# A single dietary factor, daily consumption of a fermented beverage, can modulate the gut microbiome within the same ethnic community

**Santanu Das [a,b,#], Ezgi Özkurt [c,d,#], Tulsi Kumari Joishy [a], Dibyayan Deb [a], Ashis K. Mukherjee [a], Falk Hildebrand [c,d *] Mojibur R. Khan [a]***

[a] Life Sciences Division, Institute of Advanced Study in Science and Technology (IASST), Guwahati-781035, Assam, India

[b] Department of Molecular Biology and Biotechnology, Cotton University, Guwahati- 781001, Assam, India

[c] Gut microbes and health, Quadram Institute Bioscience, Norwich, UK

[d] Digital Biology, Earlham Institute, Norwich, UK

# Both the authors have contributed equally

*Corresponding authors:

Dr. M. R. Khan, Life Sciences Division, Institute of Advanced Study in Science and Technology (IASST), Guwahati-781035, Assam, India. Email: mojibur.khan@gmail.com

Dr. Falk Hildebrand, Gut Microbes & Health, Quadram Institute Bioscience, Norwich Research Park, Norwich, Norfolk, NR4 7UA, UK & Earlham Institute, Norwich ResearchPark, Norwich, Norfolk, NR4 7UA, UK. Email: falk.hildebrand@quadram.ac.uk

**Abstract:**

In this study, the impact of traditional rice-based fermented alcoholic beverages (*Apong*) on the gut microbiome and health of the *Mishing* community in India was examined. Two groups that consumed one of these beverages were compared to a control group that did not consume either beverage. Gut microbial composition was analyzed by sequencing 16S rRNA of fecal metagenomes and analyzing untargeted fecal metabolites, and short-chain fatty acids (SCFAs). We also collected data on anthropometric measures and serum biochemical markers. Our results showed that *Apong* drinkers had higher blood pressure, but lower blood glucose and total protein levels than other non-drinkers. Also, gut microbiome composition was found to be affected by the choice of *Apong*, with *Apong* drinkers having a more diverse and distinct microbiome compared to non-drinkers. *Apong* drink type or being a non-drinker explained even a higher variation of fecal metabolome composition than microbiome composition and *Apong* drinkers had lower levels of the SCFA isovaleric acid than non-drinkers. Overall, this study shows that a single dietary factor can significantly impact the gut microbiome of a community and highlights the potential role of traditional fermented beverages in maintaining gut health.

**Key words:** Gut microbiome, alcoholic beverage, fermented beverage, short chain fatty acids, fecal metabolites

## Introduction:

The gut microbiome is a collection of microorganisms in the human intestine that performs many important functions, including digestion, transformation of nutrients, and immune system regulation[1]. A well-balanced gut microbiome composition can provide health benefits, while imbalances can lead to disorders related to metabolism and immunity[2]. The gut microbiome is influenced by various factors, including diet, medication, and age. Diet is a particularly important factor that affects the gut microbiome composition and its interactions with the host[5-7]. Fermented foods and beverages, such as yogurt, kefir, fermented cottage cheese, kimchi, and kombucha tea, are rich in microorganisms that can have an effect on the gut microbiome and increase overall microbial diversity[8-12].

Rice-based fermented beverages are an important part of the diet and cultural heritage of Mongoloid communities in South-East Asia[13]. In Assam, India, the Mishing community consumes two types of rice-based alcoholic beverages called *Poro Apong* and *Nogin Apong*. *Poro Apong* is prepared with roasted rice, ash of rice husk, and a starter cake, and undergoes solid-state fermentation for 7-10 days. It is filtered through ash to produce a dark, clear liquid with physical and sensory properties similar to stout beer. *Nogin Apong* is made with steamed rice, and has physical and sensory properties similar to *Maakoli*, a fermented beverage from South Korea. Within the Mishing community, some subgroups consume only one type of *Apong*, despite having similar lifestyles and dietary habits. We previously found that *Nogin Apong* had a diverse array of lactic acid bacteria and was rich in saccharides and amino acids, while *Poro Apong* was dominated by *Lactobacillus*[14]. In this study, we aimed to determine how the different compositions of microbes and metabolites in these two types of *Apong* may affect the gut microbiome in volun-

teers from the same ethnicity who have similar dietary patterns, but differ in their choice of *Apong*.

## Results:

### *Apong* drinkers have distinct levels of biochemical markers compared to non-drinkers, including differences in blood pressure, glucose, protein levels, and liver enzymes

Most individuals in this study had normal physiological and biochemical test results regardless of their lifestyles and dietary habits. However, blood pressure was significantly higher in *Poro* drinkers compared to the controls. **(Figure 1B, a-b)**. Both *Nogin* and *Poro* drinkers had lower total protein and albumin levels in their blood than non-drinkers **(Figure 1B, c-d)**, with *Nogin* drinkers having a lower albumin to globulin ratio **(Figure 1B, f)**. Both *Nogin* and *Poro* drinkers also had lower blood glucose levels than non-drinkers **(Figure 1B, g)**, which is consistent with the blood glucose-lowering properties of fermented foods and beverages[15]. Lipid levels, as measured by triglycerides, were within the normal range and comparable in all participants **(Figure 1B, h)**. *Poro* drinkers had significantly lower levels of SGPT **(Figure 1B, i)** and SGOT **(Figure 1B, j)**, which is a sign of a healthy liver[16], and lower levels of bilirubin total and bilirubin direct **(Figure 1B, k-l)**. Although high bilirubin levels can indicate liver damage, low levels are not a concern for health.

### The gut microbiota composition varies between non-drinkers and *Apong* drinkers

We compared the gut microbial composition of *Apong* drinkers and non-drinkers. We found significant differences between *Nogin* and *Poro* drinkers and drinkers and non-drinkers **(Figure**

**2A)**. Overall, the participants were dominated by *Bacteroidota* and *Firmicutes*, with smaller proportions of *Proteobacteria*. However, while *Bacteriodata* made up a significant proportion of the gut bacterial community in *Nogin* drinkers and non-drinkers, it made up a smaller proportion in *Poro* drinkers. Notably, *Actinobacteria* was only detected in non-drinkers. Next, we identified taxa at the family level that significantly differed in abundance among the *Apong* drinkers and non-drinkers **(Figure 2B)**. *Apong* drinkers had higher colonization by Enterobacteriaceae and Ruminococcaceae compared to non-drinkers. On the other hand, *Poro* drinkers had significantly less Prevotellaceae than *Nogin* drinkers and non-drinkers.

We also identified microbes that were specific to *Apong* drinkers **(Figure 2C)**. To do this, we identified and compared ASVs that were detected in at least 50% of all participants at a minimum abundance of 0.1% in *Apong* drinkers and non-drinkers. Most of the ASVs were shared (n=52) among all participants and were part of the "core" microbiome. However, certain ASVs were only detected in *Nogin* (n=9) or *Poro* drinkers (n=6). Four of the *Poro*-specific ASVs were from the Lachnospiraceae family, while the others were from the Succinivibrionaceae family and an unknown family of Bacteroidales. The majority of the *Nogin*-specific ASVs were from the Prevotellaceae family (n=7), with the remaining ASVs coming from the *Ruminococcaceae* and *Enterobacteriaceae* families. The only ASV that was detected in non-drinkers but not in *Apong* drinkers was from the *Prevotellaceae* family.

## Gut microbial diversity in *Apong* drinkers is higher than in non-drinkers, and high-frequency Nogin drinkers have lower gut microbial diversity than other Nogin drinkers

We estimated the gut microbial diversity in *Apong* drinkers and non-drinkers using three different diversity metrics **(Figure 3A)**. Gut microbial diversity was significantly higher in the *Apong* drinkers than the non-drinkers.

We divided the participants into three categories based on their *Apong* consumption frequency: high (HD), medium (MD), and low (LD). *Poro* consumption frequency did not have an effect on microbial diversity **(Supplementary Figure 1)**. However, among Nogin drinkers, high-frequency drinkers had significantly lower gut microbial diversity.

### *Apong* consumption and frequency has a significant effect in gut microbial composition

We investigated microbial composition between participants by computing Bray-Curtis and weighted UniFrac beta-diversity **(Figure 3)**. The distance matrices revealed small but significant differences among the non-drinkers, *Nogin* drinkers, and *Poro* drinkers, as well as among sub-groups of *Apong* drinkers based on frequency of consumption **(Figure 3A)**.

Although location also had a significant effect on microbial composition, *Apong* usage had a larger overall effect. **(Figure 3A)**. When blocking "location" in the perMANOVA, drink type (non-drinker vs. *Nogin*, or *Poro* drinker) still explains a significant portion of the variation in microbial composition (Bray-Curtis: $R^2$= 0.034; p= 0.003 & Unifrac: $R^2$= 0.035; p= 0.001). Despite inhabiting the same location (Majuli), the gut microbial composition of non-drinkers (n=24) and *Nogin* drinkers (n=18) formed two different clusters of PCoAs ($R^2$= 4.79; p=0.005) **(Supplementary Figure 2)**.

