
1 **Supplementary Materials**

2 **Mapping the human gut mycobiome in middle-aged and elderly adults: multi-** 3 **omics insights and implications for host metabolic health**

4 5 **This word file includes:**

6 **Methods**

7 **Fig. S1 to S12**

8 Fig. S1. Study design for the present analyses.

9 Fig. S2. The composition of gut mycobiome in different cohorts.

10 Fig. S3. Comparisons of the gut mycobiome between baseline and follow-up visit in
11 the population with repeated measures (N = 184).

12 Fig. S4. Comparison of the composition between gut fungi and bacteria.

13 Fig. S5. Associations between gut fungal and bacterial composition.

14 Fig. S6. Associations between the α -diversity indices of gut mycobiome and host
15 phenotypes.

16 Fig. S7. The number of associations within the microbiome.

17 Fig. S8. Associations between the α -diversity indices of gut bacteria and gut
18 mycobiota.

19 Fig. S9. Associations between gut mycobiome and faecal metabolome.

20 Fig. S10. Associations between gut mycobiome and metabolic traits in replication
21 cohort.

22 Fig. S11. Forest plot of the interaction of gut fungi with bacterial diversity on
23 metabolic traits.

24 Fig. S12. Mediation linkages among the gut mycobiome, faecal metabolites and
25 metabolic traits.

26 Fig. S13. Association between *Saccharomycetales spp.* and *Pichia*.

27 **Methods**

28 **Study participants and stool sample collection**

29 The current study was based on the Guangzhou Nutrition and Health Study (GNHS)
30 cohort that involved 4048 participants aged 45-70 years at recruitment between 2008-
31 2013. All participants had lived in Guangzhou city (China) for at least 5 years and
32 were followed up approximately every 3 years (Fig. S1). Detailed information on the
33 study design has been reported previously¹. The study was registered at
34 clinicaltrials.gov (NCT03179657). In the present study, 1244 participants with ITS2
35 sequencing data were included in the discovery cohort (Table S1). To validate the
36 interaction analysis, we used the control arm of a case-control study of hip fracture
37 with ITS2 and 16S rRNA sequencing data (n = 94) as an independent replication
38 cohort (Table S15). The study design of the replication cohort has previously been
39 described in detail². Stool samples of 1244 participants were collected during on-site
40 study visits between 2014-2018. Among these participants, 184 had stool samples
41 collected twice at different follow-up visits with a median follow-up time of 3.2 years.
42
43 Stool samples were collected on-site at Sun Yat-sen University, temporarily stored in
44 an ice box, manually stirred and then dispensed into centrifuge tubes and stored in a
45 refrigerator at -80°C within 4 hours. After each sample was collected, we immediately
46 checked whether the sample was sufficient and whether the collection date was
47 recorded. The stool samples were stored at -80 °C before further processing. The
48 study protocol was approved by the Ethics Committee of the School of Public Health

49 at Sun Yat-sen University and Ethics Committee of Westlake University. All

50 participants provided written informed consent.

51

52 **Phenotype data**

53 Phenotypic data assessed in the present study included demographics (e.g., age,

54 gender, BMI), 14 physiologic traits, 16 dietary intakes and 3 diseases (e.g., obesity,

55 type 2 diabetes, metabolic syndrome). Information on demographics, lifestyle and

56 physical activity was collected by questionnaires. Habitual dietary intakes at baseline

57 were estimated from a validated food frequency questionnaire (FFQ), which recorded

58 the frequencies of foods consumed in the past 12 months. Food intake was divided

59 into 16 food groups according to the Guidelines for Measuring Household and

60 Individual Dietary Diversity. Anthropometrics data were obtained by the

61 measurements of trained staff on site. Fasting venous blood samples were collected at

62 both baseline and follow-up visits. Insulin was measured using the

63 electrochemiluminescence immunoassay (ECLIA, Roche cobas 8000 e602) method.

64 The homeostatic model assessment of insulin resistance (HOMA-IR) was calculated

65 as fasting blood glucose (FBG) (mmol/L) times fasting insulin (mIU/L) divided by

66 22.5. The measurements of the other blood tests (e.g., FBG, triglycerides, total

67 cholesterol) had been described elsewhere³.

68

69 Obesity was defined as a BMI ≥ 28 kg/m². Type 2 diabetes (T2D) was defined in

70 adults as meeting one of the criteria: FBG ≥ 7.0 mmol/L, glycated hemoglobin

71 (HbA1c) \geq 6.5% or self-reported medical treatment for diabetes. Metabolic syndrome
72 was defined as meeting three of the five following items: (i) waist circumference $>$ 90
73 cm (male) or waist circumference $>$ 85 cm (female), (ii) FBG \geq 6.1 mmol/L (110
74 mg/dl) or previously diagnosed with T2D, (iii) triglycerides \geq 1.7 mmol/L (150
75 mg/dl), (iv) high-density lipoprotein (HDL) cholesterol $<$ 1.04 mmol/L (40 mg/dl) and
76 (v) systolic blood pressure (SBP) / diastolic blood pressure (DBP) \geq 130/85 mm Hg or
77 previously diagnosed with high blood pressure.

78

79 **Bioinformatic analyses**

80 *ITS2 data to detect fungal identity, diversity and abundance*

81 Microbial DNA was extracted using the E.Z.N.A.[®] soil DNA Kit (Omega Bio-tek,
82 Norcross, GA, U.S.) according to manufacturer's protocols. The final DNA
83 concentration and purification were determined by NanoDrop 2000 UV-vis
84 spectrophotometer (Thermo Scientific, Wilmington, USA), and DNA quality was
85 checked by 1% agarose gel electrophoresis. The ITS2 hypervariable regions of the
86 fungal ITS rRNA gene were amplified with primers ITS3F:
87 GCATCGATGAAGAACGCAGC and ITS4R: TCCTCCGCTTATTGATATGC by
88 thermocycler PCR system (GeneAmp 9700, ABI, USA). The PCR reactions were
89 conducted using the following program: 3 min of initial denaturation at 95 °C, 35
90 cycles of 30 s at 95 °C, 30s for annealing at 55 °C, and 45s for elongation at 72 °C,
91 and a final extension at 72 °C for 10 min. PCR reactions were performed in triplicate
92 in a 20 μ L mixture containing 4 μ L of 5 \times FastPfu Buffer, 2 μ L of 2.5 mM dNTPs, 0.8

93 μL of each primer (5 μM), 0.4 μL of FastPfu Polymerase and 10 ng of template DNA.

94 The resultant PCR products were extracted from a 2% agarose gel and further purified

95 using the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union City, CA,

96 USA) and quantified using QuantiFluorTM-ST (Promega, USA) according to the

97 manufacturer's protocol. Purified amplicons were pooled in equimolar and paired-end

98 sequenced (2×300) on an Illumina MiSeq platform (Illumina, San Diego, USA)

99 according to the standard protocols by Majorbio Bio-Pharm Technology Co. Ltd.

