



OPEN

Gut Microbiota Profiles of Treated Metabolic Syndrome Patients and their Relationship with Metabolic Health

Montree Wutthi-in^{1,2}, Supapon Cheevadhanarak^{2,3}, Sakawdaun Yasom^{4,5,6}, Sasiwan Kerdphoo^{4,5}, Parameth Thiennimitr^{4,5,6}, Arintaya Phrommintikul^{4,5,7}, Nipon Chattipakorn^{4,5}, Weerayuth Kittichotirat^{1,2}✉ & Siriporn Chattipakorn^{4,5,8}✉

Metabolic syndrome (MetS) has become a worldwide health issue. Recent studies reveal that the human gut microbiota exerts a significant role in the pathogenesis of this disease. While drug treatments may greatly improve metabolic symptoms, little is known about the gut microbiota composition of these treated MetS patients. This study aimed to characterize the gut microbiota composition of treated-MetS patients and analyse the possibility of using gut microbiota as an indicator of metabolic conditions. 16S rRNA metagenomic sequencing approach was used to profile gut microbiota of 111 treated MetS patients from The Cohort of patients at a high Risk of Cardiovascular Events (CORE)-Thailand registry. Our results show that the gut microbiota profiles of MetS patients are diverse across individuals, but can be classified based on their similarity into three groups or enterotypes. We also showed several associations between species abundance and metabolic parameters that are enterotype specific. These findings suggest that information on the gut microbiota can be useful for assessing treatment options for MetS patients. In addition, any correlations between species abundance and human properties are likely specific to each microbial community.

Metabolic syndrome (MetS) is characterized by a cluster of metabolic symptoms including obesity, hypertension, dyslipidemia, hyperglycemia, and insulin resistance¹. People with MetS have an increased risk of mortality from cardiovascular and cerebrovascular diseases, as well as being at greater risk of neurodegeneration². The criteria for diagnosing MetS are designated by values for obesity (e.g. waist circumference or body mass index (BMI)), triglyceride, high-density lipoprotein cholesterol (HDL-C), hypertension, and urine or blood albumin^{3,4}. According to American Heart Association criteria, approximately 35% of adults in United States and 50% of people who are older than 60 years have metabolic syndrome^{3,5}. For decades a great deal of research has been undertaken to better understand the factors that can lead to MetS. The development of MetS may be due to genetic factors as well as lifestyle including sedentary habits and high-fat/high-sugar dietary consumption⁶. Several recent studies have demonstrated that the human gut microbiota, the complex microbial community living inside the human gastrointestinal tract, plays a significant role in the pathogenesis of MetS^{7–11}. Previous studies have shown that interactions between gut microbiota and host metabolism could result in either promotion or protection of the host from metabolic disorders^{12–17}. In general, human and mice share two similar major phyla within their gut

¹Bioinformatics and Systems Biology Program, School of Bioresources and Technology and School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, 10150, Thailand. ²Systems Biology and Bioinformatics Research Group, Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Bangkok, 10150, Thailand. ³School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok, 10150, Thailand. ⁴Cardiac Electrophysiology Research and Training Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, 50200, Thailand. ⁵Center of Excellence in Cardiac Electrophysiology Research, Faculty of Medicine, Chiang Mai University, Chiang Mai, 50200, Thailand. ⁶Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, 50200, Thailand. ⁷Department of Internal Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, 50200, Thailand. ⁸Department of Oral Biology and Diagnostic Sciences, Faculty of Dentistry, Chiang Mai University, Chiang Mai, 50200, Thailand. ✉e-mail: weerayuth.kit@kmutt.ac.th; siriporn.c@cmu.ac.th

Parameters	Values
Male (%)	49 (44.14)
Age (years)	64.297 ± 8.419
Height (cm)	157.486 ± 7.352
Weight (kg)	68.358 ± 15.252
BMI (kg/m ²)	27.514 ± 5.764
Waist circumference (cm ²)	94.935 ± 13.394
Obese (%)	32 (28.83)
Hypertension (%)	97 (87.4)
Diabetes mellitus (%)	88 (79.28)
Chronic Kidney Disease (%)	23 (20.72)
Dyslipidemia (%)	96 (84.49)
Coronary artery disease (%)	35 (31.5)
Cerebrovascular disease (%)	29 (26.1)
Smoke frequency (days per week)	0.405 ± 0.578
Alcohol consumption (days per week)	0.486 ± 0.785
High-density lipoprotein cholesterol (mg/dL)	49.667 ± 16.771
Low-Density lipoprotein cholesterol (mg/dL)	94.467 ± 36.188
Very-low-density lipoprotein cholesterol (mg/dL)	28.111 ± 15.877
Triglyceride (mg/dL)	133.408 ± 74.089
Glucose (mg/dL)	130.140 ± 44.771
Hemoglobin A1C (mg/dL)	7.159 ± 1.463
Fibroblast growth factor 21 (pg/mL)	389.794 ± 508.598
Insulin (mg/dL)	7.988 ± 8.931
Homeostatic Model Assessment score	2.600 ± 3.097

Table 1. Sample demographic data of metabolic syndrome patients (N = 111) in this study. Categorical variables are expressed as frequency and percentage. Continuous variables are expressed as mean ± standard deviation.

microbiota, specifically Firmicutes, and Bacteroidetes¹⁸. The development of obesity has been shown to be associated with an increased Firmicutes/Bacteroidetes (F/B) ratio and higher Proteobacteria levels in the gastrointestinal tract^{19–22}. However, recent characterization studies based on large cohorts also demonstrated inconsistent findings where F/B ratios were not significantly different among subjects with different body weights^{10,23}. Besides that, Arumugam *et al.* showed that human gut microbiota variation can be stratified into a limited number of common microbial abundance profiles, called enterotypes, that might respond differently to external factors such as diets or drug intakes²⁴. While such stratification has been demonstrated to be useful in the analysis human gut microbiota, the result relies heavily on clustering criteria used in each study (e.g. choice of taxonomic levels, clustering algorithms, and distance functions) and therefore standardization of this approach is still an active area of research²⁵.

Clinically, a patient is considered to have MetS when three or more of the following five conditions exist, which are (i) waist circumference ≥90 cm in men or ≥80 cm in women, (ii) blood pressure ≥130/85 mmHg, (iii) triglycerides ≥150 mg/dl, (iv) HDL-C < 40 mg/dl in men or <50 mg/dl in women, and (v) fasting glucose ≥100 mg/dl²⁶. Various drugs are usually administered to treat MetS patients based on their metabolic symptoms. These include, for instance, Angiotensin converting enzyme inhibitor, Beta blocker, Calcium channel blocker, and Angiotensin receptor blocker for the treatment of high blood pressure, Statin, Fibrate, Niacin, and Ezetimibe for lowering lipid, Insulin, Sulfonyl urea, Metformin, Thiazolidinedione, and DPP4 inhibitor for lowering glucose, and Aspirin, Clopidogrel, and Warfarin as antithrombotic agents. While these drugs are very effective in improving patient outcomes in term of metabolic parameters, little is known about the condition of the gut microbiota of these treated patients. To this end, we hypothesize that common patterns of gut microbiota abundance or enterotypes exist among treated MetS patients and knowledge of these enterotypes may be useful for physicians to detect gut dysbiosis as well as understand the associations between microbial abundance and metabolic parameters. Therefore, the present study aimed to characterize the gut microbiota of clinically-treated MetS patients and identify existing enterotypes as well as relationship between microbiota profiles and metabolic outcomes of these patients.