Notably, blood serum biochemistry markers, such as total protein, albumin and globulin levels, and blood glucose and pressure, explained significant variation in the presence or absence of certain microbes, but not in the phylogenetic diversity of gut microbial composition.

## *Nogin* drinkers have a more homogenous gut microbial community than *Poro* drinkers and non-drinkers

The gut microbial community of *Nogin* drinkers clustered together in the principal coordinate analysis (PCoA) based on both weighted UniFrac and Bray-Curtis distances **(Figure 3C)**. However, non-drinkers and *Poro* drinkers had more heterogeneous microbial communities than *Nogin* drinkers. This was also demonstrated by the higher distance to centroid (Multivariate homogeneity of groups dispersions) **(Figure 3D)**, where *Nogin* drinkers had the lowest distance to centroid, so highest homogeneity whereas *Poro* drinkers had a high distance, so more heterogeneity in the weighted UniFrac distance (considers phylogenetic relatedness).

## Fecal microbial metabolite profiles are different between the *Apong* drinkers and non-drinkers

To study the metabolic activity in the gut ecosystem of the cohorts, an untargeted metabolite profiling with GC-MS analysis was performed. Metabolites of microbial origins were identified using the human metabolome database (HMDB) for subsequent analysis. We extracted a total of 384 metabolites which comprises mainly amino acids, bile acids, fatty acids, indoles, and saccharides. *Apong* consumption led to depletion of certain metabolites, such as acetamid, benzestrol,

butanedioic acid, and cyclopropanecarboxylic acid, while *Poro* drinkers had higher levels of undecanoic acid than non-drinkers and *Nogin* drinkers **(Figure 4)**.

We agglomerated gut microbes at family level and correlated these families to fecal metabolites using the mbOmic package in R[17]. We found 113 significant correlations (adjusted p-values < 0.05) between microbial families and fecal metabolites. Only 23 of these correlations had a rho value higher than 0.70. Top ten correlations are listed in **Figure 5A**, and the majority of these correlations were with Clostridia or unknown Bacteroidota. All the correlations are listed in **Supplementary Table 1**.

Although having relatively heterogeneous gut microbial composition among participants **(Figure 3C-D)**, *Poro* drinkers had a highly uniform composition of fecal metabolites **(Figure 5B)**. Drink type (non-drinker, *Nogin* drinker, or *Poro* drinker) had even a larger effect on metabolite composition than on gut microbiome composition **(Figure 5C)**.

Taken together, these results show that *Apong* drinkers had distinct blood serum markers, gut microbial composition, and fecal metabolites compared to non-drinkers.

### *Apong* consumers had lower levels of iso-valeric acid among their fecal SCFAs compared to non-drinkers

The gut microbiota helps to break down undigested food through fermentation, producing short-chain fatty acids (SCFAs) as a result[18]. The SCFAs are important for gut homeostasis and health and their levels are affected by diet. Although BCFAs (branched-chain fatty acids) may be im-

164 portant in the gut and could potentially serve as markers of gut microbial metabolism, they have
165 received less attention than the major SCFAs[19].

166 We measured the levels of three major SCFAs and one BCFA (acetic acid, butyric acid, propion-
167 ic acid, and iso-valeric acid) in fecal samples from both *Apong* drinkers and non-drinkers using
168 HPLC. Propionic acid was the most abundant SCFA in both groups, while butyric acid was the
169 least **(Figure 6)**. Some volunteers had very high levels of acetic acid, but it was not correlated
170 with any other data **(Supplementary Figure 3)**. Isovaleric acid, a BCFA and considered to be
171 harmful to the colon epithelium[20], was significantly lower in *Apong* drinkers compared to non-
172 drinkers, but there was no significant difference for the other SCFAs **(Figure 6)**.

173

174 **Discussion**

175 The composition of the gut microbiome is influenced by various factors, including diet and life-
176 style[4, 21-23], but the impact of a single component of diet within a population of the same ethnicity
177 on the gut microbiome has not been studied before. This study investigates the effect of two
178 types of a traditional, rice-based fermented alcoholic beverage (*Apong*) on the gut microbiome
179 and health of the *Mishing* community in India. All volunteers in the study were of the same eth-
180 nicity, healthy, and had similar dietary habits, except for being an *Apong* drinker or not.

181 Our study showed that the composition of the gut microbiome is affected by the choice of
182 *Apong*. Despite the frequent consumption of an alcoholic beverage, the volunteers in the study
183 had normal BMI and healthy vital organ function. Previous research has shown that both varie-
184 ties of *Apong* contain mild alcohol (~10%) and have a high content of phenolics[14].

185 This study found that *Apong* consumers had a more diverse gut microbiome compared to the
186 non-drinkers, which had a stable community with fewer ASVs. This suggests that a single die-
187 tary factor alone can significantly impact the gut microbiome of a community, consistent with
188 previous research on the gut microbiome of children in Burkina Faso[24]. Our study also observed
189 a strong association between variation in microbial composition and blood glucose levels and
190 blood pressure, which has not been reported in previous studies. Further research is needed to
191 understand the causal relationship between these factors and gut bacterial diversity, which may
192 allow for the development of microbiome-based biomarkers for predicting lifestyle diseases.

193 In this study, the gut microbiome of the *Mishing* community was found to be dominated by the
194 *Prevotellaceae* family, a signature of the Indian population[25-30], which has been previously asso-
195 ciated with a vegetarian or carbohydrate-rich diet[7]. However, *Poro* drinkers had lower levels of
196 *Prevotellaceae* than non-drinkers and *Nogin* drinkers, even though their gut microbiomes were
197 colonized to a high extent by *Prevotellaceae*.

198 Lastly, we found that the gut microbiome of the *Mishing* population was colonized to a high ex-
199 tent by *Succinivibrio*, a bacterium not previously reported in the Indian population[25-27]. This bac-
200 terium is commonly found among hunter-gatherers and foragers[31, 32]. The presence of
201 *Succinivibrio* in the microbiome was previously reported in the *Hadza* hunter gatherers and tradi-
202 tional Peruvian populations[31,33]. We speculate that the high abundance of *Succinivibrio* in the
203 *Mishing* population may be due to co-habitation with domesticated animals. The co-presence of
204 butyrate producers and other essential gut bacteria along with expected levels of SCFAs but low-
205 er levels of BCFA and blood serum measurements suggests that *Apong* does not have a detri-
206 mental effect on the structure and function of the gut microbiome.

**Conclusion:** In conclusion, this study found that the choice of *Apong*, a traditional rice-based fermented alcoholic beverage, significantly impacts the gut microbiome composition and blood serum markers in the *Mishing* community in India. The gut microbiomes of *Apong* consumers were more diverse than those of non-drinkers, and *Poro* drinkers had lower levels of Prevotellaceae than non-drinkers and *Nogin* drinkers. The gut microbiome of the *Mishing* community was also colonized by *Succinivibrio*, a bacterium not previously reported in the Indian population. The differences in gut metabolites between *Apong* drinkers and non-drinkers were even greater than those in the gut microbiome. These findings suggest that a single dietary factor can significantly impact the gut of a population and highlight the need for further research on the causal relationship between these factors and gut bacterial diversity for the development of microbiome-based biomarkers for predicting lifestyle diseases.

## Materials and methods

### Materials

All the reagents, media and chemicals used in this study were of analytical grade. RNA-later solution (Cat. No RM49049), short chain fatty acids standards viz. Acetic acid (Cat. No. 5438080100), Butyric acid (Cat. No 19215), Isovaleric acid (Cat. No 78651), and Propionic acid (Cat No. 94425) were procured from Sigma-Aldrich, USA. QiAmp DNA stool mini kit (Cat. No: 19590) was procured from Qiagen Inc, Germany. Blood collection vials K3-EDTA vial (Cat. No 368860) and clot vials (Cat. No 368975) were procured from BD Diagnostics, Oxford, UK.

Blood serum biomarkers kits were procured from CCS ® Coral Clinical Systems, Tulip Diagnostics (P) Ltd., Goa, India.

**Ethics statement**

This study was approved by the Ethics Committee of the Institute of Advanced Study in Science Technology (IASST), Guwahati, India (Approval number: IEC(HS)/IASST/1082/2014-15/6) and was conducted following the guidelines and regulations. Written informed consents were taken from the volunteers along with some standard questionnaires.

