,2].

Reads quality filtering

Shotgun metagenomics sequencing was performed on an Illumina HiSeq2500 platform with 150-bp read length to a targeted 5 Gbp sequencing depth. The sequence quality filtering process involved several steps and started by removing reads that contained ambiguous bases (**Supplementary Figure S1B**). The elimination of reads with *PhiX* DNA contamination was then performed, followed by trimming the sequencing adapter and 3'-end low quality reads using cutadapt (version 1.9.1)[4]. The adapter sequences were removed by cutadapt using the option of "-a CTGTCTCTTATACACATCTCCGAGCCCACGAGAC -o 33" for the forward primer sequence and "-a CTGTCTCTTATACACATCTGACGCTGCCGACGA -o 32" for the reverse primer sequence.

We continued the filtering process by applying a read length and average quality cut-off. The reads of less than 50 bp in length and an average quality score of less than 25 were removed. Furthermore, contaminating human DNA sequences were removed by screening against the human genome (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/001/405/GCA_000001405.15_GRCh38/GCA000001405.15_GRCh38_assembly_structure/Primary_Assembly/assembled_chromosomes/FASTA/). We used Bowtie2 (version 2.2.9)[5] to eliminate reads with *PhiX* and human genome contamination. Only paired-end reads were selected as high-quality reads for further analysis.

Taxonomic profiling by the mOTU and MetaPhlAn2 pipeline

Taxonomic profiling was performed using the metagenomics operational taxonomic units (mOTUs)[6] and MetaPhlAn2 pipeline[7] that will be explained below.

Erawijantari PP, *et al. Gut* 2020;0:1–12. doi: 10.1136/gutjnl-2019-319188

mOTU pipeline

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

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

MetaPhlAn2 pipeline

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

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

Erawijantari PP, *et al. Gut* 2020;0:1–12. doi: 10.1136/gutjnl-2019-319188

bases was then divided by the length of the predicted genes to determine the gene abundances. In other words, the predicted gene abundance was defined by the number of reads that mapped to the predicted gene. After the abundance of the annotated gene was calculated, total sum scaling normalization (TSS) was performed to account for uneven sequencing depth across samples. The abundance value was divided by the total abundance per sample to generate its relative abundance. The KEGG were then annotated into the KEGG orthologous (KO) group profile[14].

HUMAnN2 pipeline

Functional profiling was performed using HUMAnN2 v0.11.1 in UniRef90 mode[8]. High-quality reads were initially mapped to the pangenomes species identified during the taxonomic profiling by MetaPhlAn2 using Bowtie2. The pangenomes have been pre-annotated to its respective UniRef90 families. Subsequently the unmapped reads were mapped to UniRef90 by translated search using DIAMOND. The gene-level outputs were produced in reads per kilobases units and stratified according to known/unclassified community contribution to relative abundance unit. The gene relative abundance was regrouped into its respective KO using the UniRef90-KEGG linked file.

Microbial community structure analysis

To visualize the microbial community structure between the gastrectomy and the control group, we calculated between-sample diversity score (Bray-Curtis distance) using the relative abundances of taxa and the concentration of metabolites. We visualized the separation using the Principal coordinates analysis (PCoA). The distance calculation and PCoA analysis were performed by R package phyloseq[15]. Subsequently, we performed permutational multivariate analysis of variance (PERMANOVA) ('adonis' function, vegan package, R) on the matrices between-sample diversity score to test the differences of microbial community between group. We also performed PERMANOVA using the other predictor which are the medical history that were extracted from the questionnaire and the clinical parameters (BMI, serum glucose and total cholesterol) from participants records. For the medical history, we marked the participants who had a history of other diseases or history of medications before the sample collections as "Yes". We excluded the samples who did not have the data for particular parameter (label as NA) in the PERMANOVA analysis. Our gastrectomy participants have undergone different types of

surgery (total and subtotal gastrectomy) and different gastrectomy reconstructions. Thus, we also assess how those factors could explain the variance in the microbiome and metabolome compositions in the gastrectomy subjects. The results were reported in **Supplementary Table S4.**

Species richness and the Shannon diversity index were evaluated to estimate the microbial diversity between the gastrectomy patients and controls participants. While richness indicates the number of different species in a community, diversity takes into account both the richness and evenness (relative abundance of species). The analysis was carried out using the skbio.diversity.alpha module from the Python package of scikit-bio (version 0.4.2) using the richness calculation metric of 'chao1' to estimate the Chao1 richness. The metric 'shannon' was used for defining the microbial Shannon-Wiener alpha-diversity index. The two-sided Mann–Whitney *U* (MWU) test and significant differences were stated if the P value was below 0.05.