Results

The demographic data of enrolled metabolic syndrome patients and metagenomic sequencing results. A total of 111 clinically-treated MetS patients were recruited into this study. Table 1 shows the demographic data of MetS patients together with some their clinical parameters. A summary of gut microbiota metagenomic sequencing result of these patients can be found in Supplementary Fig. S1 and Table S1. Briefly, the

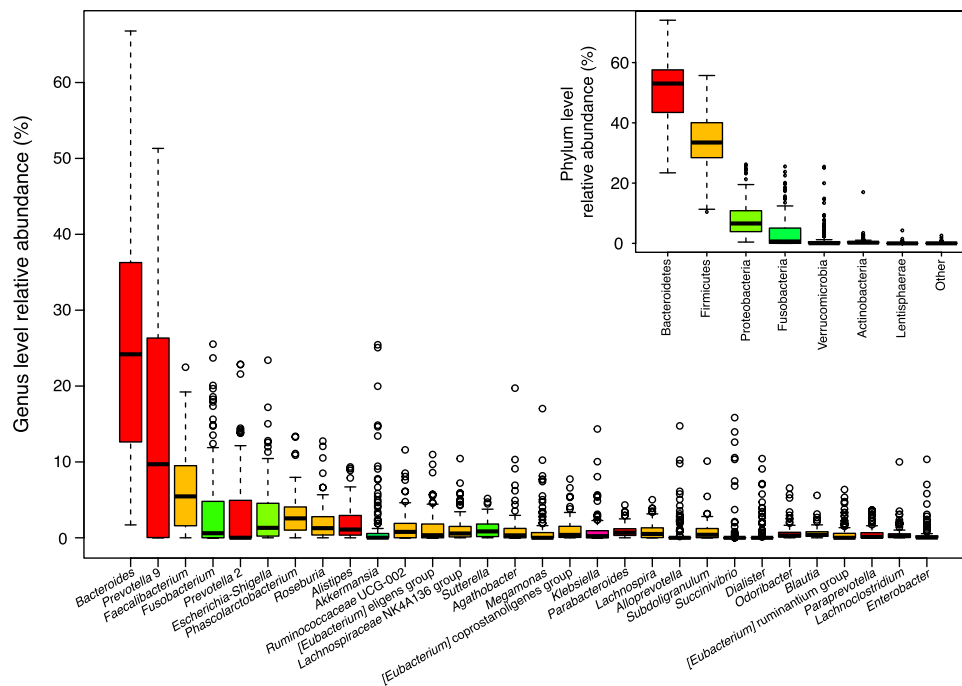


Figure 1. Box plots showing distributions of genera and phyla relative abundances identified across 111 metabolic syndrome patients. Only the top 30 genera in terms of relative abundance are shown in the plot. Relationship between each phylum and respective genera are shown using the same color bar.

sequencing of the V3–V4 hypervariable region of 16S rRNA gene produced between 36,400 and 83,728 paired-end reads that are 250 bases in length for each sample (mean number of reads of $66,557 \pm 8,460$). After pre-processing and amplicon sequence variant (ASV) identification, the number of remaining sequences for each sample was between 17,057 and 44,761 reads (mean number of reads of $30,797 \pm 5,239$) with a total of $8,754 \pm 1,239$ different ASVs assigned across all samples. Some of the most abundant bacterial phyla found within these samples were Bacteroidetes, Firmicutes, Proteobacteria, Fusobacteria, and Verrucomicrobia as summarized in the inset of Fig. 1 where the relative abundance values of each taxa for all samples provided in Supplementary Data 1. Also shown in Fig. 1, these common phyla appeared to be made up of genera including *Bacteroides*, *Prevotella 9*, *Faecalibacterium*, *Fusobacterium*, *Prevotella 2*, and *Escherichia-Shigella* as displayed by the same color bar. While *Bacteroides* was the found in high relative abundance in most samples, its presence also exhibited the greatest variation in terms of relative abundance, which is consistent with data reported in previous studies^{24,27}.

Variation of gut microbiota composition among metabolic syndrome patients. To better understand the gut microbiota variation among treated MetS patients, we first grouped their gut microbiota profiles based on the genus relative abundance similarity using a multidimensional clustering algorithm. Our clustering result revealed that gut microbiota profile of 111 MetS patients can be optimally classified into three groups or enterotypes that share similar genus abundance profiles (Supplementary Fig. S2). Specifically, enterotype 1, 2, and 3 could be assigned to 38, 36, and 37 patients respectively as shown by Fig. 2 and Supplementary Table S2. Principal component analysis showed that while most samples from each enterotype were well separated, a small number of samples appeared to be overlapping between enterotypes indicating that some microbiota profiles may possibly be classified to more than one enterotypes (Fig. 3a). The alpha diversity analysis using Shannon index showed that enterotype 1 exhibited the highest diversity when compared to other enterotypes ($p < 0.001$) as shown in Fig. 3b. Differential abundance analysis using ANalysis of COmposition of Microbiomes (ANCOM)²⁸ showed that *Ruminococcaceae* (UCG-002, UCG-005), *Prevotella 9*, and *Bacteroides* were significantly different between enterotypes (Supplementary Data 2) and their abundance variations can be used to identify enterotype 1, 2, and 3 respectively as shown by Fig. 3c. Additionally, differential abundance analysis using Linear discriminant analysis Effect Size (LEfSe²⁹) was carried out and a similar result was produced where *Prevotella 9* and *Bacteroides* were the top genera for enterotype 2 and 3 respectively in term of their differential abundance (Supplementary Data 3). On the other hand, LEfSe identified *Escherichia-Shigella* as the top genus in term of differential abundance for enterotype 1 followed by *Ruminococcaceae* UCG-002 and *Ruminococcaceae* UCG-010. This inconsistent result for enterotype 1 suggests that this enterotype may be less well defined by any single genus or could be further classified into smaller subgroups.