**Study sites, volunteers, and sample collection**

In this study, potential volunteers were identified through an electoral database and surveyed in locations dominated by the *Mishing* community. The control group consisted of individuals who do not consume *Apong* but follow the same dietary pattern, while the experimental group included 166 individuals from the *Mishing* community and 24 individuals from a Vaishnavite satra with slightly different dietary habits. A bi-lingual survey questionnaire was used to collect information on dietary patterns, age, sex, medical history, family lineage, and demographics. Inclusion criteria were age (18-50 years) and ethnicity, while exclusion criteria were the use of antibiotics, health supplements and other drugs, consumption of any other liquor except *Apong*, and medical history. The *Mishing* population was stratified based on the amount of rice beer consumed per week, with non-drinkers classified as those who did not consume *Apong*, low-drinkers as those who consumed less than 250 ml, medium-drinkers as those who consumed 250-500 ml, and high-drinkers as those who consumed more than 500 ml per day. The control group included monks from *Vaishnavite* satras, who avoided alcoholic beverages and red meat but consumed fish as part of their regular diet.

Fecal and blood samples were collected from the Mishing population living in the Telam, Dhemaji, and Majuli districts of Assam. A customized kit was used to collect fecal samples, with one container, including 2 ml of RNA later solution, used for DNA extraction and another for metabolomics studies. Three ml of blood was withdrawn from individuals by a phlebotomist. Blood was collected in a K3-EDTA vial and clot vials to separate plasma and serum, respectively. All the samples were immediately frozen after collection and were transported to the laboratory in frozen condition. Fecal and blood samples were stored at -80 ℃, until processed. In addition to the samples and questionnaires, anthropometric data such as height, weight, BMI, and blood group were also collected from the volunteers.

**Analysis of biochemical parameters of plasma and serum**

The plasma samples were analyzed using standard biochemical assay kits (CCS ® Coral Clinical Systems, Tulip Diagnostics (P) Ltd., Goa, India), following the manufacturer's instructions. Serum glucose, HDL cholesterol, albumin, globulin, total protein, and liver function tests, for example, serum glutamic oxaloacetic transaminase (SGOT), serum glutamic pyruvic transaminase (SGPT), and alkaline phosphatase (ALP), direct and total bilirubin contents were determined.

**DNA extraction from faecal sample**

DNA extraction was performed within 72 h of faecal sample collection. QiAmp DNA stool mini kit (Cat. No: 19590 Qiagen Inc, Germany) was used for the extraction of metagenomic DNA from the faecal samples following the manufacturer's instruction. Briefly, a 400 µl of faecal sample was mixed to 1400 µl of ASL buffer (provided with the kit) and incubated at 90 ℃ for 10

min. The supernatant was collected after brief centrifugation at 13,000 rpm for 2 min, to which Inhibitex tablet and ProteinaseK (supplied with the kit) were added. After a short incubation of 10 min at 70 ℃, the mixture was centrifuged at 13,000 rpm for 3 min. The supernatant was collected in a filter with a silica column. The column was washed twice with a wash buffer (provided with the kit) and the bound DNA was eluted with a preheated elution buffer (supplied with the kit). The amount of dsDNA was quantified using a dsDNA estimation kit with a Fluorometer (Quantiflour, Promega, Madison, USA).

**Library preparation, 16S rDNA amplicon sequencing and analyses**

The V3-V4 region of the 16S rDNA was amplified using the primer pairs 341F & 805R[34~~41~~]. The indexing and library preparation of the amplified DNA fragments were carried out using Nextera XT library preparation and indexing kits according to the Illumina MiSeq protocol[35]. DNA fragments were multiplexed and subjected to 2 X 300 bp paired-end sequencing in an Illumina MiSeq machine with the sequencing service provider (Macrogen Inc., Seoul, Republic of Korea).

**Bioinformatics analyses of the amplicon dataset**

The paired-end reads generated from Illumina sequencing were processed using the LotuS2 pipeline[36]. Reads having less than 170 bases in length were filtered out from the analysis. In LotuS2, DADA2 algorithm[37] was used to cluster sequences into amplicon sequence variants (ASVs). Using the options for LULU (-lulu)[38~~45~~] and UNCROSS2 (-xtalk), sequence clusters were curated and refined. ASVs were aligned with Lambda[39] to SILVA 138.1[40], to obtain taxonomic assignments for ASVs using the LotuS2 LCA algorithm. Otherwise, default options in LotuS2 were

used. As a result, 28,083,099 total reads were clustered into 7550 ASVs in the final matrix, summing to 14,893,374 reads after quality filtering and removal of contaminants[41].

The processed samples were further analysed with phyloseq package[42] in R (version 3.6.1). Samples were rarefied to an even depth using the rtk tool[43] for diversity analysis. For estimation and calculation of diversity indices, vegan[44] package was used.

**Metabolomics analyses of fecal samples**

The following techniques were used for the analyses of fecal samples.

**Gas chromatography-mass spectroscopy (GC-MS) analysis**

Untargeted fecal metabolite profiles were determined by GC-MS analysis. A 40 mg of lyophilized fecal sample was extracted with 1 ml of HPLC grade methanol (Merck, Mumbai, India) and kept in a shaker overnight at 1200 rpm. The sample was centrifuged at 10000 rpm for 10 min at 10°C. The extract was then dried at room temperature using a vacuum desiccator and re-suspended in a mixture of 40 µl of pyridine and 20 mg/ml of methoxyamine hydrochloride. After a brief vortex, the solution was incubated at 30 °C for 90 min and then derivatized with 20 µl of N- methyl-N-trimethylsilyltrifluoroacetamide with 1% trimethylchlorosilane (Merck, USA) at 70 °C for 30 min[54]. The sample was then centrifuged at 3000 rpm for 5 min and used for GC-MS analysis.

Samples were run in a Shimadzu GC 2010Plus-triple quadrupole (TP-8030) system fitted with an EB-5MS column (length: 30 m, thickness: 0.25 µm, ID: 0.25 mm). A 1 µl of the sample was injected in split less mode at 300°C using helium as carrier gas at a 1 ml/min flow rate. The oven program was set at 70°C initially, ramped at 1°C/min for 5 min up to 75°C. Subsequently, it was

311 increased at 10°C/min up to 150 °C, and held for 5 min, followed by increasing at the same rate
312 up to 300 °C and held for 5 min. The mass spectrometer was operated at a continuous scan from
313 45 to 600 m/z in the electron ionization (EI) mode at 70ev with 200 °C as the source tempera-
314 ture. Peak identification was performed using the National Institute of Standards and Technology
315 library, USA, by matching the mass spectra.

316 **Quantification of faecal short-chain fatty acids (SCFAs) by RP-HPLC analysis**

317 Faecal samples were lyophilized, and 100 mg of each sample was dissolved in a solvent prepared
318 using acetonitrile (Cat. No 271004 Merck Millipore, Germany) and 10-mM $KH_2PO_4$ (pH 2.4)
319 (Cat. No P5379, Sigma Aldrich, USA) in 1:1 ratio. The solutions were then vortexed vigorously
320 for 5 min. ensuring proper mixture, centrifuged at 4000 rpm for 5 min. The supernatant was col-
321 lected and filtered through a 0.22 µm syringe filter for downstream analysis.

322 Quantification of the organic acids was carried out in an analytical HPLC instrument (Waters,
323 USA) with 5 µm ODS2 (4.6 x 250 mm, Waters SPHERISORB) reversed- phase C18 analytical
324 column. Two HPLC grade solvents (solvent A was 10-mM $KH_2PO_4$, pH 2.4 with phosphoric ac-
325 id, while solvent B was 100% acetonitrile) were used in a gradient system with a flow rate of 1.5
326 ml/min. The absorbance of the eluted compound was monitored at 210 nm by the PDA detector.

327 **Statistical analyses:**

328 All the statistical tests were performed in the R platform by using base functions and calling spe-
329 cialized packages such as phyloseq[42], vegan[44], microeco[45], microbiome[46], microbiomeutilities[10],
330 mbOmic[17], rstatitix[51]. Comparisons among the anthropometric measures and serum biochemical
331 markers were carried out using the Kruskal-Wallis test. To compare the microbial diversity

among *Apong* drinkers and non-drinkers, the samples were first rarefied to an equal depth. Then, the alpha diversity indices were calculated for each sample, including the number of unique features (richness), the Shannon diversity, and Chao1 metric. To determine if there were significant differences in microbial diversity among *Apong* drinkers and non-drinkers, a Kruskal-Wallis test was performed. Before calculating the beta distance (Bray-Curtis and weighted UniFrac), the table was normalized to relative abundances. “adonis” and “mantel” functions in the vegan package were used to run PERMANOVA (Permutational Multivariate Analysis of Variance) and Mantel tests to calculate metadata variables explaining variation based on beta diversity distances. “betadisper” function in the vegan package was used to estimate the homogeneity of drink types groups in the PCoAs. The "corr" function from the mbOmic package was used to calculate Pearson correlations between the metabolomics data and microbial taxa at the family level and adjusted p values for multiple comparisons were used.