100 (Shanghai, China).

101

102 Raw data was demultiplexed by the MiSeq Controller Software (Illumina Inc.).

103 QIIME2 (version 2020.2) was used for the downstream analysis⁴. The demultiplexed

104 ITS2 sequences were denoised and grouped into amplicon sequence variants (ASVs;

105 i.e., 100% exact sequence match) using DADA2⁵. The ASV features that were

106 presented in only one sample were excluded for all cohorts. The individual ASVs

107 were taxonomically classified based on the UNITE (version 8.2, 99%) database using

108 the VSEARCH tool wrapped in QIIME2. α -diversity analysis was conducted through

109 the q2-diversity plugin at the sampling depth of 4500. α -diversity was estimated by

110 Shannon's diversity index (or Shannon; a quantitative measure of community richness

111 and evenness), Observed Features (or Richness; a qualitative measure of community

112 richness), Evenness (or Pielou's Evenness; a measure of community evenness), and

113 Faith's PD (or Faith's Phylogenetic Diversity; a qualitative measure of community

114 richness that incorporates phylogenetic relationships between the observed features).

115
116 To compare the gut mycobiome composition with different ethnicity, we also analysed
117 the ITS2 sequencing data from the Human Microbiome Project (HMP) cohort
118 (PRJNA356769, n = 313) and Danish cohort (PRJEB34758, n = 99). The detailed
119 information about both cohorts had been described previously^{6 7}.

120

121 ***16S rRNA data to detect bacterial identity, diversity and abundance***

122 For the 16S analysis, raw data was obtained using the same method as previously
123 described. Fastq files were demultiplexed by the MiSeq Controller Software (Illumina
124 Inc.) using a mapping file as input. Sequences were merge-paired, quality filtered and
125 analysed using QIIME2 (version 2019.10). As described above, we used DADA2
126 denoised-paired plugin in QIIME2 to process the fastq files. We filtered the features
127 that were present in only a single sample. The taxonomies of ASVs were subsequently
128 determined using the Naive Bayes classifier trained on the Sliva_132 99% reference
129 database. α -diversity analysis was conducted at the sampling depth of 1000. α -
130 diversity of the gut bacteria was estimated by the indices the same as ITS2 data.

131

132 ***Metagenome data***

133 Faecal DNA extractions were carried out by a standardized CTAB procedure. DNA
134 concentration was measured using Qubit dsDNA Assay Kit in Qubit 2.0 Fluorometer
135 (Life Technologies, CA, USA). For DNA library preparation, a total amount of 1 μ g
136 DNA per sample was used as input material. In addition, the NEBNext Ultra DNA

137 Library Prep Kit (NEB, USA) was used following manufacturer's recommendations
138 and index codes were added to attribute sequences to each sample. The DNA samples
139 were fragmented by sonication to a size of approximately 350 bp. Then, the DNA
140 fragments were end-polished, A-tailed, and ligated with the full-length adaptor for
141 Illumina sequencing with further PCR amplification. After that, PCR products were
142 purified (AMPure XP system) and libraries were analysed for size distribution by
143 Agilent2100 Bioanalyzer and quantified using real-time PCR. The clustering of the
144 index-coded samples was performed on a cBot Cluster Generation System according
145 to the manufacturer's instructions. After cluster generation, the library preparations
146 were sequenced on an Illumina HiSeq platform and 150 bp paired-end reads were
147 generated. Finally, we obtained on average 42.4 million paired-end raw reads for each
148 sample.

149

150 Next, raw sequencing reads were first quality-controlled with PRINSEQ (v0.20.4)⁸: 1)
151 to trim the reads by quality score from the 5' end and 3' end with a quality threshold
152 of 20; 2) removed read pairs when either read was < 60 bp, contained "N" bases or
153 quality score mean below 30; and 3) deduplicate the reads. Reads that could be
154 aligned to the human genome (*H. sapiens*, UCSC hg19) were removed (aligned with
155 Bowtie2 v2.2.5 using --reorder --no-contain --dovetail).

156

157 Functional profiling was performed with HUMAnN2 v2.8.1⁹, which maps sample
158 reads against the sample-specific reference database to quantify gene presence and

159 abundance in a species-stratified manner, with unmapped reads further used in a
160 translated search against Uniref90 to include taxonomically unclassified but
161 functionally distinct gene family abundances. We extracted the Uniref90 gene families
162 of gut bacteria for downstream analyses. The Uniref90 gene families were then
163 converted into KEGG Orthologs (KOs). The abundances of KOs were normalized
164 into relative abundance for each sample. In addition, the KOs presented in less than
165 10% of the 1009 samples were exclude from the downstream analysis.

166

167 **Targeted faecal metabolome profiling**

168 The absolute quantification of faecal samples (n = 913) was performed by an ultra-
169 performance liquid chromatography coupled to tandem mass spectrometry (UPLC-
170 MS/MS) system. Detailed information about the measurements has been described
171 previously¹⁰. The same list of metabolites was selected to capture the microbiota-
172 related metabolites and some key host metabolites. The metabolites mainly included
173 amino acids, bile acids, fatty acids, carbohydrates, organic acids, benzenoids,
174 pyridines, and carnitines.

175

176 **Statistical analysis**

177 All statistical analyses were performed using Stata version 15 or R version 4.0.2. To
178 compare the inter-individual variability of the human gut fungi and bacteria, we
179 applied a Mann-Whitney U test on the Bray-Curtis dissimilarity. The dissimilarity
180 matrix was calculated based on the composition at the genus level using the *vegdist()*

181 function from the *vegan* R package¹¹. We used a linear mixed model to examine the
182 association of sequencing depth and sequencing run (as a random effect) with z-score
183 normalized α -diversity indices. To examine the association between α -diversity of gut
184 fungi and bacteria, the residuals of the linear mixed model were taken as technique-
185 adjusted α -diversity indices for the Spearman correlation analysis. Procrustes analysis
186 was performed to investigate the overall relationship between gut fungal and bacterial
187 taxa, and the p value was generated based on 999 permutations. We then used the
188 *envfit* function in *vegan* to explore the contributors for the variation of the gut fungal
189 community. In addition, we excluded the fungal genera that were present in less than
190 5% of the 1244 participants. The remaining genera were identified as a group of core
191 taxa in our population. To explore the stability of the human gut mycobiome, the
192 relative abundance of the core taxa between baseline and follow-up visits were
193 compared using the paired two-sided Wilcoxon signed-rank test, among 184
194 participants with repeated measures of ITS2 data. To investigate the relationship
195 between age and gut mycobiome, we further performed a Wilcoxon rank-sum test and
196 a linear regression analysis to study the gut mycobial differences between middle-
197 aged and elderly adults. The regression model was adjusted for gender and BMI.
198
199 To investigate the determinants of the gut fungal composition, 40 factors (including
200 demographics, physiologic traits, diseases, and habitual dietary intakes) were
201 correlated to the gut fungal distance matrix (Bray-Curtis) using permutational
202 multivariate analysis of variance (PERMANOVA; permutations = 999). The linear

203 regression models were performed to examine the prospective associations between
204 the core taxa and the habitual dietary intakes (e.g., dairy, fish, cereals), adjusted for
205 age, gender, BMI, total energy intake, smoking status, drinking status, education
206 attainment, income level, physical activity, and Bristol stool scale. In these models,
207 the rank-based inverse normal transformation was applied for the relative abundance
208 of the gut fungi. The associations between the core taxa of gut fungi and technique-
209 adjusted bacterial α -diversity indices were investigated using Spearman correlation
210 analysis.