To further understand the potential influence of each enterotype's main contributor in overall gut microbiota, a correlation analysis of co-occurring genera was carried out and the result is shown in Fig. 3d. For enterotype 1, the enriched *Ruminococcaceae* UCG-002 and UCG-005 genera appeared to strongly co-occur with other diverse genera such as *Megamonas*, *Butyricimonas*, and *Christensenellaceae* R-7 group. On the other hand, the abundances of the enriched *Prevotella 9* in enterotype 2 and *Bacteroides* in enterotype 3 did not show a positive

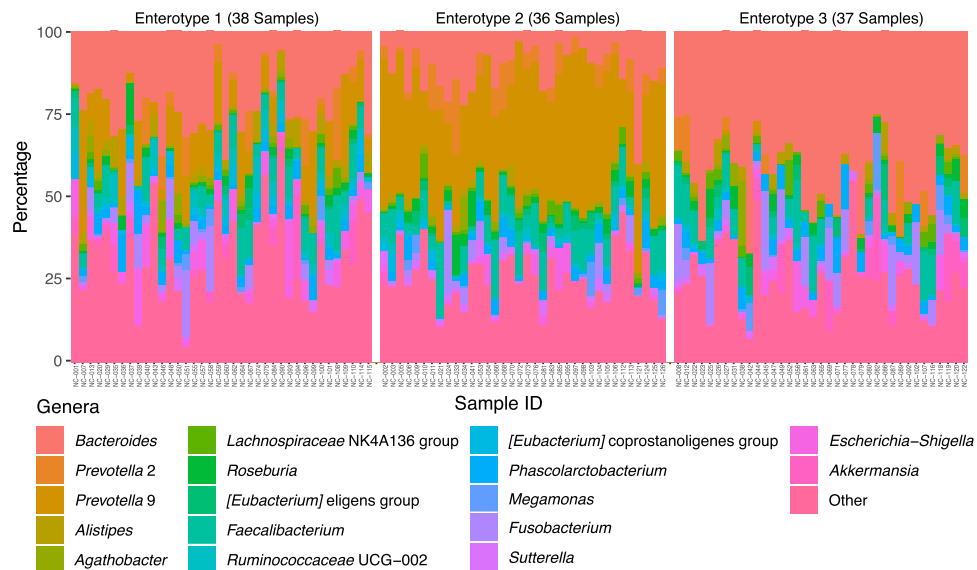


Figure 2. Genus relative abundance profiles of 111 metabolic syndrome patients separated by their enterotype assignments. Genera that exhibit less than 1 percent abundance across the samples are combined and shown as other.

correlation with the abundance of any other genera. Specifically, the presence of *Prevotella 9* in high relative abundance among enterotype 2 samples was found to negatively associated with the presence of genera including *Fusobacterium*, *Lachnospiraceae* NK4A136 group, and *Bacteroides*. Similarly, the high abundance of *Bacteroides* in enterotype 3 samples appeared to be negatively correlated with the abundance of *Eubacterium ventriosum* group, *Subdoligranulum*, *Akkermansia*, *Butyricimonas*, *Ruminiclostridium 6*, and *Ruminococcaceae* UCG-005 (Fig. 3d). The co-occurrence analysis results for enterotype 2 and 3 are consistent with the reduction in diversity found among these samples (Fig. 3b).

In term of MetS patients' clinical properties, parameters including smoke frequency, alcohol consumption, and F/B ratio (both qPCR and metagenomics) were significantly different between enterotypes (Supplementary Table S3). Specifically, MetS patients classified as enterotype 2 appeared to have significantly higher smoke frequency at 0.61 ± 0.69 days per week ($p = 0.041$) when compared to other enterotypes (0.37 ± 0.54 and 0.24 ± 0.43 days per week in enterotype 1 and 3, respectively). Similarly, alcohol consumption was also found to be significantly higher among enterotype 2 patients at 0.86 ± 0.87 days per week ($p = 0.001$) relative to 0.34 ± 0.71 days per week in enterotype 1 and 0.27 ± 0.65 days per week in enterotype 3. On the other hand, F/B ratio was found to be significantly increased among enterotype 1 samples (0.89 ± 0.46 , $p = 0.007$) compared to the other enterotypes (0.59 ± 0.16 and 0.67 ± 0.23 in enterotype 1 and 3 respectively). While variations could also be observed in other clinical parameters, their differences between enterotypes were not statistically significant (Supplementary Table S3). These include, for instance, age (enterotype 1: 64.45 ± 9.10 , enterotype 2: 63.36 ± 7.58 , and enterotype 3: 65.05 ± 8.62), BMI (enterotype 1: 27.35 ± 6.08 , enterotype 2: 28.49 ± 5.65 , and enterotype 3: 26.74 ± 5.56), HDL cholesterol (enterotype 1: 52.42 ± 19.54 , enterotype 2: 48.24 ± 16.73 , and enterotype 3: 48.23 ± 13.55), LDL cholesterol (enterotype 1: 100.84 ± 39.38 , enterotype 2: 93.97 ± 39.24 , and enterotype 3: 88.39 ± 28.95), glucose (enterotype 1: 130.86 ± 48.91 , enterotype 2: 128.82 ± 44.82 , and enterotype 3: 130.65 ± 41.62) and insulin (enterotype 1: 6.31 ± 5.49 , enterotype 2: 8.64 ± 7.37 , and enterotype 3: 9.06 ± 12.40).

Correlation analysis between bacterial abundance and metabolic parameters. Correlation analyses between each genus abundance and patient's clinical parameters were performed separately for each group of MetS patients (enterotype 1, 2, and 3) to identify associations between host clinical parameters and bacterial abundances. The results are summarized in Fig. 4 and the detailed results of Fig. 4 are shown in Supplementary Data 4. Notably, the same correlation analyses were also carried out using all gut microbiota data (111 samples) and no strong association ($r \geq 0.4$ or ≤ -0.4) between the abundance of any genus and clinical parameters were identified (Supplementary Data 5).

For enterotype 1 MetS patients, several significant associations were detected. Among all positive correlations, levels of interleukin 1 beta (IL 1B) and interleukin 18 (IL 18) showed the strongest associations with the abundance of *Faecalibacterium* ($r = 0.839$, $p < 0.001$) and *Agathobacter* ($r = 0.781$, $p = 0.003$) respectively. On the other hand, the abundance of *Parabacteroides* appeared to be associated to lipid profiles as the results showed significant positive correlations with parameters including total cholesterol ($r = 0.570$, $p < 0.001$), HDL-C ($r = 0.331$, $p = 0.048$), and low-density lipoprotein cholesterol (LDL-C) ($r = 0.428$, $p = 0.008$). Since these lipid related parameters are potentially collinear, we also analysed their correlations and found that total cholesterol and LDL-C were in fact highly correlated among enterotype 1 MetS patients ($r = 0.878$, $p < 0.001$) as shown in Supplementary Data 6. Among bacterial abundances that exhibit significant negative correlations with clinical parameters, *[Eubacterium] eligens* group appeared to be related to heart function as it is negatively correlated