**Acknowledgement:** The authors acknowledge the Institutional Level Biotech Hub (DBT, Govt. of India) and Central instrumentation facility of IASST for providing the facilities. This research was funded by Department of Biotechnology (DBT) under Unit of Excellence Project (BT/550/NE/U-Excel/2014) and SC/ST community development programme in IASST (SEED/TITE/2019/103) funded by DST, Govt. of India. SD is thankful to IASST for providing financial support to carry out research work. The authors also acknowledge Mr. Pinku Rajbongshi for assisting in collection of blood samples. We are thankful to Mr. Dhrubajyoti Regon, Mr. James Doley, Mr. Deepak Mili and Mr. Samujjal Saikia for assistance in recruitment of volunteers and sample collection. EO, FH were supported by European Research Council H2020 StG (erc-stg-948219, EPYC). CF, RA, EO, FH were supported by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme Gut Microbes

and Health BB/r012490/1 and its constituent project BBS/e/F/000Pr10355, Core Capability Grant BB/CCG1720/1 and the work delivered via the Scientific Computing group, as well as support for the physical HPC infrastructure and data centre delivered via the NBI Computing infrastructure for Science (CiS) group.

**Availability of data:** Sequencing data are available on the NCBI SRA server under the BioProject ID: PRJNA906264

**REFERENCES:**

1. Cani PD. Human gut microbiome: hopes, threats and promises. Gut 2018; 67:1716-25.
2. Heintz-Buschart A, Wilmes P. Human Gut Microbiome: Function Matters. Trends in microbiology 2018; 26:563-74.
3. Hildebrand F, Gossmann TI, Frioux C, Özkurt E, Myers PN, Ferretti P, et al. Dispersal strategies shape persistence and evolution of human gut bacteria. Cell host & microbe 2021; 29:1167-76.e9.
4. Spor A, Koren O, Ley R. Unravelling the effects of the environment and host genotype on the gut microbiome. Nature reviews Microbiology 2011; 9:279-90.
5. Xu Z, Knight R. Dietary effects on human gut microbiome diversity. The British journal of nutrition 2015; 113 Suppl:S1-5.
6. Hasan N, Yang H. Factors affecting the composition of the gut microbiota, and its modulation. PeerJ 2019; 7:e7502.

7. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature 2014; 505:559-63.

8. Tanes C, Bittinger K, Gao Y, Friedman ES, Nessel L, Paladhi UR, et al. Role of dietary fiber in the recovery of the human gut microbiome and its metabolome. Cell host & microbe 2021; 29:394-407.e5.

9. Melini F, Melini V, Luziatelli F, Ficca AG, Ruzzi M. Health-Promoting Components in Fermented Foods: An Up-to-Date Systematic Review. Nutrients 2019; 11.

10. Taylor BC, Lejzerowicz F, Poirel M, Shaffer JP, Jiang L, Aksenov A, et al. Consumption of Fermented Foods Is Associated with Systematic Differences in the Gut Microbiome and Metabolome. mSystems 2020; 5.

11. Lee JE, Ha JS, Park HY, Lee E. Alteration of gut microbiota composition by short-term low-dose alcohol intake is restored by fermented rice liquor in mice. Food research international (Ottawa, Ont) 2020; 128:108800.

12. Le Roy CI, Kurilshikov A, Leeming ER, Visconti A, Bowyer RCE, Menni C, et al. Yoghurt consumption is associated with changes in the composition of the human gut microbiome and metabolome. BMC Microbiol 2022; 22:39.

13. Green R, Milner J, Joy EJ, Agrawal S, Dangour AD. Dietary patterns in India: a systematic review. The British journal of nutrition 2016; 116:142-8.

14. Das S, Deb D, Adak A, Khan MR. Exploring the microbiota and metabolites of traditional rice beer varieties of Assam and their functionalities. 3 Biotech 2019; 9:174.

15. Kok CR, Hutkins R. Yogurt and other fermented foods as sources of health-promoting bacteria. Nutrition reviews 2018; 76:4-15.

16. Cohen JA, Kaplan MM. The SGOT/SGPT ratio--an indicator of alcoholic liver disease. Digestive diseases and sciences 1979; 24:835-8.

17. McHardy IH, Goudarzi M, Tong M, Ruegger PM, Schwager E, Weger JR, et al. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. Microbiome 2013; 1:17.

18. van der Beek CM, Dejong CHC, Troost FJ, Masclee AAM, Lenaerts K. Role of short-chain fatty acids in colonic inflammation, carcinogenesis, and mucosal protection and healing. Nutrition reviews 2017; 75:286-305.

19. Rios-Covian D, González S, Nogacka AM, Arboleya S, Salazar N, Gueimonde M, et al. An Overview on Fecal Branched Short-Chain Fatty Acids Along Human Life and as Related With Body Mass Index: Associated Dietary and Anthropometric Factors. 2020; 11.

20. Vieira ADS, de Souza CB, Padilha M, Zoetendal EG, Smidt H, Saad SMI, et al. Impact of a fermented soy beverage supplemented with acerola by-product on the gut microbiota from lean and obese subjects using an in vitro model of the human colon. Appl Microbiol Biotechnol 2021; 105:3771-85.

21. Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, et al. Factors influencing the composition of the intestinal microbiota in early infancy. Pediatrics 2006; 118:511-21.

22. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. Succession of microbial consortia in the developing infant gut microbiome. Proceedings of the National Academy of Sciences of the United States of America 2011; 108 Suppl 1:4578-85.

23. O'Hara AM, Shanahan F. The gut flora as a forgotten organ. EMBO reports 2006; 7:688-93.

422 24. De Filippo C, Di Paola M, Ramazzotti M, Albanese D, Pieraccini G, Banci E, et al. Diet,
423 Environments, and Gut Microbiota. A Preliminary Investigation in Children Living in Rural and
424 Urban Burkina Faso and Italy. Front Microbiol 2017; 8:1979.
425 25. Dehingia M, Devi KT, Talukdar NC, Talukdar R, Reddy N, Mande SS, et al. Gut bacteri-
426 al diversity of the tribes of India and comparison with the worldwide data. Scientific reports
427 2015; 5:18563.
428 26. Bhute S, Pande P, Shetty SA, Shelar R, Mane S, Kumbhare SV, et al. Molecular Charac-
429 terization and Meta-Analysis of Gut Microbial Communities Illustrate Enrichment of Prevotella
430 and Megasphaera in Indian Subjects. Front Microbiol 2016; 7:660.
431 27. Dhakan DB, Maji A, Sharma AK, Saxena R, Pulikkan J, Grace T, et al. The unique com-
432 position of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered
433 using multi-omics approaches. GigaScience 2019; 8.
434 28. Maji A, Misra R, Dhakan DB, Gupta V, Mahato NK, Saxena R, et al. Gut microbiome
435 contributes to impairment of immunity in pulmonary tuberculosis patients by alteration of butyr-
436 ate and propionate producers. Environmental microbiology 2018; 20:402-19.
437 29. Das B, Ghosh TS, Kedia S, Rampal R, Saxena S, Bag S, et al. Analysis of the Gut
438 Microbiome of Rural and Urban Healthy Indians Living in Sea Level and High Altitude Areas.
439 Scientific reports 2018; 8:10104.
440 30. Hildebrand F, Moitinho-Silva L, Blasche S, Jahn MT, Gossmann TI, Huerta-Cepas J, et
441 al. Antibiotics-induced monodominance of a novel gut bacterial order. Gut 2019; 68:1781-90.
442 31. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, et al. Gut
443 microbiome of the Hadza hunter-gatherers. Nat Commun 2014; 5:3654.

444 32. Graf D, Di Cagno R, Fåk F, Flint HJ, Nyman M, Saarela M, et al. Contribution of diet to
445 the composition of the human gut microbiota. Microbial ecology in health and disease 2015;
446 26:26164.
447 33. Obregon-Tito AJ, Tito RY, Metcalf J, Sankaranarayanan K, Clemente JC, Ursell LK, et
448 al. Subsistence strategies in traditional societies distinguish gut microbiomes. Nat Commun
449 2015; 6:6505.
450 34. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of
451 general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-
452 based diversity studies. Nucleic acids research 2013; 41:e1.
453 35. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-
454 high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The
455 ISME Journal 2012; 6:1621-4.
456 36. Özkurt E, Fritscher J, Soranzo N, Ng DYK, Davey RP, Bahram M, et al. LotuS2: an ul-
457 trafast and highly accurate tool for amplicon sequencing analysis. Microbiome 2022; 10:176.
458 37. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2:
459 High-resolution sample inference from Illumina amplicon data. Nature methods 2016; 13:581-3.
460 38. Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, et al. Algorithm
461 for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nat
462 Commun 2017; 8:1188.
463 39. Hauswedell H, Singer J, Reinert K. Lambda: the local aligner for massive biological data.
464 Bioinformatics 2014; 30:i349-55.

40. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic acids research 2013; 41:D590-6.