211

212 To further explore the ecological connections between gut fungi and bacteria, a 5%
213 prevalence filtering for the bacterial genera was performed. The SparCC algorithm,
214 wrapped in the FastSpar application¹², was used to estimate the cross-kingdom
215 taxonomic associations based on the raw read counts tables. FastSpar is a fast and
216 parallelizable implementation of the SparCC algorithm with an unbiased p -value
217 estimator, which was designed to provide rapid and robust correlation estimation for
218 compositional data. The analysis was conducted with default parameters and 1000
219 bootstrap samples were used to infer pseudo p values. To estimate the robustness of
220 our SparCC-inferred associations, we compared them to the associations from the
221 Spearman correlation analysis. In addition, we examined the associations between the
222 relative abundance of the gut mycobiome and bacterial KOs using the Spearman
223 correlation analyses. The significant KOs were then mapped to the PATHWAY,
224 MODULE and BRITE databases (<https://www.genome.jp/kegg/ko.html>) that stored

225 the cellular and organism-level functions.

226

227 We used the Spearman correlation analysis to examine the associations between gut
228 mycobiome and faecal metabolites. Procrustes analysis was performed to investigate
229 the overall relationship between gut mycobiome and faecal metabolome, and the p
230 value was generated based on 999 permutations. To evaluate the link between the gut
231 fungal composition and faecal metabolites, we conducted the PERMANOVA analysis
232 based on the Bray-Curtis distance matrix of the gut mycobiome (permutations = 999).
233 p values were controlled by Benjamini Hochberg method for multiple tests. FDR
234 corrected (q values) or raw p values < 0.05 were considered significant.

235

236 We applied two linear regression models to examine the associations between the gut
237 mycobiome and metabolic traits, including BMI, waist circumference, triglycerides,
238 total cholesterol (TC), HDL cholesterol, LDL cholesterol, TC/HDL, fasting glucose,
239 HbA1c, insulin and HOMA-IR: dependent variable (metabolic traits) \sim (intercept) +
240 independent variable (fungi, continuous / binary) + age + gender. The ‘continuous’
241 model used the relative abundance of fungi as independent variable. As for the
242 ‘binary’ model, we transformed the relative abundance of fungal into binary variables
243 (presence or absence). The dependent variables with skewed distribution were log-
244 transformed before analyses (fasting glucose, insulin, triglycerides, TC/HDL and
245 HOMA-IR). In addition, we used a linear regression model to examine the
246 associations of the changes in relative abundance of gut mycobiome between baseline

247 and follow-up with the parallel changes in metabolic traits, adjusted for age and
248 gender. To test the potential effect modification of the gut mycobiome, we examined
249 the interaction of gut fungi with bacterial α -diversity on the metabolic traits in the
250 discovery cohort. We further validated the results from the interaction analyses in an
251 independent replication cohort. To evaluate whether bacterial genes or faecal
252 metabolites can mediate the mycobiomic impact on the host metabolic traits, we first
253 checked whether the bacterial KOs/faecal metabolites were associated with the gut
254 mycobiome. Then we applied the bi-directional mediation analysis using the *mediate*
255 function from the R package *mediation* to infer the causal role of the gut mycobiome
256 in contributing to the human phenotype through bacterial gene function or faecal
257 metabolites¹³. Here, the host metabolic traits were the change between the baseline
258 and the follow-up visit while the gut mycobiome and bacterial gene/faecal metabolites
259 were from the baseline. We conducted the z-score normalization for the host
260 phenotype before the mediation analysis. The mediation models were adjusted for age
261 and gender. We estimated the total effect, direct effect, and indirect effect of the gut
262 mycobiome or bacterial KOs/faecal metabolites on the metabolic traits from these
263 models. The results were confirmed by the simulation exercises bootstrapped 1000
264 times. To estimate whether the association between the *Saccharomycetales spp.* and
265 *Pichia* changes with increased age, we compared the association not only between the
266 middle-aged and the elderly based on the cross-sectional data, but also between the
267 baseline and follow-up visits based on the fungal repeated measures.

Fig. S1. Study design for the present analyses. Our study was based on the Guangzhou Nutrition and Health Study (GNHS) which is an ongoing community based cohort study in China. We included 1244 participants in the current study. We profiled the gut mycobiome using ITS2 sequencing. 184 individuals have collected stool sample twice at different time points. The dietary information was collected using the food frequency questionnaires (FFQs) every three years since 2008. To examine the associations between gut fungi and bacterial taxonomy and function, we applied qiime2 and HUMAnN2 to perform the taxonomic and functional profiling from 16S rRNA sequencing data and stool metagenomics data. In addition, we investigated the associations between gut mycobiome and fecal metabolome. We employed linear regression models to examine the interaction between gut fungi and bacterial diversity on the cardiometabolic health. Finally, the bi-directional mediation analyses were performed to check whether gut mycobiome can influence the human phenotypes through gut bacterial gene function. Figure created with BioRender.com.