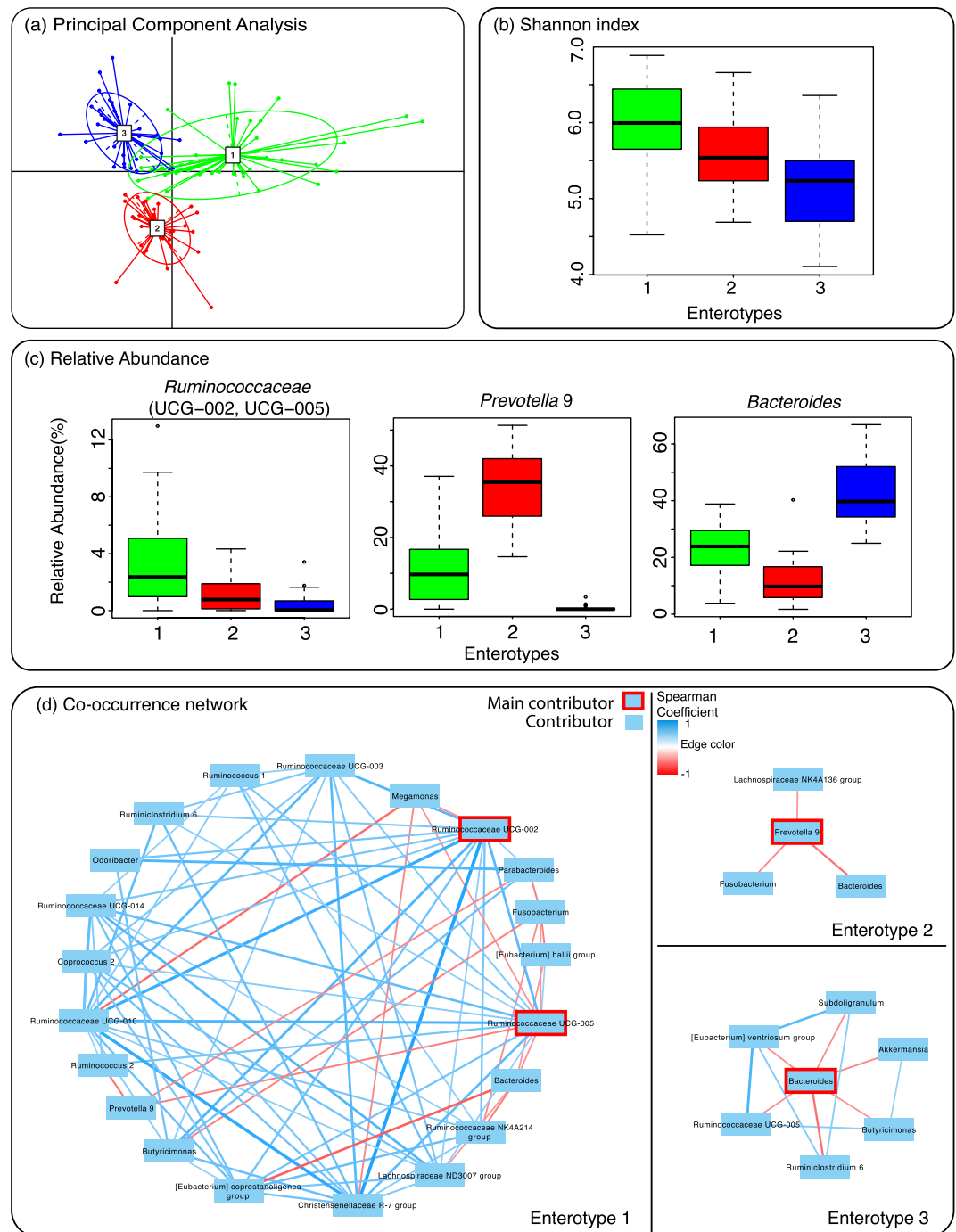


Figure 3. Variations in gut microbiota composition of metabolic syndrome patients. **(a)** Grouping of microbiota profiles based on similarity in species compositions resulted in 3 clusters or enterotypes as visualized by principal component analysis. **(b)** Diversity analysis using index for each enterotype of metabolic syndrome patients. **(c)** Relative abundance of main contributor in each enterotype. **(d)** Co-occurrence network of the main contributor(s) and other genera in each enterotype.

with several cardiac rhythm properties including percentage of intervals different from the preceding interval by at least 50 ms (pNN50) ($r = -0.418$, $p = 0.030$), percentage of intervals 50 ms longer than the preceding interval (pNN50a) ($r = -0.426$, $p = 0.034$), number of intervals different from the preceding interval by at least 50 ms (BB50) ($r = -0.441$, $p = 0.015$), and number of intervals at Least 50 ms Longer Than the Preceding Interval (BB50a) ($r = -0.515$, $p = 0.004$). However, we also note that correlation analyses of these cardiac rhythm properties showed that they are highly collinear with r values between 0.836 and 0.965 as shown in Supplementary Data 6.

Different correlation patterns between genus abundances and clinical parameters were found for enterotype 2 MetS patients as shown in Fig. 4. Of all the positive correlations, levels of Glyceraldehyde 3-phosphate