Different microbial features between the control and gastrectomy group

The differences in relative abundance of the features (taxonomy, KO module, and metabolites) were determined by linear discriminant analysis (LDA) effect size (LEfSe) analysis, which emphasizes statistical significance, biological consistency, and effect relevance[16]. LEfSe first identifies features that are statistically different between control and gastrectomy groups using the non-parametric Kruskal–Wallis sum-rank test ($P<0.05$). We modified the default calculation by controlling the multiple testing using Benjamini–Hochberg (BH) false discovery rate (FDR) correction procedure. The features that pass the threshold ($q<0.1$) were subsequently tested by a set of pairwise tests among the sub-group using the Wilcoxon rank-sum test ($P<0.05$) to investigate its biological consistency. LDA coupled with effect size measurement was finally performed to identify bacterial taxa, annotated functions, and metabolome whose sequences were differentially abundant between the controls and gastrectomy participants. In addition to detect significant features, LEfSe also ranks features by effect size, which places features that can explain most of the biological difference at the top[16]. A log-transformed LDA score of 2.0 (default parameter) was used as the threshold for significance in the KEGG module, taxon and metabolites to narrow down the focus. For taxonomic analysis, two types of tables were used for LEfSe analysis. The first table contains hierarchical taxonomic data in which abundances are computed at different

taxonomic levels, such as kingdom as the highest level and species as the lowest level, and the results are visualized as a cladogram. The second table only contains taxa at the species level and the results are visualized as a histogram.

The P-value and FDR-value for each test are shown in **Supplementary Tables S10, S11, S12 and S14**. A log2-fold change in a feature was also calculated by dividing the log of the relative abundance in gastrectomy patients to controls participants. A pseudo-count was added to the relative abundances, which was the lowest relative abundance observed for the entire data. This calculation was carried out to indicate in which group the feature was over-represented to confirm the LEfSe calculation.

Associations of the detected features to the demographic data

Patient demographic data was acquired from the questionnaire and from demographic measurements (**Supplementary Table S2**) and dietary information (**Supplementary Table S16**). First, statistical differences were tested for the demographic data between control and gastrectomy group to test the nature of the possible confounding factors. The two-sided MWU test (scipy.stats.mannwhitneyu version 0.18.1) was performed on numerical data (such as BMI, age, and dietary component information) and Fisher's exact test (FisherExact 1.4.2) was performed on categorical data (such as medical history, gender, smoking status, and alcohol consumption status).

In addition, the associations coefficient (variance explained) between the demographic data (such as age, gender), medical history (such as history of diseases and drug usage) and clinical parameters (BMI, serum glucose and total cholesterol) as explanatory variables and the detected microbial features as response were tested by the multivariate associations with linear models (MaAsLin) R package (default parameters except for the fAllvAll=TRUE, dMinAbd=0.0, dMinSamp=0.05 and dSignificanceLevel = 0.1 were applied). The relative abundance of species and KEGG modules and metabolite concentrations were transformed by arcsin-square root transformation. The MaAsLin package was assessing the association using boosted additive generalized linear models with the calculation of significance level of multiple testing correction[17]. The significant associations were reported, and the associations were stated as significant if the P value was below 0.05 and FDR (q value) was below 0.1. Additionally, the variance explained calculated using the participant's group as explanatory variables (defined as crude coefficient) were also compared to those adjusted by potential confounder which are

Erawijantari PP, *et al. Gut* 2020;0:1–12. doi: 10.1136/gutjnl-2019-319188

(BMI, total cholesterol, diabetes medications status , and gastric acids medications, age, and gender) (strForcedPredictors=''BMI, total cholesterol, diabetes medications, and gastric acids medications, age,gender" was applied).

Oral microbes categorizations and microbial phenotype predictions

To examine the effect of gastrointestinal reconstruction, we further categorized the annotated microbiome by performing oral microbes categorizations and phenotype predictions. The categorizations of oral microbes were performed based on the expanded Human Oral Microbiome Database (eHOMD)[18]. The reference is in a tabularized view of all human oral microbial taxa that are defined and curated by eHOMD. The list of microbes retrieved from both mOTU and MetaPhlAn2 pipeline was categorized as oral microbes if it is listed on eHOMD and the non-listed species were categorized as others. The total relative abundance of oral microbes in each participant was calculated and compared between the gastrectomy (n=50) and control (n=56) groups using two-sided MWU test.