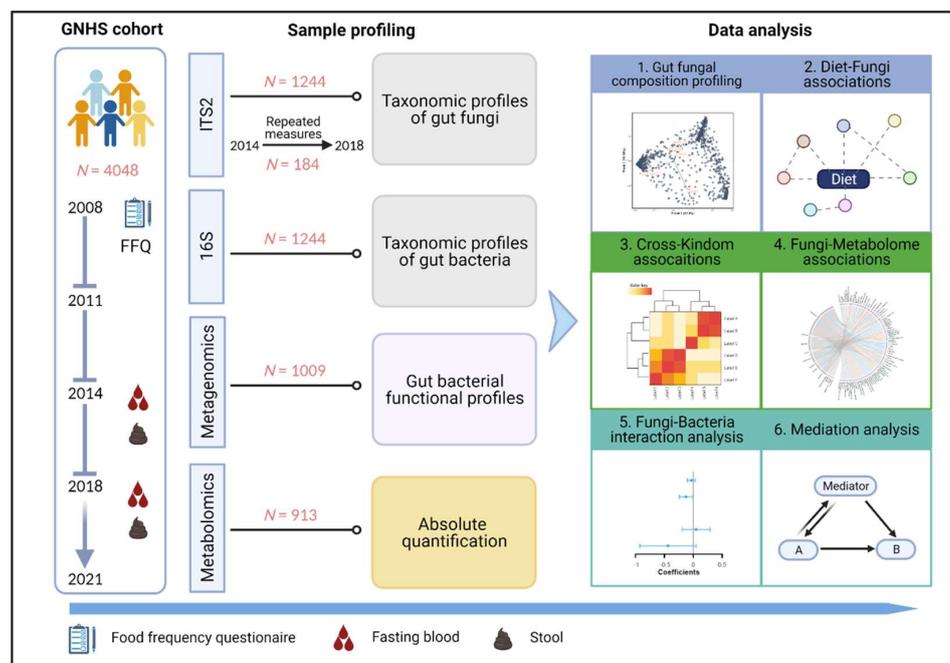


Fig. S2. The composition of gut mycobiome in different cohorts. (A) Age distribution of the GNHS cohort in the present study (N = 1244). **(B)** Bar plot shows the mean abundance ratio between the *Ascomycota* and *Basidiomycota*. **(C)** Venn plot shows the number of genera detected in three cohorts.

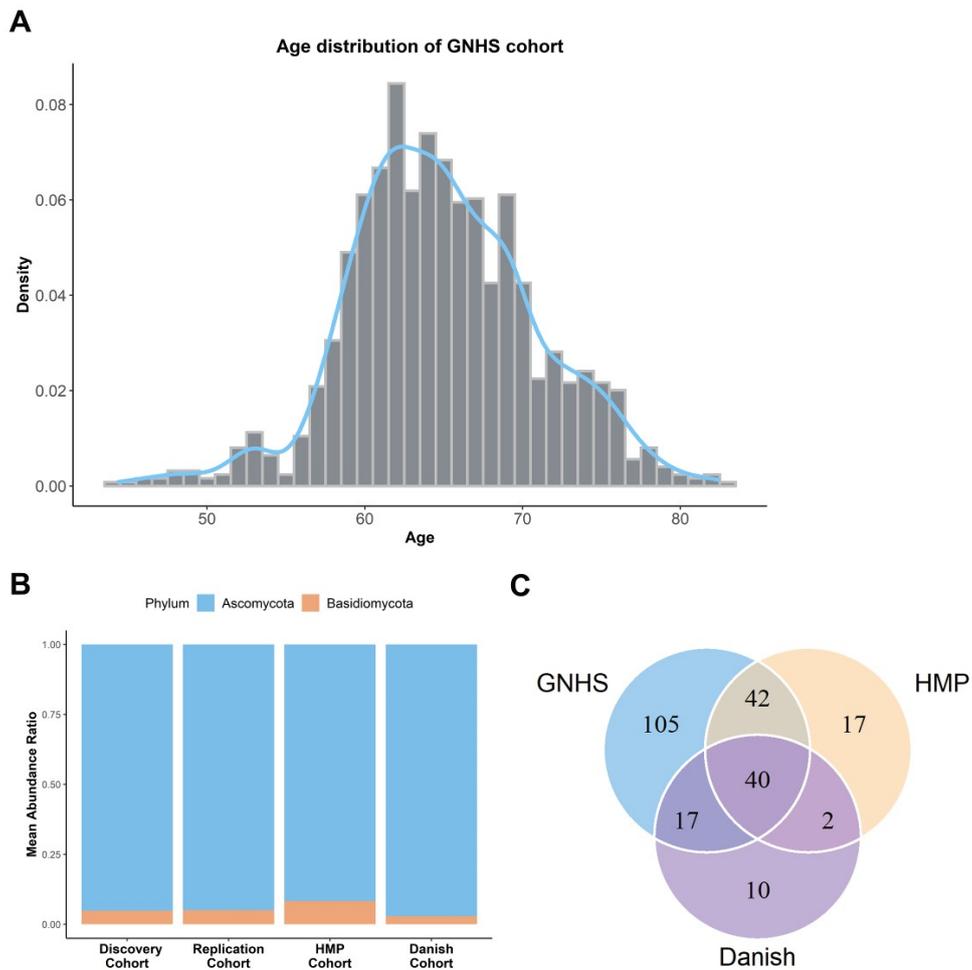


Fig. S3. Comparisons of the gut mycobiome between baseline and follow-up visit in the population with repeated measures (N = 184). The bar plots show the mean relative abundance (A) and prevalence (B) of the gut mycobiome at genus level. The median follow-up time was 3.2 years. (C) Bar plot shows the probability that the taxonomy (prob > 40%; 2 phyla and 7 genera) could be detected repeatedly at two time points.