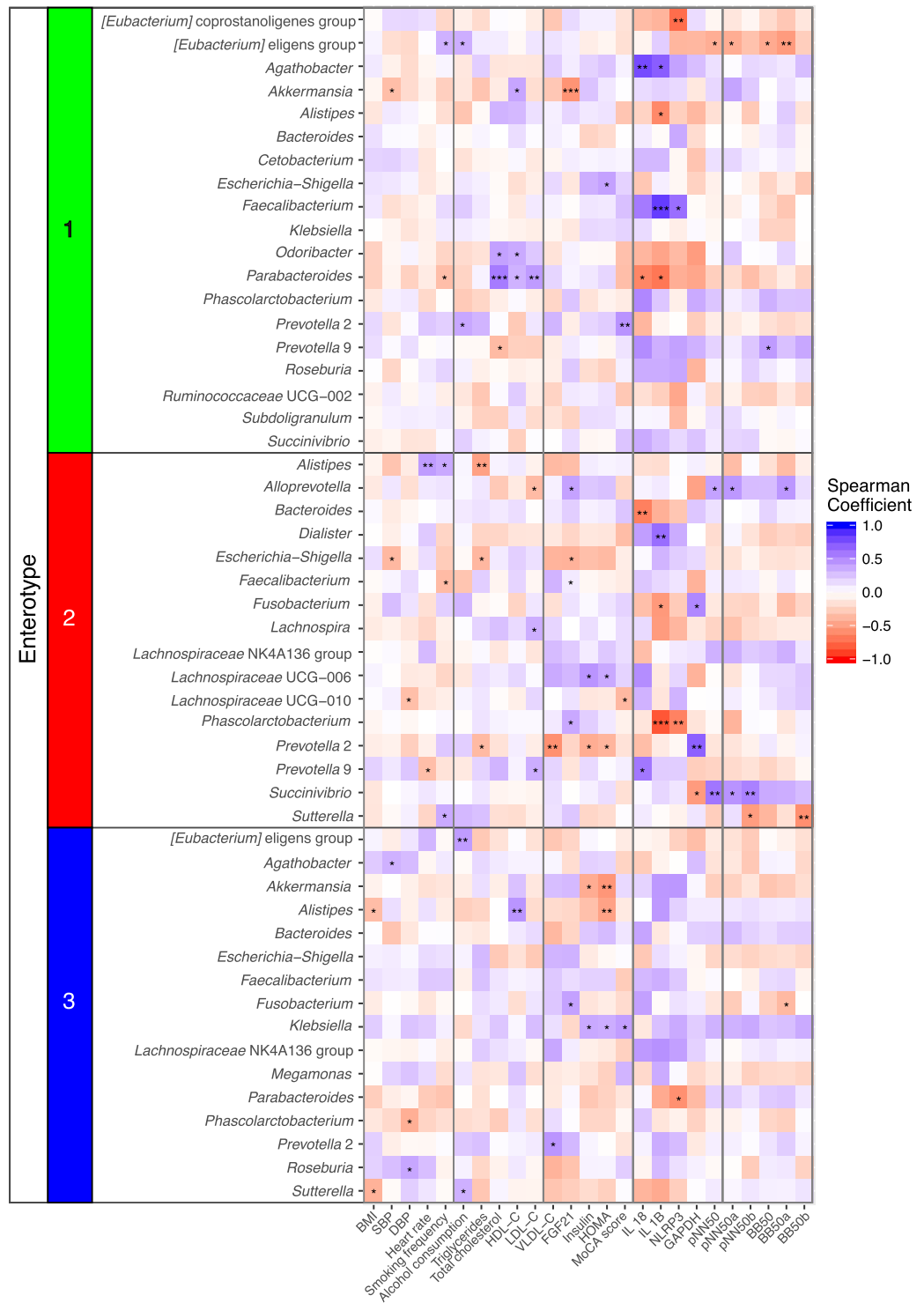


Figure 4. A heatmap showing the correlations between species abundance (rows) and clinical parameters (columns) for each enterotypes of metabolic syndrome patients. Positive correlations are shown using a gradient of blue colors. Negative correlations are shown using a gradient of red colors. Correlations with p-values that are statistically significant are noted by asterisk(s). See Methods section for meaning of all abbreviations.

dehydrogenase (GAPDH) and IL 1B exhibit the strongest association with the abundance of *Prevotella 2* ($r = 0.690$, $p = 0.003$) and *Dialister* ($r = 0.632$, $p = 0.009$) respectively. For this enterotype, the abundance of genus *Alloprevotella* showed the highest number of significant positive correlations with clinical parameters including Fibroblast growth factor 21 (FGF21) ($r = 0.398$, $p = 0.029$), pNN50 ($r = 0.402$, $p = 0.042$), pNN50a ($r = 0.444$, $p = 0.039$), and BB50a ($r = 0.450$, $p = 0.027$). While the cardiac rhythm values can be collinear against one another as observed in enterotype 1, they did not exhibit high collinearity with the Fibroblast growth factor 21

value (r values between 0.246 and 0.619). On the other hand, our analyses also reveal that some genera exhibit strong negative correlations with clinical parameters for this group of patients. Specifically, *Phascolarctobacterium* abundance was found to have negative correlations with IL 1B ($r = -0.818$, $p = 0.001$) and NOD-like receptor family, pyrin domain-containing protein 3 inflammasome (NLRP3) levels ($r = -0.633$, $p = 0.009$) where IL 1B and NLRP3 levels exhibit positive correlation of $r = 0.715$ ($p = 0.002$). Similarly, *Bacteroides* abundance showed a significant negative correlation with the IL 18 level ($r = -0.645$, $p = 0.007$).

For enterotype 3 MetS patients, a fewer number of significant correlations between genus abundances and clinical parameters were found when compared to other enterotypes (Fig. 4). Among these, genera such as [*Eubacterium*] *eligens* group and *Alistipes* exhibited strong positive correlations with alcohol consumption ($r = 0.425$, $p = 0.009$) and HDL-C ($r = 0.459$, $p = 0.006$) level respectively. Notably, genus *Klebsiella* appeared to have significant positive correlations with several clinical parameters including insulin level ($r = 0.370$, $p = 0.037$), Homeostatic Model Assessment (HOMA) score ($r = 0.350$, $p = 0.049$), and Montreal Cognitive Assessment (MoCA) score ($r = 0.415$, $p = 0.016$) where insulin value and HOMA score appeared to be collinear with one another ($r = 0.934$, $p < 0.001$). On the other hand, HOMA also appeared to be significantly associated with *Akkermansia* ($r = -0.537$, $p = 0.002$), and *Alistipes* ($r = -0.479$, $p = 0.006$) in a negative manner.

Discussion

The major findings of this study are: (1) gut microbiota profiles of 111 treated MetS patients are highly diverse across individuals (beta diversity); (2) although gut microbiota profiles of these MetS patients are diverse, they can be classified into three groups or enterotypes based on the similarity of their genera compositions where enterotype 1 is identifiable by *Ruminococcaceae* UCG-002 and UCG-005 abundances, enterotype 2 is identifiable by *Prevotella* 9 abundance, and enterotypes 3 is identifiable by *Bacteroides* abundance. (3) In relation to clinical parameters, enterotype 2 appeared to be significantly enriched in MetS patients with high levels of smoking frequency and alcohol consumption. (4) Many significantly strong associations between clinical parameters and genus abundances were found to be enterotype-specific. For instance, the IL 1B level was positively correlated with the abundance of *Faecalibacterium* in enterotype 1 group whereas the same parameter was positively correlated with *Dialister* in enterotype 2 group.

Metabolic syndrome (MetS) has become a global public health issue. Previous studies have shown that gut microbiota plays an important role in the pathogenesis of this disease^{30–33}. Better understanding of gut microbiota variability among clinically-treated MetS patients can potentially be useful in assessing and classifying patients as well as informing more suitable treatment options. Our result revealed that gut microbiota of treated MetS patients can be classified into three distinct enterotypes (Fig. 2) similar to the previously published result of human enterotypes by Arumugam *et al.*²⁴. The differential abundance analysis results show that enterotype 1 group is enriched with *Ruminococcaceae*, a genus that has been linked to the reduction of BMI³⁴. On the other hand, enterotype 2 group was enriched with genus *Prevotella* 9. Comparisons of these *Prevotella* 9 sequences with the NCBI database using BLAST shows that *Prevotella copri* DSM 18205 was repeatedly the best hit (387 out of 390 sequences) with 96.59% average pairwise sequence similarity (Supplementary Data 7). *Prevotella* has been shown to be a large genus that is made up of very genetically diverse species^{35,36}. In addition, presence of *Prevotella copri* has been reported to increase the risk of diarrhea and irritable bowel syndrome^{37,38}. Lastly, genera including *Bacteroides*, *Hungatella*, and *Akkermansia* were enriched in gut microbiota of enterotype 3 MetS patients. Genus *Akkermansia* contains probiotic bacterial species such as *Akkermansia muciniphila*³⁹, which has previously been linked to mucus layer-protection and is known to possess potential anti-inflammatory properties, both of which can improve gut function⁴⁰. Further comparisons of *Akkermansia* sequences from this enterotype against NCBI database showed that *Akkermansia muciniphila* was repeatedly the best hit (64 out of 66 sequences) with 95.13% average of pairwise sequence similarity (Supplementary Data 8). The enrichment of this bacterium may be beneficial to enterotype 3 MetS patients. *Bifidobacterium* spp., which is a known beneficial bacterium⁴¹, was not significantly enriched in any of the three groups of MetS patients (Supplementary Data 2 and Data 3).