41. Bedarf JR, Beraza N, Khazneh H, Özkurt E, Baker D, Borger V, et al. Much ado about nothing? Off-target amplification can lead to false-positive bacterial brain microbiome detection in healthy and Parkinson's disease individuals. Microbiome 2021; 9:75.

42. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PloS one 2013; 8:e61217.

43. Saary P, Forslund K, Bork P, Hildebrand F. RTK: efficient rarefaction analysis of large datasets. Bioinformatics 2017; 33:2594-5.

44. Dixon P. VEGAN, a package of R functions for community ecology. 2003; 14:927-30.

45. Liu C, Cui Y, Li X, Yao M. microeco: an R package for data mining in microbial community ecology. FEMS Microbiology Ecology 2020; 97.

46. Shetty SA, Hugenholtz F, Lahti L, Smidt H, de Vos WM. Intestinal microbiome landscaping: insight in community assemblage and implications for microbial modulation strategies. FEMS Microbiol Rev 2017; 41:182-99.

47. Sudarshan A. Shetty, & Leo Lahti. (2020, October). microbiomeutilities: Utilities for Microbiome Analytics.

**Figures legends:**

**Figure 1. A)** Summary of the cohort. **B)** A comparison of anthropological factors and serum biochemical markers among *Apong* drinkers and non-drinkers (a) Systolic Blood pressure, (b) Diastolic Blood pressure, (c) Total protein (d) Albumin, (e) Globulin, (f) Albumin/Globulin ratio, (g) Blood Glucose, (h) Triglycerides, (i) serum glutamic-pyruvic transaminase SGPT, (j) Serum glutamic-oxaloacetic transaminase (SGOT), (k) Bilirubin total and, (l) Bilirubin direct. The probability of significance is denoted by *'s, where **** depicts significance level of 0.0001, *** depicts 0.001, ** depicts 0.01, * depicts 0.05. Only significant p-values are indicated.

**Figure 2**. **A)** Relative abundance of microbial taxa at phylum level in *Apong* drinkers and non-drinkers. **B)** Top ten microbial families that are differentially abundant among *Apong* drinkers and non-drinkers. Only significant p-values are indicated. **C)** Venn diagram of microbial features (ASVs) shared between different drink type groups or special to each group.

**Figure 3. A)** Microbial diversity of *Apong* drinkers and non-drinkers as estimated by ASV richness and Shannon and Chao1 indices. **B)** The inferred variance (adjusted R2) explained by each identified covariate as determined by PERMANOVA, calculated based on weighted UniFrac and Bray–Curtis dissimilarities. Statistically significant covariates with an adjusted $p < 0.05$ using the Benjamini–Hochberg (BH) method are shown. **C)** PCoA of the weighted UniFrac and Bray-Curtis distances of the gut bacterial composition of *Apong* drinkers and non-drinkers. **D)** Dispersion (i.e. distance to centroid of the groups) of each drink type group in the PCoA plots.

**Figure 4.** Top ten fecal metabolites that are differentially abundant among *Apong* drinkers and non-drinkers. Only significant p-values are indicated.

**Figure 5. A)** Top ten highest correlations between gut microbial taxa at family level and fecal metabolites. **B)** PCoA of the Bray-Curtis distances of the fecal metabolites of *Apong* drinkers and non-drinkers. The table shows explained variance by drink type (*Nogo*, *Poro*, and non-drinker) for microbial and fecal metabolites composition.

**Figure 6.** Composition of the four short chain fatty acids (SCFAs) in the fecal samples of *Apong* drinkers and non-drinkers. Only significant p-values are indicated. "ms" denotes marginal significance.

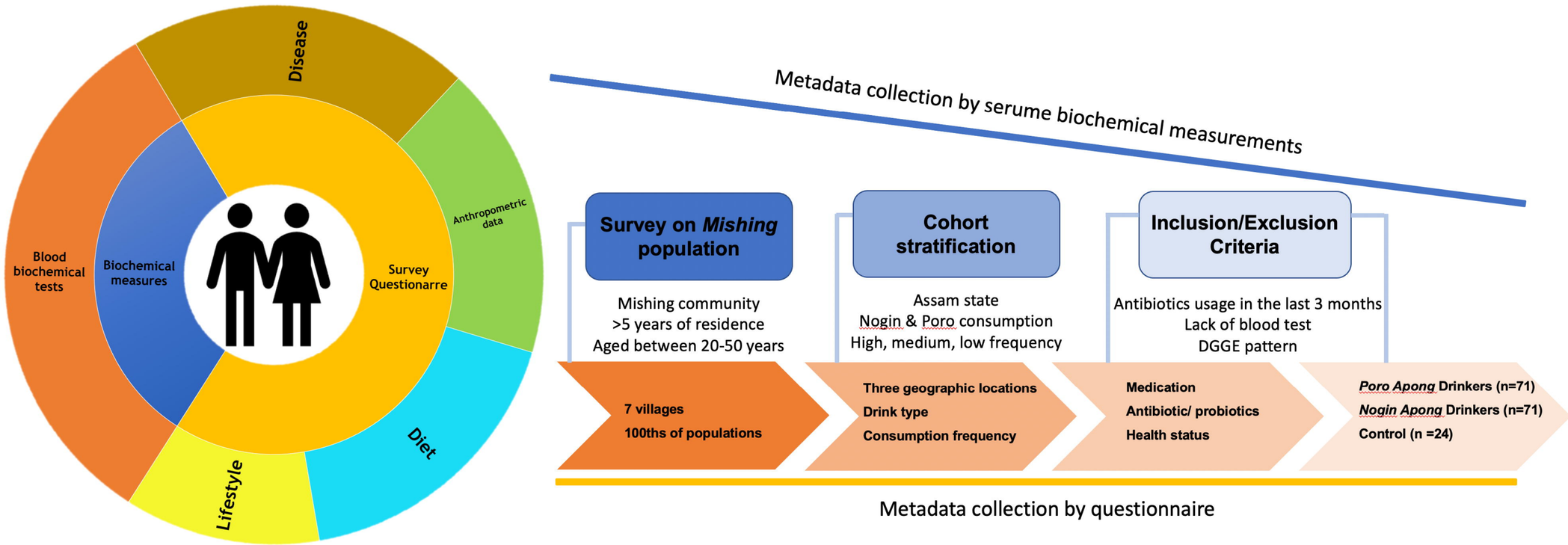
**Supplementary Figure 1:** Microbial diversity of *Apong* drinkers estimated by ASV richness: "low-drinkers" (less than 250 ml per day), "medium-drinkers" (250-500 ml per day), and "high-drinkers" (more than 500 ml per day).

**Supplementary Figure 2:** PCoA of the weighted UniFrac and Bray-Curtis distances of the gut bacterial composition of *Apong* drinkers and non-drinkers where colors depict *Nogin*, *Poro*, and non-drinkers and shapes depict different locations.

**Supplementary Figure 3:** Composition of the four short chain fatty acids (SCFAs) in the fecal samples of the participants