To reveal the phenotypic properties regarding the oxygen requirement of the species, we used the BugBase tool[19] on the metagenomics data. BugBase determines the proportions for each microbiome sample of Gram positive, Gram negative, biofilm-forming, pathogenic, potential mobile element-containing, oxygen utilizing, and oxidative stress-tolerant microorganisms. For the metagenomics shotgun sequencing data, we selected the OTUs using the IMG database before processing in BugBase. The detected phenotypes were plotted and the statistical P-values from two-sided MWU test between gastrectomy and control groups were also reported.

Correlation between genus and species in each group by SparCC

We estimated the microbial association (genus and species level) in each group using SparCC (bootstrap n=5000)[20]. First, we selected the significantly different genus and species between the gastrectomy and control groups that were overlapped in the mOTU and MetaPhlAn2 (LEfSe: $P<0.05$, $q<0.1$, $LDA>2.0$) pipeline annotation. SparCC correlation value between the genus or species counts (mOTU annotated) were then calculated in each group (control and gastrectomy) independently. SparCC has been widely used to estimate the correlation values from compositional data. Significant co-occurrence and co-excluding interaction (SparCC correlation scores $\rho<-0.2$ or $\rho>0.2$ for genus and

ρ<−0.4 or ρ>0.4 for species, with a P<0.05) were visualized and analyzed using *igraph*. We calculated the degree, betweenness and strength of each node to estimate its importance to the network.

Correlation between microbiome and metabolite in each group

We performed Procrustes analysis to determine the congruence of two-dimensional shapes produced by superimpositions of principal component analysis from microbiome and metabolome datasets based on the Euclidean distances of eigenvalues from both matrix (protest functions in vegan R package, n of permutations=1000)[21]. We selected significantly different (LEfSe: P<0.05; q<0.1; LDA>2 for comparison of gastrectomy(n=50) *vs* control (n=56)) metabolites and species that were overlap between mOTU and MetaPhlAn2 annotations for the analysis. First, we generated principal component from each measurement and then measured their inter-omics associations. Procrustes analysis superimpose and scales principal component plots and allows for quantification of non-random conformance between two different measurements from similar participants (**Supplementary Figure S9**).

Additionally, to evaluate the possible microbe–metabolite associations a Spearman rank-based correlation matrix was employed. Correlation analysis was performed on the relative abundance of 16 statistically-significant genus (LEfSe: P<0.05; q<0.1; LDA score ≥2.0 both confirmed by mOTU and MetaPhlAn2 pipelines, genus relative abundances annotated by mOTU pipeline were used for calculation) and the concentrations of 34 statistically-significant metabolites (LEfSe: P<0.05; q<0.1; LDA score≥3.0) among 44 gastrectomy patients and 54 controls participants. The correlation analysis was performed in Python 2.7 using spearmanr modules from the scipy.stats (version 0.18.0) package[22]. The P value of the correlation analysis was controlled by the BH-FDR multiple correction test and was stated as the q-value. The correlation was defined as significant if the P-value was below 0.05 and the FDR value was below 0.10. The correlation coefficients were between -1 and 1, for which 0 implied no correlation, while -1 or 1 implied an exact correlation. The cluster heatmap was created from the spearman correlation matrices (without q value thresholding) using the seaborn.clustermap based on the Euclidean distance metric. The results were visualized as a heatmap in which the red color implied a positive correlation and the blue color implied a negative correlation[23]. Significant correlations (P<0.05; q<0.10) were

indicated by an asterisk (*).

Predicting the species contribution on metabolite by Model-based Integration of Metabolite Observations and Species Abundances (MIMOSA)

The functional contribution of each species was retrieved from the HUMAnN2 annotation. The table of KO relative abundance stratified by species information was utilized to calculate the community metabolic potential (CMP) of each species based on the KEGG reaction information. This method integrates information about gene abundances in terms of KOs and KEGG reaction definition describing the quantitative relationship between genes and metabolites to provide an estimate of the way the species composition may impact each metabolite's abundance. The CMP score was then compared to the measured metabolomic data using Mantel test for each metabolite. The test assesses Spearman correlation between pairwise differences in CMP scores (across all pairs of samples) and the corresponding pairwise differences in measured metabolite concentration[24]. A significant positive pairwise correlation (Mantel test: $P \leq 0.01$; $q \leq 0.01$) was determined as "consistent" while a negative correlation was determined as "contrasting" trend. We also retrieved the information of key contributor species to the synthesis or degradation process for each metabolite (**Supplementary Figure 10, Supplementary Table 10**).

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

18 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.

19 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.

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

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

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

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

24 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**. doi:10.1128/mSystems.00013-15