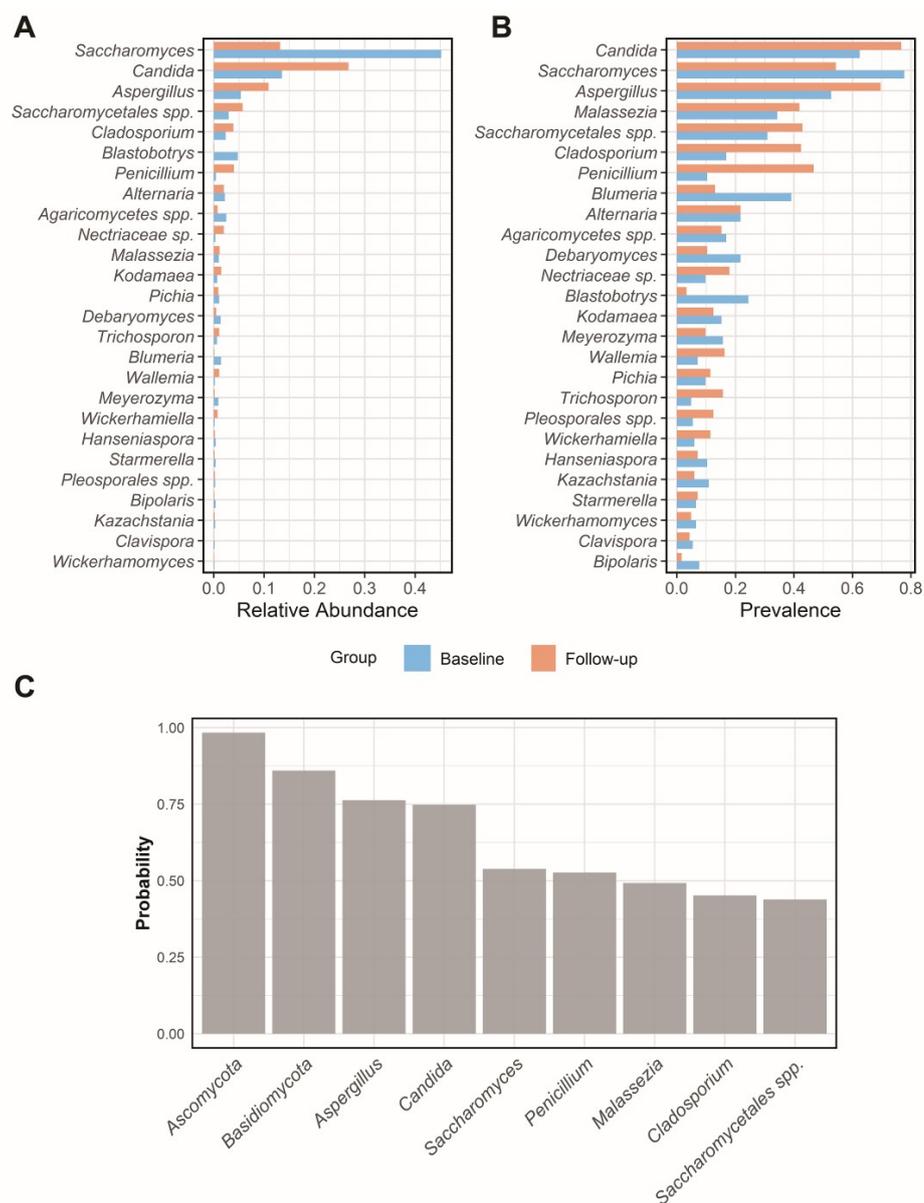


Fig. S4. Comparison of the composition between gut fungi and bacteria. A, The comparison of α -diversity indices between gut fungi and bacteria. **B,** The comparison of the inter-individual Bray-Curtis distance of gut fungi and bacteria. All box plots are the median with the interquartile range. p value was from Mann-Whitney U test.

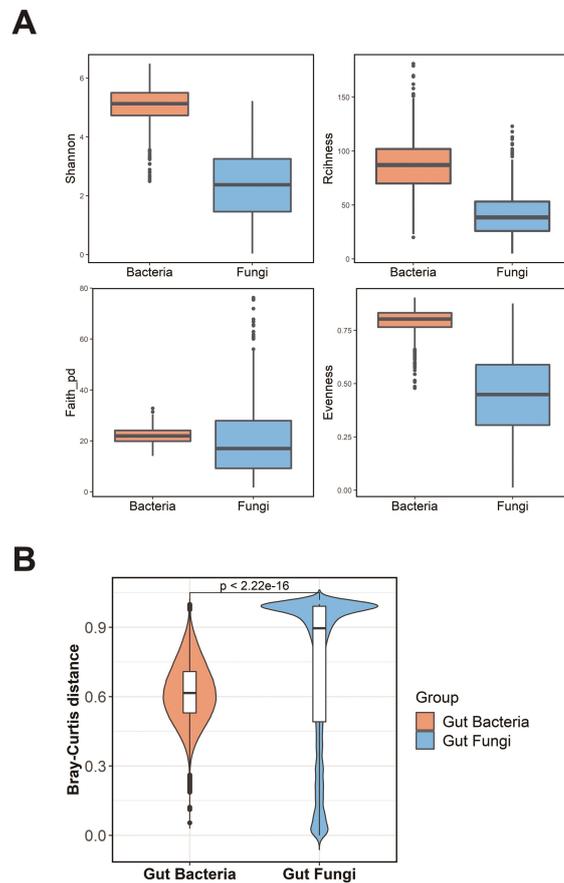


Fig. S5. Associations between gut fungal and bacterial composition. Procrustes analysis of gut fungi (ITS2) versus gut bacteria (16S) in 1244 participants. Fungi and bacteria are shown as orange and blue dots, respectively. Fungi and bacteria from the same individual are connected by grey lines.

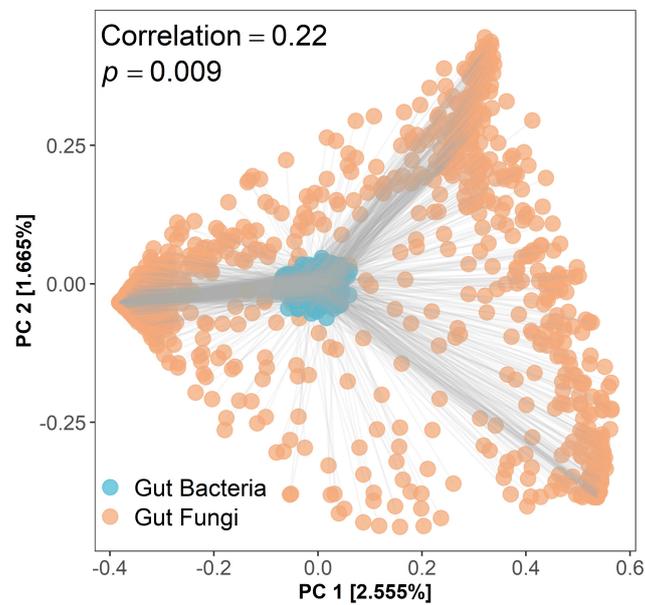


Fig. S6. Associations between the α -diversity indices of gut microbiome and host phenotypes. Heatmap shows the Spearman correlation coefficients between the α -diversity indices and host phenotypes. Total number of participants in each analysis was 1235 for waist circumference and hip circumference, 1242 for neck circumference, 1210 for fasting glucose, 1019 for insulin, 1212 for HbA1c, 1211 for triglycerides, total cholesterol, HDL-C, LDL-C and TC/HDL, 1239 for diet and 1244 for the other variables. bA1c, glycated hemoglobin; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC/HDL, total Cholesterol/HDL Cholesterol ratio; Middle-aged, age < 65 (N = 669); Elderly, age \geq 65 (N = 575).