Most parameters that are directly related to MetS (such as BMI, LDL-C, triglyceride, glucose, and insulin) are not significantly different between the three enterotypes (Supplementary Table S3). This could be due to the fact that MetS patients in this study were clinically treated and their metabolic parameters were well controlled with their medications. However, our results show that enterotype 2 MetS patients have significantly higher alcohol consumption and smoking frequency when compared to patients in another enterotypes. This result may suggest potential relationship between these behaviors and the enrichment of genus *Prevotella* 9 in these patients.

To better understand the possible effects of gut microbiota on patient physiology, correlation analyses between each genus abundance and clinical parameters were carried out. When the analysis was performed using the microbiota profile of all 111 MetS patients at both phylum and genus levels, we did not find any significantly strong correlation ($r \geq 0.4$ or ≤ -0.4) between taxon and clinical parameters. However, when the analyses were done separately for each group of patients, many interesting correlations were identified as summarized in Fig. 4 and Supplementary Data 4. Notably, the same genus abundance may show significant positive or negative correlations with different metabolic parameters in different enterotypes of MetS patients. For instance, our result shows that the genus *Faecalibacterium* exhibited a strong significant positive correlation with the IL 1B level in enterotype 1 group while it exhibits a significant negative correlation with smoking frequency in enterotype 2 group. This example shows that the correlations between each genus abundance and clinical parameters may potentially be enterotype/group specific. On the other hand, it is also possible that the same genus presented in different groups may represent different bacterial species or strains. More detailed analyses using an approach such as whole-genome shotgun metagenomic sequencing to more accurately determine bacterial species/strains as well as functional genes that are present in the gut microbiota of MetS patients will be required to better understand these microbes-host correlations.

In summary, we have shown that gut microbiota profiles may be useful for classifying MetS patients. We have also described examples where several correlations between microbial abundance and patients' clinical parameters are group/enterotype specific. Despite these results, we note that our study does contain some limitations. Since, our results are based on 16S rRNA metagenomics data of a cohort from the same region (Northern part of Thailand), usage of the findings reported here must be proceeded with cautions. More studies based on a larger and ethnically different population using whole-genome shotgun metagenomic sequencing and metabolomics approaches should be carried out to further explore variations at functional levels (e.g. genes, and metabolites) within the gut microbial community to better understand their relationship with host parameters. While several enterotype-specific associations between individual genera and clinical parameters that we discovered may be relevant for personalised treatment, the mechanisms driving these associations are still unclear and our current data cannot delineate the exact cause-effect of these relationships. In addition, we note that while enterotype concept has been previously demonstrated to be useful in analysing the complex human gut microbiota, Costea and colleagues, for instance, has shown the controversy of this concept and standardization of this approach is still an active area of research²⁵.

Methods

Subjects selection. The study protocol was approved by the institutional Ethics Committee of the Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand. Written informed consents were obtained from all patients (n = 111) prior to the study. All methods were performed in accordance with the relevant guidelines and regulations. This study is a sub-study of metabolic syndrome (MetS) patients in The Cohort Of patients at a high Risk for Cardiovascular Events (CORE) Thailand registry, which is an ongoing prospective cohort of Thai patients with a high atherosclerotic risk. Patients were recruited from outpatient clinic at Maharaj Nakorn Chiang Mai Hospital during the period between April 2011 and March 2014 using these criteria: 1) 45 years or older and 2) diagnosed with coronary artery disease (CAD), cerebrovascular disease (CVD), peripheral arterial disease (PAD), or at least 3 atherosclerosis risk factors (multiple risk factors, MRF). Some patients were excluded from the cohort for the following reasons: had an acute atherosclerotic event within 3 months, had a large aortic aneurysm indicated for surgery, participating in a blinded clinical trial, had a limited life expectancy from a non-cardiovascular condition such as cancer or a documented human immunodeficiency virus (HIV) infection, or might have difficulty returning for a follow-up visit. Patients were invited to participate in this study when they met three or more of the following five criteria for metabolic syndrome: elevated waist circumference (≥ 90 cm in men and ≥ 80 cm in women), blood pressure $\geq 130/85$ mmHg or treated, triglycerides ≥ 150 mg/dl or treated, high density lipoprotein cholesterol (HDL-C) (< 40 mg/dl in men and < 50 mg/dl in women), and fasting glucose ≥ 100 mg/dl or treated²⁶.

Data and sample collection. MetS patients were first medically treated to control their metabolic symptoms before clinical data and sample collections were performed. All data and sample collections were carried out on the same day. These include history taking, physical examination, blood sampling and stool collection. The homeostasis model assessment (HOMA) index was used to assess the severity of peripheral insulin resistance as previously described⁴². Each patient's stool sample was collected in a sterile plastic cup and kept in a refrigerator before sending to the Maharaj Nakorn Chiang Mai hospital within 24 hours.

Metabolic parameter determination. Evaluation and calculation of the fasting plasma glucose, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), triglyceride, insulin and fibroblast growth factor 21 (FGF21) levels were carried out for all fasting blood samples that were obtained from participants. Colorimetric assay (ERBA diagnostic, Mannheim, Germany) was used to determine the fasting plasma glucose and triglyceride levels. A sandwich enzyme-linked immunosorbent assay (ELISA) kit (Millipore, MI, USA) was used to evaluate the fasting plasma insulin levels.

DNA extraction. A commercial genomic DNA isolation kit (QIAGEN, Germany) was used to extract bacterial genomic DNA from patient fecal samples. Briefly, the fecal sample (0.25 g) was homogenized in QIAGEN ASL lysis buffer using a Minibeadbeater (BioSpec Products, Bartlesville, US). Then, following the manufacturer's instructions, the bacterial genomic DNA was extracted from the feces. The fractions of bacterial microbiota population representing the Firmicutes/Bacteroidetes (F/B) ratio were quantified using quantitative polymerase chain reaction (qPCR) as described previously^{43,44}.

PCR amplification and sequencing of the 16S rRNA Gene. Amplification of the V3–V4 hypervariable region of the microbial 16S rRNA gene was performed using the universal primers 341 F, 5'ACACTGACGACATGGTTCTACACCTACGGGNGGCWGCAG-3' and 805 R, 5'-TACGGTAGCAGAGACTTGGTCTGACTACHVGGGTATCTAATCC-3'. Amplification consisted of the following steps: an initial denaturing at 94 °C for 3 minutes followed by 35 cycles of 94 °C for 45 seconds, 50 °C for 1 minute, 72 °C for 90 seconds, and an elongation step of 72 °C for 10 minutes⁴⁵. After PCR amplification, paired-end sequencing was performed using the Illumina MiSeq platform at Génome Québec, Canada.