**Supplementary Table 1:** Correlations between gut microbial taxa at family level and fecal metabolites.

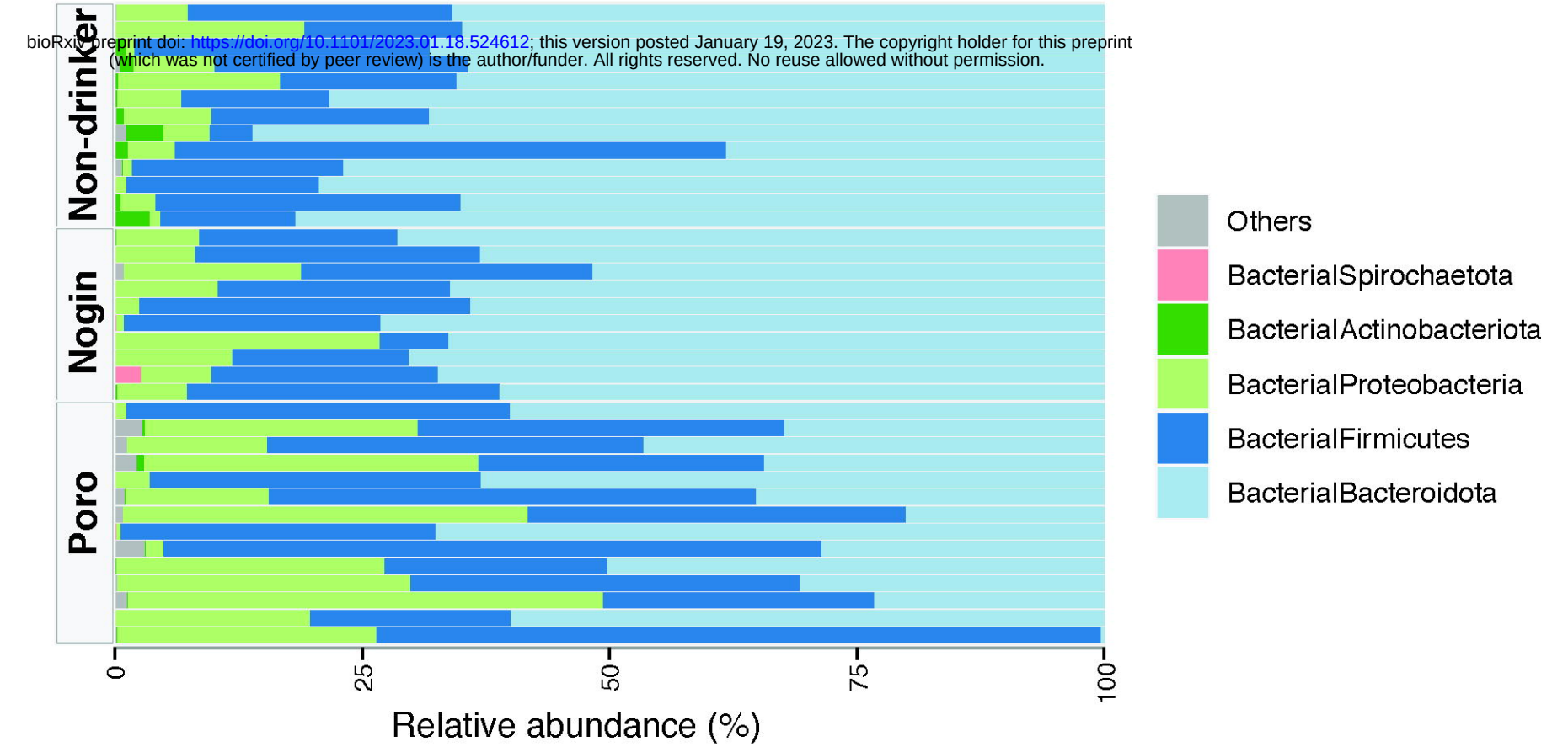
**A**



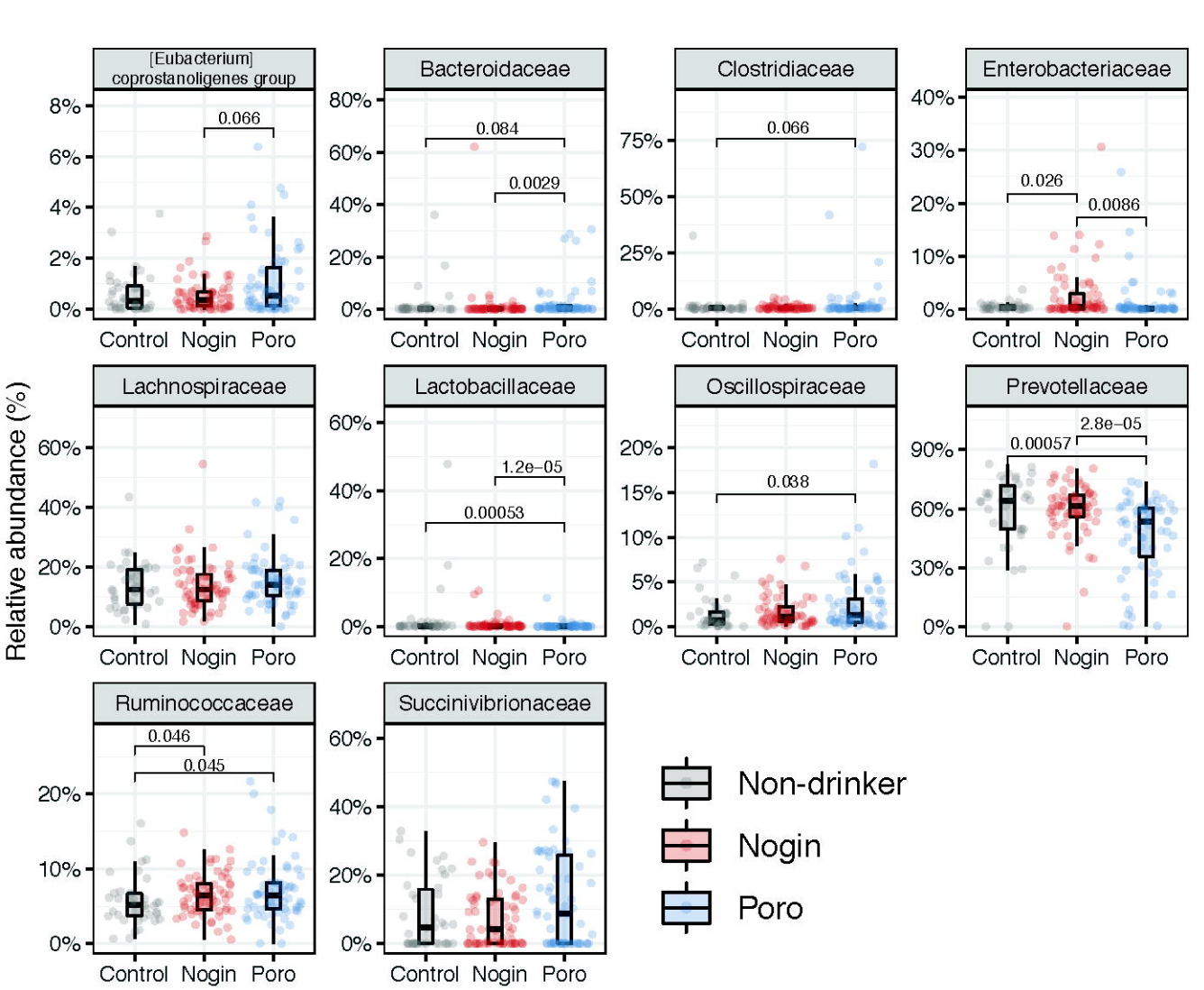

**B**

## Summary of the biochemical measurements of the participants

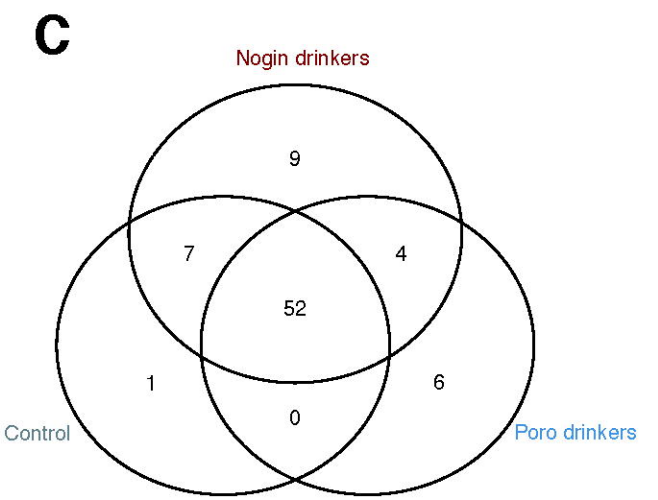
A



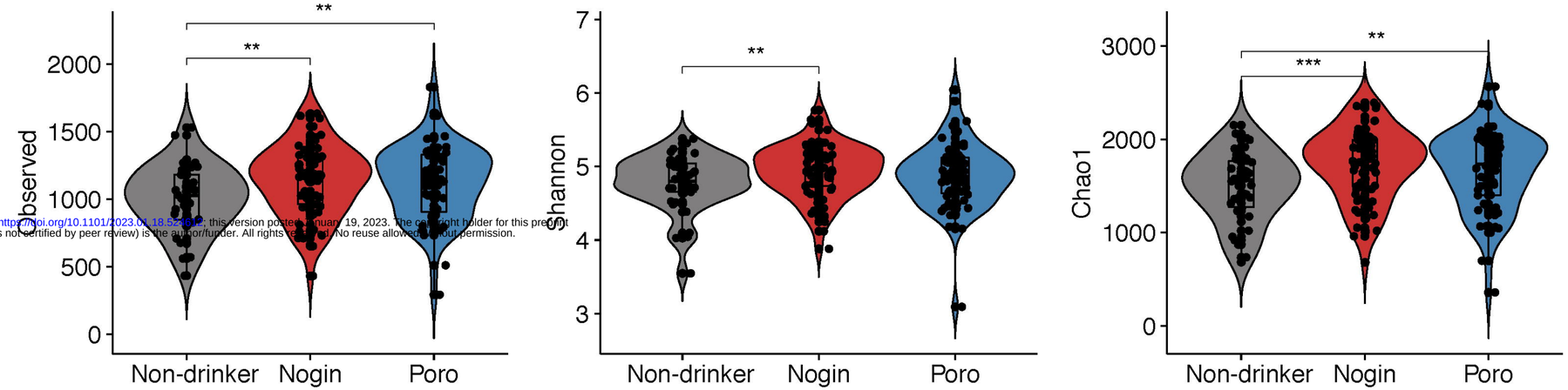

B



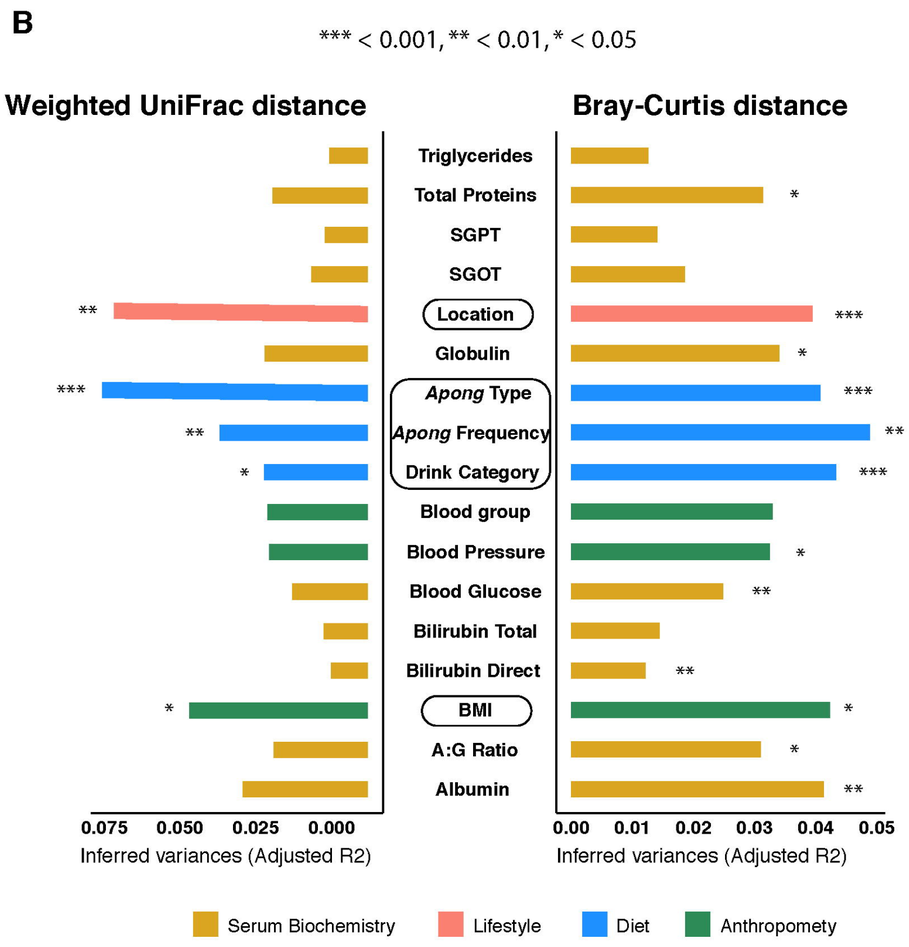

C



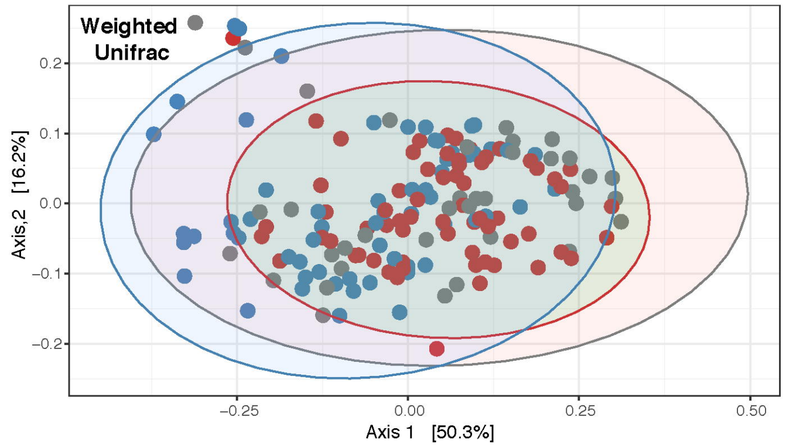
A

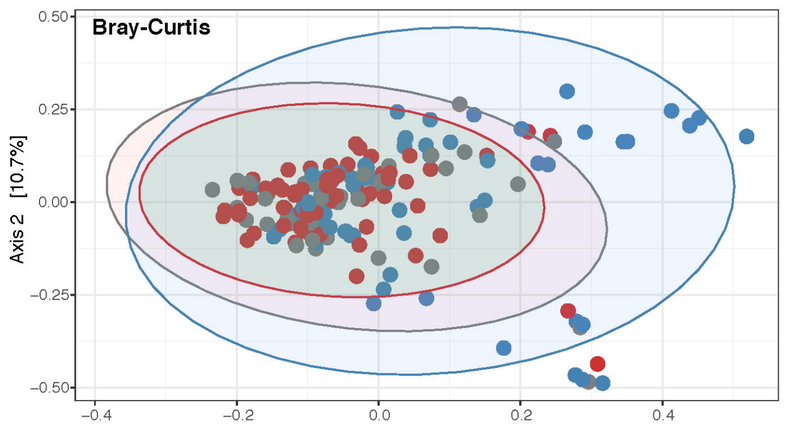



B

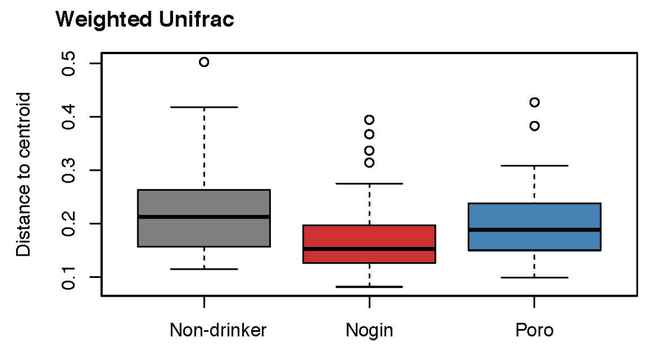


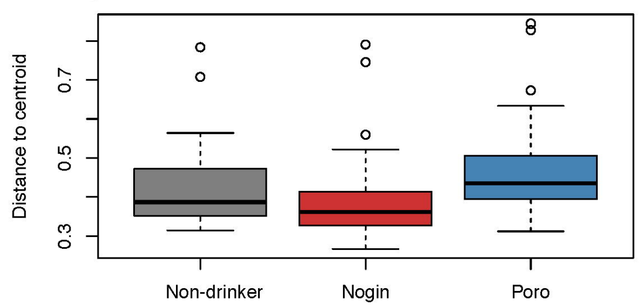

C

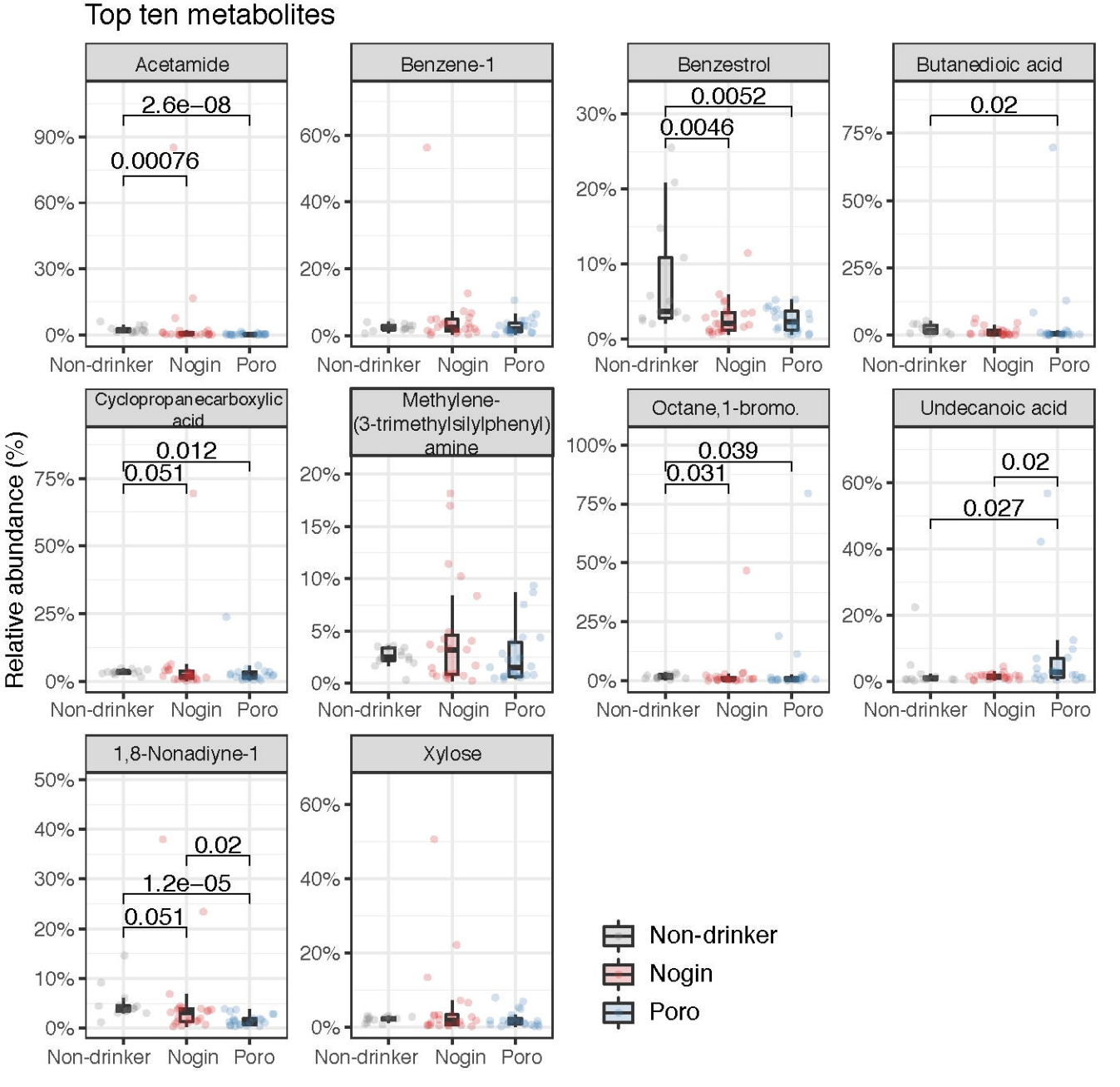



D



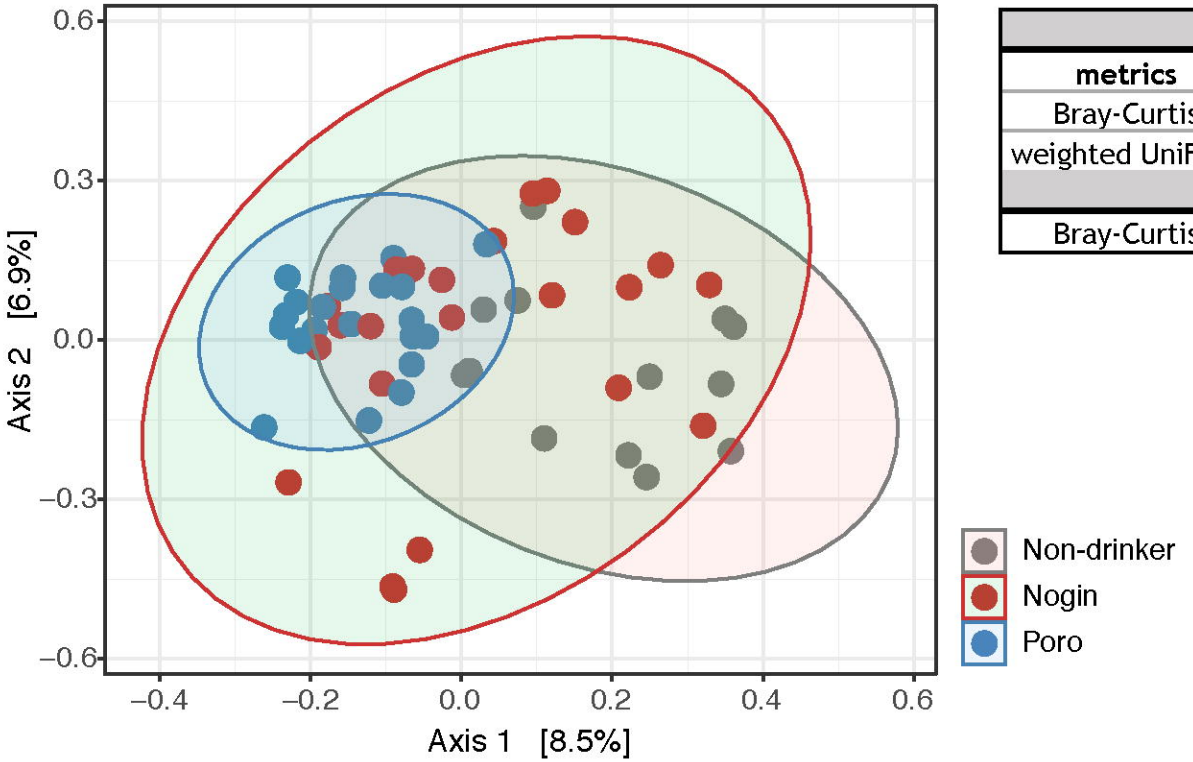


Top ten metabolites
Acetamide
2.6e−08
0.00076
Benzene-1
Benzestrol
0.0052
0.0046
Butanedioic acid
0.02