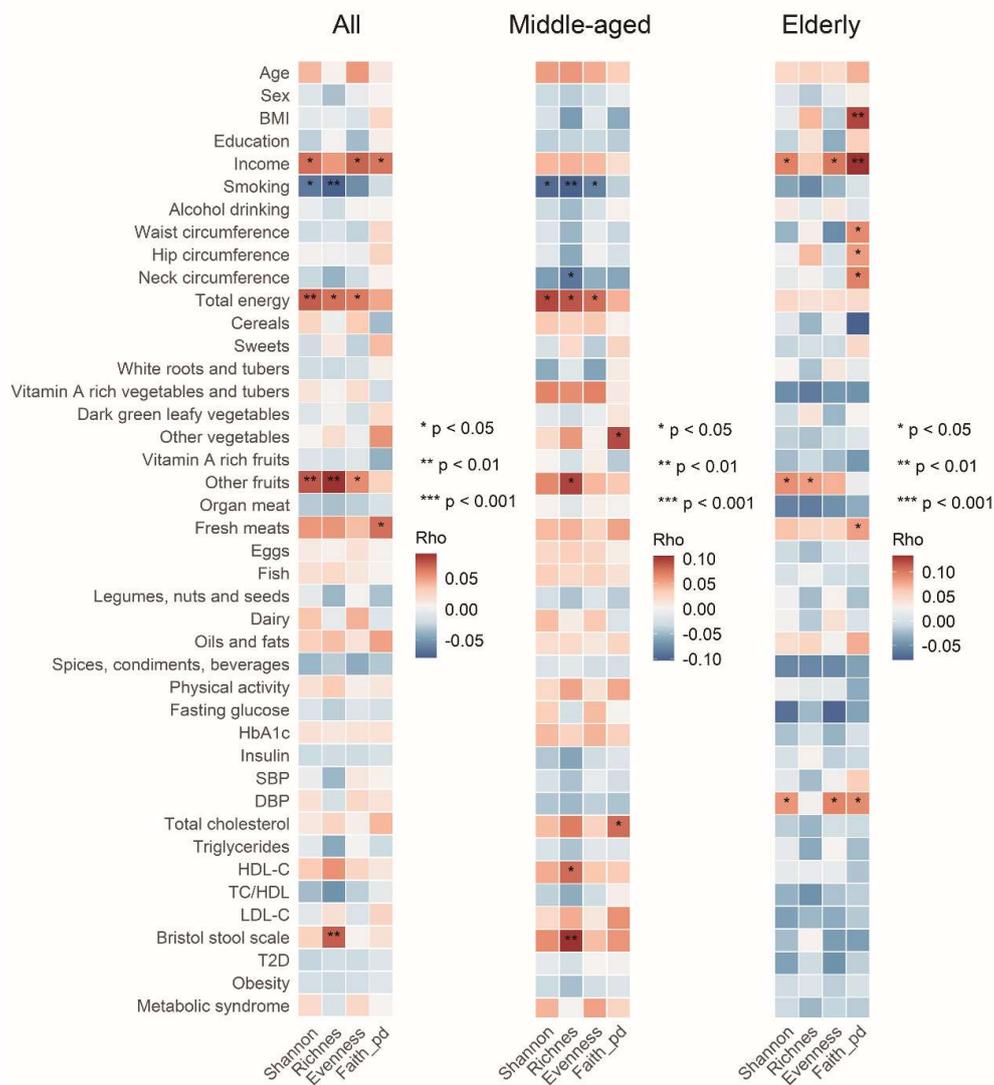


Fig. S7. The number of associations within the microbiome. **A**, Venn plot shows the number of associations between the gut fungi and bacteria assessed by SparCC and Spearman correlation analysis. **B**, Venn plot shows the number of inter- and intra-kingdom associations between gut bacteria and fungi.

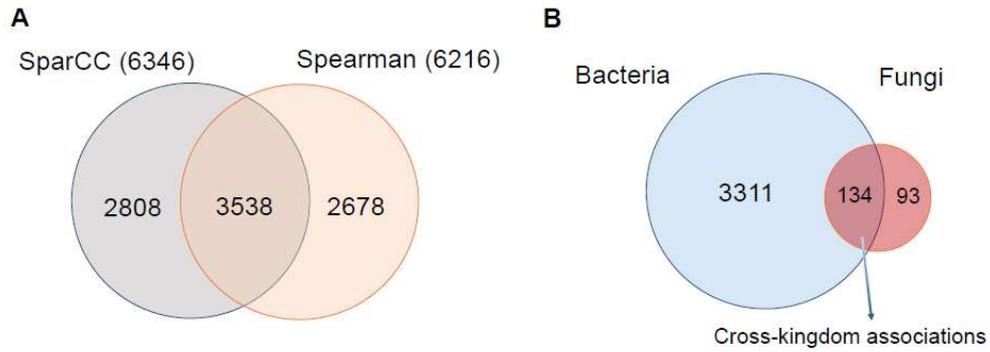


Fig. S8. Associations between the α -diversity indices of gut bacteria and gut mycobiota. Heatmap shows the Spearman correlation coefficients between specific gut fungi and the α -diversity indices of gut bacteria. *sp.* or *spp.* here represents unidentified fungal genus.

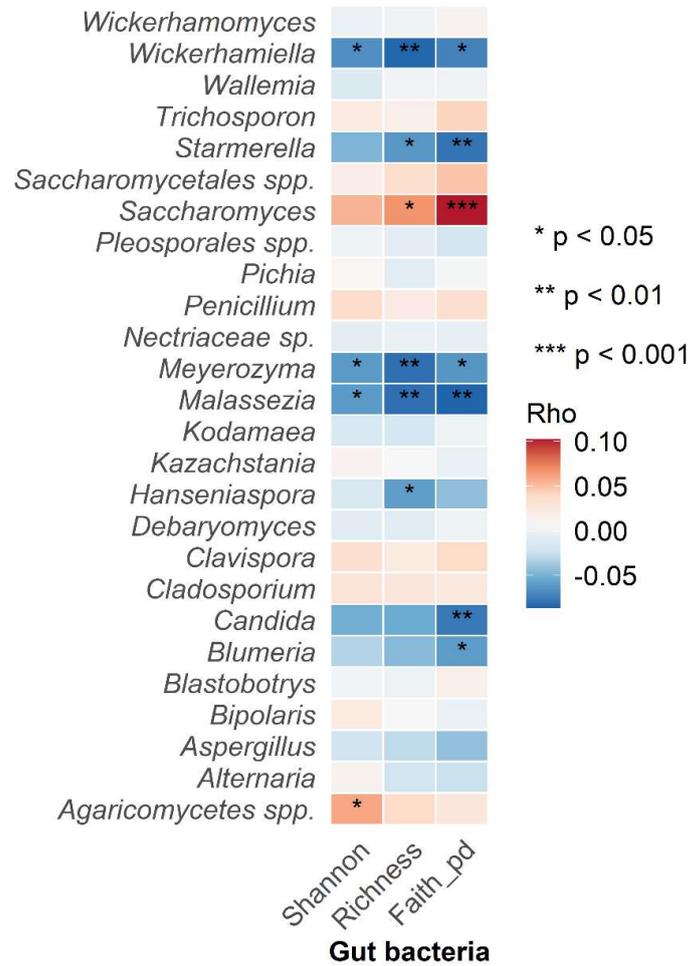


Fig. S9. Associations between gut mycobiome and fecal metabolome. A, Procrustes analysis of gut mycobiome versus fecal metabolome in 913 participants. Fungi and metabolites are shown as purple and green dots, respectively. Fungi and metabolites from the same individual are connected by grey lines. **B,** Venn plot shows the number of metabolites associated with gut fungal composition and specific fungi.