Gut microbiota analysis. The raw paired-end sequences were analysed by using the Quantitative Insights Into Microbial Ecology 2 (QIIME2; version 2019.7.0) package^{46,47}. The DADA2 pipeline within QIIME2 was used to filter low-quality reads, denoise sequences, join paired-end reads, and remove chimeric sequences⁴⁸. The taxonomy was assigned using the scikit-learn naive Bayes machine-learning classifier⁴⁹ where the classifier was created by using sequences from the SILVA database version 132⁵⁰. Rarefying normalization⁵¹ was applied to normalize all samples using the 15,000 reads per sample criterion. To group samples based on similar microbiota profiles at genus level, a clustering method previously described by Arumugam *et al.* was applied²⁴. Briefly, the

partitioning around medoid (PAM) clustering algorithm with Jensen–Shannon distance was employed to cluster samples based on relative genus abundances where the optimal number of clusters was determined by using Calinski–Harabasz (CH) index. The clustering result was then used to assign samples into groups that share similar relative genus abundances (enterotypes). Alpha diversity of each enterotype was determined using Shannon index. Finally, the enterotypes of all MetS samples were visualized using principal component analysis.

Statistical analysis. Continuous variables are expressed as mean \pm standard deviation. The analysis of non-normally distributed variables was done using the Mann–Whitney U and Kruskal–Wallis Tests^{52,53}. Frequencies and percentages are used for all of the categorical variables. Pearson χ^2 test was used to compare between groups. The differential abundance test between groups was done using analysis of composition of microbiomes (ANCOM) and linear discriminant analysis effect size (LEfSe)^{28,29}. The metadata of samples used in this study consisted of the non-normally distributed variables such as age, height, weight, body mass index (BMI), waist circumference, smoking frequency, alcohol consumption frequency, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), very-low-density lipoprotein cholesterol (VLDL-C), heart rate, systolic blood pressure (SBP), diastolic blood pressure (DBP), triglyceride, total cholesterol, glucose, insulin, homeostatic model assessment (HOMA) score, Montreal cognitive assessment (MoCA) score, interleukin 18 (IL 18), interleukin 1 beta (IL 1B), NOD-like receptor family, pyrin domain-containing protein 3 inflammasome (NLRP3), and glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Cardiovascular disease and co-morbidities as well as cardiac rhythms were also recorded. The correlations between the microbial abundances and clinical parameters were analysed using Spearman’s rank correlation coefficient⁵⁴. Note that species with abundance of less than 1 percent were excluded from the correlation analysis. Spearman’s rank correlation was also used to compare clinical parameters to identify collinearity between parameter values. P values of less than 0.05 based on 2-tailed probability tests were regarded as statistically significant. All statistical calculations were performed using the SPSS version 17 (SPSS Inc., IL, USA).

Data availability

The datasets used in this study are available from the corresponding authors on reasonable request.

Received: 12 November 2019; Accepted: 28 May 2020;

Published online: 22 June 2020

References

1. Nguyen, D. M. & El-Serag, H. B. The epidemiology of obesity. *Gastroenterology Clinics* **39**, 1–7 (2010).
2. Mathieu, P., Pibarot, P. & Després, J.-P. Metabolic syndrome: the danger signal in atherosclerosis. *Vascular health and risk management* **2**, 285 (2006).
3. Srikanthan, K., Feyh, A., Visweshwar, H., Shapiro, J. I. & Sodhi, K. Systematic Review of Metabolic Syndrome Biomarkers: A Panel for Early Detection, Management, and Risk Stratification in the West Virginian Population. *Int J Med Sci* **13**, 25–38, <https://doi.org/10.7150/ijms.13800> (2016).
4. Ford, E. S. The metabolic syndrome and mortality from cardiovascular disease and all-causes: findings from the National Health and Nutrition Examination Survey II Mortality Study. *Atherosclerosis* **173**, 309–314, <https://doi.org/10.1016/j.atherosclerosis.2003.12.022> (2004).
5. Aguilar, M., Bhuket, T., Torres, S., Liu, B. & Wong, R. J. Prevalence of the metabolic syndrome in the United States, 2003–2012. *JAMA* **313**, 1973–1974, <https://doi.org/10.1001/jama.2015.4260> (2015).
6. Harsch, I. A. & Konturek, P. C. The Role of Gut Microbiota in Obesity and Type 2 and Type 1 Diabetes Mellitus: New Insights into “Old” Diseases. *Med Sci (Basel)* **6**, <https://doi.org/10.3390/medsci6020032> (2018).
7. Pedersen, H. K. *et al.* Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* **535**, 376–381 (2016).
8. Kootte, R. S. *et al.* Improvement of insulin sensitivity after lean donor feces in metabolic syndrome is driven by baseline intestinal microbiota composition. *Cell metabolism* **26**, 611–619. e616 (2017).
9. Lippert, K. *et al.* Gut microbiota dysbiosis associated with glucose metabolism disorders and the metabolic syndrome in older adults. *Beneficial microbes* **8**, 545–556 (2017).
10. Org, E. *et al.* Relationships between gut microbiota, plasma metabolites, and metabolic syndrome traits in the METSIM cohort. *Genome biology* **18**, 70 (2017).
11. He, Y. *et al.* Linking gut microbiota, metabolic syndrome and economic status based on a population-level analysis. *Microbiome* **6**, 1–11 (2018).
12. Ursell, L. K., Metcalf, J. L., Parfrey, L. W. & Knight, R. Defining the human microbiome. *Nutr Rev* **70**(Suppl 1), S38–44, <https://doi.org/10.1111/j.1753-4887.2012.00493.x> (2012).
13. McKenzie, C., Tan, J., Macia, L. & Mackay, C. R. The nutrition-gut microbiome-physiology axis and allergic diseases. *Immunological reviews* **278**, 277–295 (2017).
14. Mazidi, M., Rezaie, P., Kengne, A. P., Mobarhan, M. G. & Ferns, G. A. Gut microbiome and metabolic syndrome. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **10**, S150–S157 (2016).
15. Turnbaugh, P. J. & Gordon, J. I. The core gut microbiome, energy balance and obesity. *The Journal of physiology* **587**, 4153–4158 (2009).
16. Stephens, R. W., Arhire, L. & Covasa, M. Gut microbiota: from microorganisms to metabolic organ influencing obesity. *Obesity* **26**, 801–809 (2018).
17. Kho, Z. Y. & Lal, S. K. The human gut microbiome—a potential controller of wellness and disease. *Frontiers in microbiology* **9** (2018).
18. Nguyen, T. L. A., Vieira-Silva, S., Liston, A. & Raes, J. How informative is the mouse for human gut microbiota research? *Disease models & mechanisms* **8**, 1–16 (2015).
19. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480–U487, <https://doi.org/10.1038/nature07540> (2009).
20. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology - Human gut microbes associated with obesity. *Nature* **444**, 1022–1023, <https://doi.org/10.1038/nature4441022a> (2006).
21. Armougom, F., Henry, M., Viallettes, B., Raccach, D. & Raoult, D. Monitoring Bacterial Community of Human Gut Microbiota Reveals an Increase in Lactobacillus in Obese Patients and Methanogens in Anorexic Patients. *Plos One* **4**, <https://doi.org/10.1371/journal.pone.0007125> (2009).
22. Vazquez, G., Duval, S., Jacobs, D. R. & Silventoinen, K. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: A meta-analysis. *Epidemiol Rev* **29**, 115–128, <https://doi.org/10.1093/epirev/mxm008> (2007).