Cyclopropanecarboxylic acid
0.012
0.051
Methylene-(3-trimethylsilylphenyl) amine
Octane,1-bromo.
0.039
0.031
Undecanoic acid
0.02
0.027
1,8-Nonadiyne-1
0.02
1.2e−05
0.051
Xylose
Relative abundance (%)
Non-drinker
Nogin
Poro

**A**

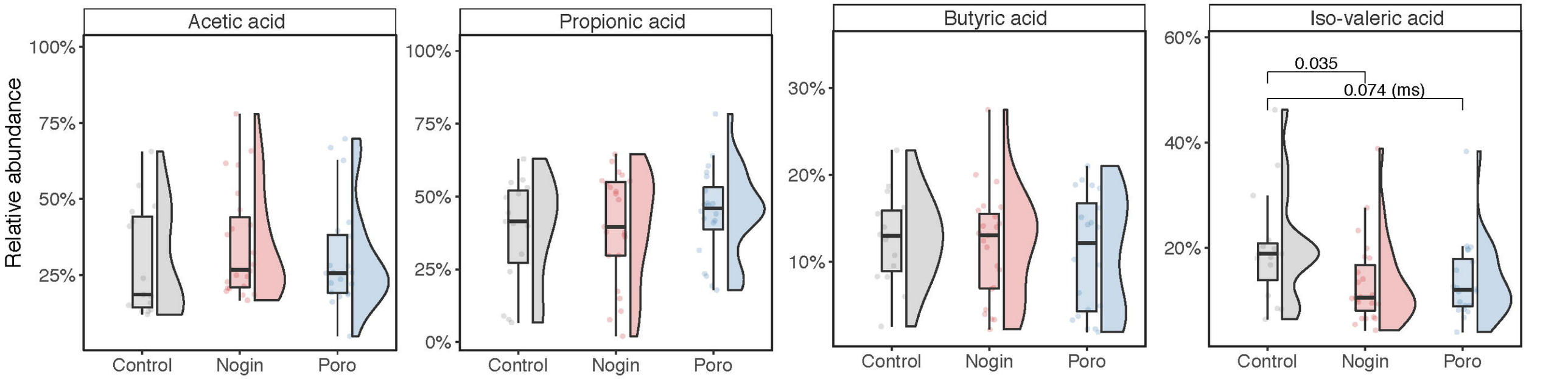
Top 10 correlating microbial taxa and metabolites

| taxa (family) | metabolite | rho | p value | adjusted p value |
|---|---|---|---|---|
| Clostridiaceae | Dimethylmalonic acid | 0.933214976 | 4.07E-26 | 1.40E-22 |
| Clostridiaceae | 1,6-Bis(2-propyn-1-yloxy)hexane | 0.932759101 | 4.88E-26 | 1.40E-22 |
| Selenomonadaceae | Decanoic acid | 0.92445005 | 1.08E-24 | 2.05E-21 |
| Clostridiaceae | Urea | 0.91609472 | 1.73E-23 | 2.47E-20 |
| Clostridiaceae | 5-(4H)-Oxazolones | 0.907582075 | 2.20E-22 | 2.51E-19 |
| Clostridiaceae | 2,4-Hexadien-1-ol | 0.881270757 | 1.51E-19 | 1.44E-16 |
| unknown Bacteroidota | Dodecanedioic acid | 0.876080983 | 4.58E-19 | 3.74E-16 |
| unknown Bacteroidota | Butanedioic acid | 0.864326188 | 4.72E-18 | 3.37E-15 |
| unknown Bacteroidota | Cyclopentane | 0.848294141 | 8.17E-17 | 5.19E-14 |
| unknown Bacteroidota | Trifluoroacetoxypentadecane | 0.832590054 | 9.87E-16 | 5.64E-13 |

**B**




| microbial composition | | |
|---|---|---|
| **metrics** | **explained variation** | **significance** |
| Bray-Curtis | 3,45 | *** |
| weighted UniFrac | 6,259 | *** |
| **metabolite composition** | | |
| Bray-Curtis | 8,66 | *** |


Acetic acid
Propionic acid
Butyric acid
Iso-valeric acid
Relative abundance
100%
75%
50%
25%
0%
30%
20%
10%
60%
40%
Control
Nogin
Poro
0.035
0.074 (ms)