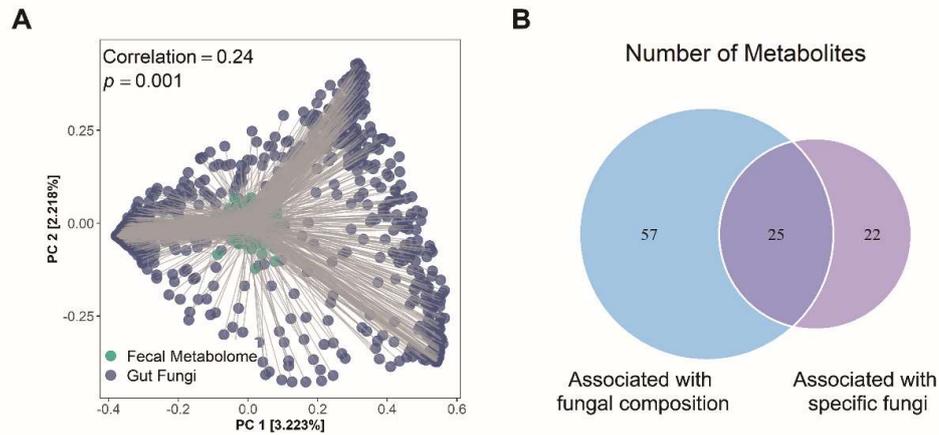


Fig. S10. Associations between gut mycobiome and metabolic traits in replication cohort. Heatmap of cross-sectional associations between gut fungi abundance and presence (left: continuous variable; right: binary variable) and metabolic traits. Linear regression models were adjusted by age and sex. Total number of participants in each analysis was 89 for fasting glucose, HDL cholesterol, HOMA-IR, LDL cholesterol, triglycerides, total cholesterol and TC/HDL, 90 for HbA1c and insulin, and 93 for BMI and waist circumference. Triglycerides, fasting glucose, TC/HDL and insulin were log-transformed.

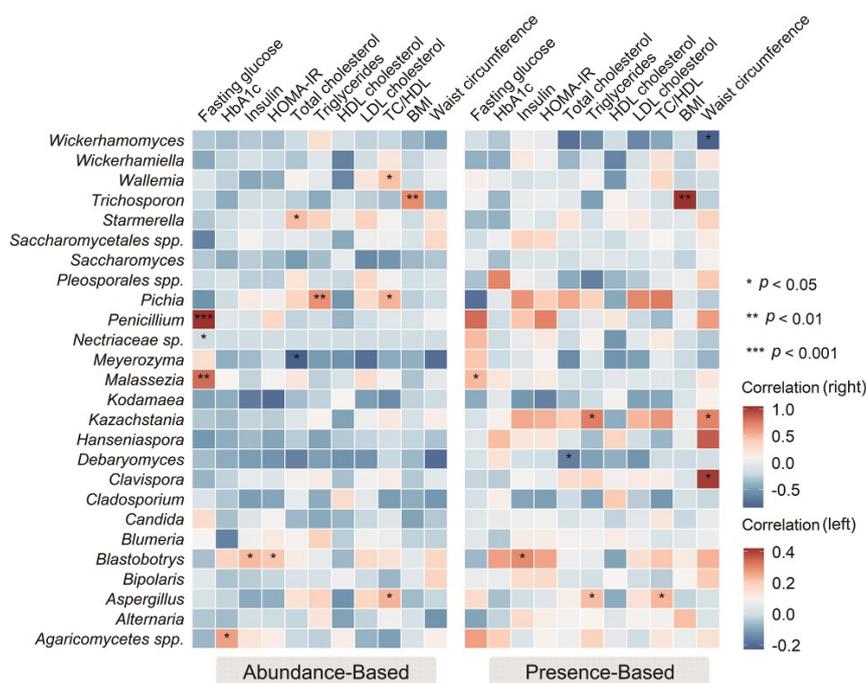


Fig. S11. Forest plot of the interaction of gut fungi with bacterial diversity on cardiometabolic traits. Associations were expressed as the difference in cardiometabolic traits (in SD unit) per 1 SD in each bacterial α -diversity index. Linear regression models were adjusted for age and sex. Associations were stratified by the presence of gut fungi. The error bars represent confidence intervals.