23. Falony, G. *et al.* Population-level analysis of gut microbiome variation. *Science* **352**, 560–564 (2016).
24. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174 (2011).
25. Costea, P. I. *et al.* Enterotypes in the landscape of gut microbial community composition. *Nature microbiology* **3**, 8–16 (2018).
26. Alberti, K. G. M. M. *et al.* Harmonizing the Metabolic Syndrome A Joint Interim Statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* **120**, 1640–1645, <https://doi.org/10.1161/Circulationaha.109.192644> (2009).
27. Tap, J. *et al.* Towards the human intestinal microbiota phylogenetic core. *Environmental microbiology* **11**, 2574–2584 (2009).
28. Mandal, S. *et al.* Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial ecology in health and disease* **26**, 27663 (2015).
29. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome biology* **12**, R60 (2011).
30. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–1023, <https://doi.org/10.1038/4441022a> (2006).
31. Furet, J. P. *et al.* Differential adaptation of human gut microbiota to bariatric surgery-induced weight loss: links with metabolic and low-grade inflammation markers. *Diabetes* **59**, 3049–3057, <https://doi.org/10.2337/db10-0253> (2010).
32. Schwartz, A. *et al.* Microbiota and SCFA in lean and overweight healthy subjects. *Obesity (Silver Spring)* **18**, 190–195, <https://doi.org/10.1038/oby.2009.167> (2010).
33. Larsen, N. *et al.* Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *Plos One* **5**, e9085, <https://doi.org/10.1371/journal.pone.0009085> (2010).
34. Wang, F. G. *et al.* Differential composition of gut microbiota among healthy volunteers, morbidly obese patients and post-bariatric surgery patients. *Experimental and therapeutic medicine* **17**, 2268–2278 (2019).
35. Gupta, V. K., Chaudhari, N. M., Iskepalli, S. & Dutta, C. Divergences in gene repertoire among the reference Prevotella genomes derived from distinct body sites of human. *Bmc Genomics* **16**, 153 (2015).
36. Ley, R. E. Gut microbiota in 2015: Prevotella in the gut: choose carefully. *Nature reviews Gastroenterology & hepatology* **13**, 69 (2016).
37. Su, T. T. *et al.* Altered Intestinal Microbiota with Increased Abundance of Prevotella Is Associated with High Risk of Diarrhea-Predominant Irritable Bowel Syndrome. *Gastroent Res Pract*, <https://doi.org/10.1155/2018/6961783> (2018).
38. Larsen, J. M. The immune response to Prevotella bacteria in chronic inflammatory disease. *Immunology* **151**, 363–374, <https://doi.org/10.1111/imm.12760> (2017).
39. Dao, M. C. *et al.* Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. *Gut* **65**, 426–436, <https://doi.org/10.1136/gutjnl-2014-308778> (2016).
40. Derrien, M., Belzer, C. & de Vos, W. M. Akkermansia muciniphila and its role in regulating host functions. *Microb Pathogenesis* **106**, 171–181, <https://doi.org/10.1016/j.micpath.2016.02.005> (2017).
41. Sugahara, H. *et al.* Probiotic Bifidobacterium longum alters gut luminal metabolism through modification of the gut microbial community. *Scientific reports* **5**, 13548 (2015).
42. Matthews, D. *et al.* Homeostasis model assessment: insulin resistance and β -cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28**, 412–419 (1985).
43. Rinttila, T., Kassinen, A., Malinen, E., Krogus, L. & Palva, A. Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. *Journal of applied microbiology* **97**, 1166–1177, <https://doi.org/10.1111/j.1365-2672.2004.02409.x> (2004).
44. Winter, S. E. *et al.* Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature* **467**, 426–429, <https://doi.org/10.1038/nature09415> (2010).
45. Herlemann, D. P. R. *et al.* Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *Isme J* **5**, 1571–1579, <https://doi.org/10.1038/ismej.2011.41> (2011).
46. Bolyen, E. *et al.* QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. Report No. 2167-9843, (PeerJ Preprints, 2018).
47. Almeida, A., Mitchell, A. L., Tarkowska, A. & Finn, R. D. Benchmarking taxonomic assignments based on 16S rRNA gene profiling of the microbiota from commonly sampled environments. *Gigascience* **7**, <https://doi.org/10.1093/gigascience/giy054> (2018).
48. Callahan, B. J. *et al.* DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods* **13**, 581 (2016).
49. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J Mach Learn Res* **12**, 2825–2830 (2011).
50. Pruesse, E. *et al.* SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**, 7188–7196, <https://doi.org/10.1093/nar/gkm864> (2007).
51. Weiss, S. *et al.* Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5**, <https://doi.org/10.1186/s40168-017-0237-y> (2017).
52. McKnight, P. E. & Najab, J. J. T. C. e. o. p. Mann-Whitney U Test. 1-1 (2010).
53. Breslow, N. A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship. *Biometrika* **57**, 579–594 (1970).
54. Hauke, J. & Kossowski, T. J. Qg Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones geographicae* **30**, 87–93 (2011).

Acknowledgements

This work was supported by Thailand Research Fund TRF-RTA6080003 (S.C.) and Senior Research Scholar grant from the National Research Council of Thailand (S.C.), a NSTDA Research Chair Grant from the National Science and Technology Development Agency Thailand (N.C.) and Chiang Mai University Excellence Center Award (N.C.). The authors also acknowledge the financial support provided by King Mongkut's University of Technology Thonburi through the "KMUTT 55th Anniversary Commemorative Fund (W.K.)" and Petchra Pra Jom Kloa Master's Degree Research Scholarship from King Mongkut's University of Technology Thonburi (M.W.).

Author contributions

M.W., W.K., N.C., and S.C. conceived and design the project. M.W., Su.C., S.Y., S.K., P.T., A.P. and W.K. acquired the data. M.W., Su.C., P.T., W.K., A.P., N.C. and S.C. analysed and interpreted the data. M.W., Su.C., P.T., W.K., A.P., N.C. and S.C. wrote and reviewed the manuscript. All authors read, revised and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-67078-3>.

Correspondence and requests for materials should be addressed to W.K. or S.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020