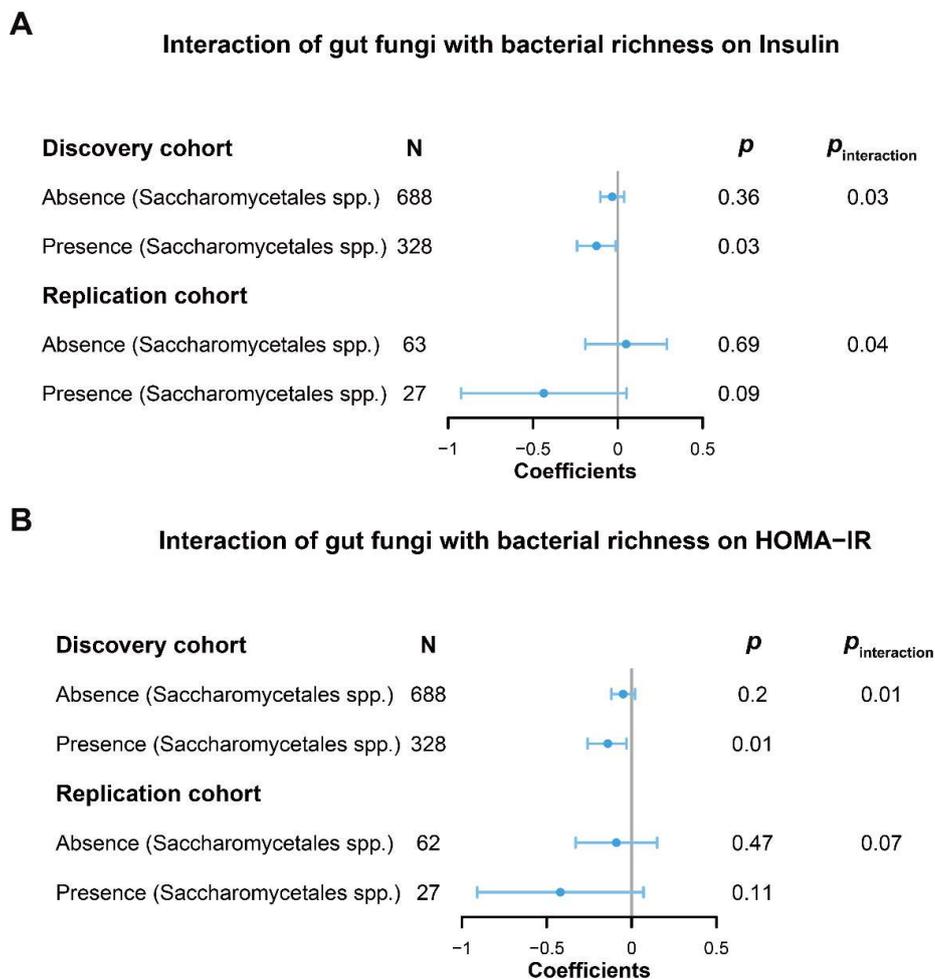


Fig. S12. Mediation linkages among the gut mycobiome, faecal metabolites and metabolic traits. p_{IDE} and $p_{\text{inv-IDE}}$ were estimated by the bi-directional mediation analysis. The metabolic traits were the longitudinal change between the baseline and follow-up visits while the gut mycobiome and faecal metabolites were from the baseline (N = 270). The red arrowed lines indicate the gut fungal effects on the changes of phenotype mediated by faecal metabolites with corresponding mediation p_{IDE} . Inverse mediation was performed to check whether faecal metabolites can influence the human phenotype through gut mycobiome. The percentage number in the center of each figure are effect ratios reflecting the proportion of the total effect of the independent variable on the dependent variable that is explained by the mediator, i.e., 29%. IDE, indirect effect; inv-IDE, inverse indirect effect. Figure created with BioRender.com.

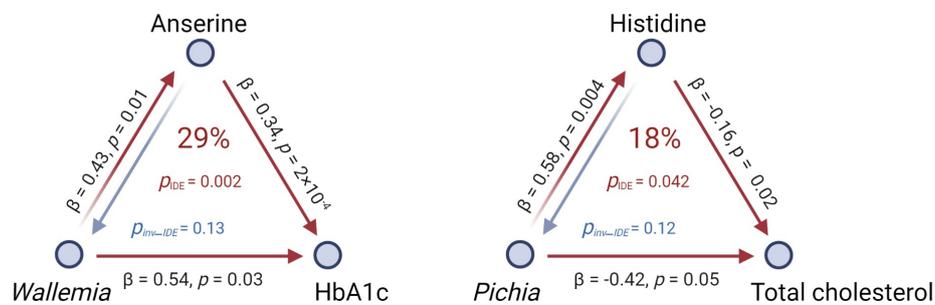
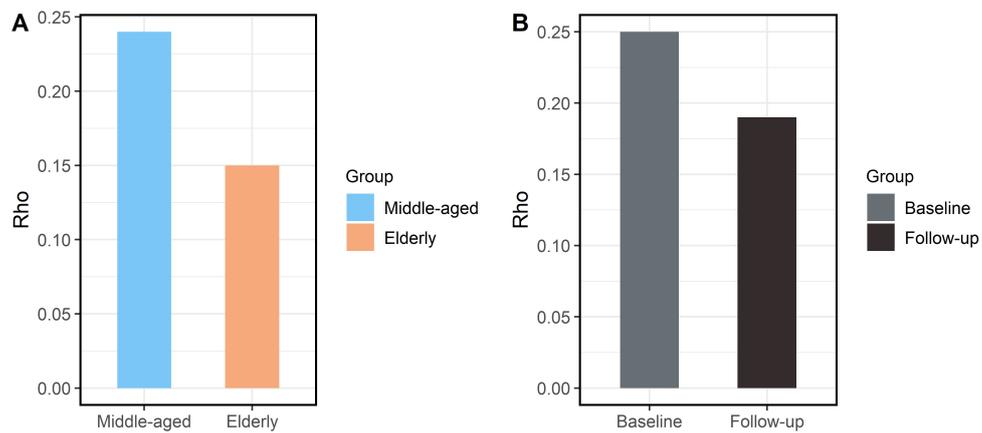


Fig. S13. Association between *Saccharomycetales spp.* and *Pichia*. **A**, Spearman correlation coefficients between the middle-aged (age < 65, N = 669, baseline) and the elderly (age ≥ 65, N = 575, baseline). **B**, Spearman correlation coefficients between the baseline and follow-up visits (N = 184). The median follow-up time was 3.2 years. All $p < 0.01$.



References

1. Zhang ZQ, He LP, Liu YH, et al. Association between dietary intake of flavonoid and bone mineral density in middle aged and elderly Chinese women and men. *Osteoporosis Int* 2014;25(10):2417-25. doi: 10.1007/s00198-014-2763-9
2. Fan F, Xue WQ, Wu BH, et al. Higher Fish Intake Is Associated with a Lower Risk of Hip Fractures in Chinese Men and Women: A Matched Case-Control Study. *Plos One* 2013;8(2) doi: ARTN e5684910.1371/journal.pone.0056849
3. Shuai ML, Zuo LSY, Miao ZL, et al. Multi-omics analyses reveal relationships among dairy consumption, gut microbiota and cardiometabolic health. *Ebiomedicine* 2021;66 doi: ARTN 10328410.1016/j.ebiom.2021.103284
4. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37(8):852-57. doi: 10.1038/s41587-019-0209-9
5. Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;13(7):581-+. doi: 10.1038/Nmeth.3869
- [dataset] 6. Nash AK, Auchtung TA, Wong MC, et al. The Intestinal Mycobium of the Human Microbiome Project Healthy Cohort. Sequence Read Archive, November 4, 2017. doi: 10.1186/s40168-017-0373-4
- [dataset] 7. Ahmad HF, Mejia JLC, Krych L, et al. Gut Mycobium Dysbiosis Is Linked to Hypertriglyceridemia among Home Dwelling Elderly Danes. European Nucleotide Archive, February 20, 2020. <https://doi.org/10.1101/2020.04.16.044693>
8. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 2011;27(6):863-4. doi: 10.1093/bioinformatics/btr026 [published Online First: 2011/02/01]
9. Franzosa EA, McIver LJ, Rahnavard G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 2018;15(11):962-68. doi: 10.1038/s41592-018-0176-y [published Online First: 2018/11/01]
10. Jiang Z, Sun TY, He Y, et al. Dietary fruit and vegetable intake, gut microbiota, and type 2 diabetes: results from two large human cohort studies. *BMC Med* 2020;18(1):371. doi: 10.1186/s12916-020-01842-0 [published Online First: 2020/12/04]
11. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci* 2003;14(6):927-30. doi: Doi 10.1658/1100-9233(2003)014[0927:Vaporf]2.0.Co;2
12. Watts SC, Ritchie SC, Inouye M, et al. FastSpar: rapid and scalable correlation estimation for compositional data. *Bioinformatics* 2019;35(6):1064-66. doi: 10.1093/bioinformatics/bty734
13. Tingley D, Yamamoto T, Hirose K, et al. mediation: R Package for Causal Mediation Analysis. *J Stat Softw* 2014;59(5)