Inter-determination of blood metabolite levels and gut microbiome supported by Mendelian randomization

Xiaomin Liu[1,#], Xin Tong[1,#], Yuanqiang Zou[1,2], Xiaoqian Lin[1,3], Hui Zhao[1,3], Liu Tian[1], Zhuye Jie[1,2], Qi Wang[1,3], Zhe Zhang[1], Haorong Lu[4], Liang Xiao[1,6,7], Xuemei Qiu[1], Jin Zi[1], Rong Wang[1], Xun Xu[1], Huanming Yang[1,8], Jian Wang[1,8], Yang Zong[1], Weibin Liu[1], Karsten Kristiansen[1,2], Yong Hou[1], Shida Zhu[1], Huijue Jia[1,5,†], Tao Zhang[1,2,†]

1. BGI-Shenzhen, Shenzhen 518083, China;
2. Department of Biology, Ole Maaløes Vej 5, University of Copenhagen, DK-2200 Copenhagen, Denmark;
3. BGI Education Center, University of Chinese Academy of Sciences, Shenzhen 518083, China;
4. China National Genebank, BGI-Shenzhen, Shenzhen 518120, China;
5. Shenzhen Key Laboratory of Human Commensal Microorganisms and Health Research, BGI-Shenzhen, Shenzhen 518083, China;
6. Shenzhen Engineering Laboratory of Detection and Intervention of Human Intestinal Microbiome, BGI-Shenzhen, Shenzhen 518083, China;
7. BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China;
8. James D. Watson Institute of Genome Sciences, Hangzhou 310058, China;

# These authors contributed equally to this work.

† To whom correspondence should be addressed: T.Z., tao.zhang@genomics.cn and H.J., jiahuijue@genomics.cn

Abstract

The gut microbiome has been implicated in a variety of physiological states. Controversy over causality, however, has always haunted microbiome studies. Here, we utilized the bidirectional Mendelian randomization (MR) approach to address questions that are not yet mature for more costly randomized interventions. From a total of 3,432 Chinese individuals with shotgun sequencing data for whole genome and whole metagenome, as well as anthropometric and blood metabolic traits, we identified 58 causal relationships between the gut microbiome and blood metabolites, and replicated 43 out of the 58. Gut microbiome could determine features in the blood. For example, increased fecal relative abundances of *Oscillibacter* and *Alistipes* were causally linked to decreased triglyceride concentration, and fecal microbial module pectin degradation might increase serum uric acid. On the other hand, blood features may determine gut microbial features, e.g. glutamic acid appeared to decrease *Oxalobacter*, and a few members of *Proteobacteria* were unidirectionally influenced by cardiometabolically important metabolites such as 5-methyltetrahydrofolic acid, alanine, as well as selenium. This study illustrates the value of human genetic information to help prioritize gut microbial features for mechanistic and clinical studies. The results are consistent with whole-body cross-talks of the microbiome and the circulating molecules.

Introduction

Metagenome-wide association studies (MWAS) using human stool samples, as well as animal models especially the germ-free mice, have pointed to a potential role of the gut microbiome in diseases such as cardiometabolic, autoimmune, neuropsychiatric diseases and cancer, with mechanistic investigations for diseases such as obesity, colorectal cancer and schizophrenia[1-4]. Twin-based heritability estimation and more recent metagenome-genome-wide association studies (M-GWAS) have questioned the traditional view of the gut microbiota as a purely environmental factor[5-9], although the extent of the genetic influence remained controversial[7,10]. Yet, all these published cohorts, except for human sequences in the metagenomic data of HMP (Human Microbiome Project), utilized array data for human genetics, and most of them had 16S rRNA gene amplicon sequencing for the fecal microbiota[5-10].

For metabolic traits, a large number of GWAS analyses have been reported[11-15]. Yet, most of them focused on imputed genotyping array data for the discovery of common variants influencing the human blood metabolome, except for two recent studies[14,15] which leveraged whole genome or exome sequencing to discover metabolic quantitative trait loci (mQTL). These studies consistently indicated high heritability of blood metabolites.

As the gut microbiome is considered to be highly dynamic, causality has always been an unresolved issue on any reported difference. Mendelian randomization (MR)[16] offers an opportunity to distinguish between causal and non-causal effects from cross-sectional data, without animal studies or randomized controlled trials. An early study used MR to look at the gut microbiota and ischemic heart disease[17]. Recently, a study used MR to confirm that increased relative abundance of bacteria producing the fecal volatile short-chain fatty acid (SCFA) butyrate was causally linked to improved insulin response to oral glucose challenge; in contrast, another fecal SCFA, propionate, were causally related to an increased risk of T2D[18]. However, both studies used genotype data, and it was not clear to what extent the genetic factors explained the microbial feature of interest.

In this study, we presented the first large-scale M-GWAS using whole genome and fecal microbiome, and bidirectional MR for the fecal microbiome and anthropometric features as well as blood metabolites. In a two-stage design from different cities in China, 58 causal links were identified from MR in the 4D-SZ discovery cohort of 2,002 individuals with high-depth whole-genome sequencing data (1,539 individuals with microbiome data for one-sample MR). 43 of the 58 causal effects were replicated in the low-depth whole-genome sequencing data of another 1,430 individuals (1,006 individuals with microbiome data for one-sample MR). In general, unidirectional causal effects could be found both from the gut to the blood and from the blood to the gut, but bidirectional effects were rarely detected. A few of the M-GWAS associations with gut microbial functional modules, e.g. module for lactose/galactose degradation

and the *ABO* loci, reached study-wise significance, illustrating the power of shotgun metagenomic data together with whole genome. The MR findings were corroborated and extended by summary statistics from the Japan Biobank study, e.g. causal effect of *Proteobacteria* on T2D (Type 2 diabetes mellitus), congestive heart disease and colorectal cancer, underscoring the significance of human genetic data to help guide microbiome intervention studies.

Results

Fecal microbiome associated with human genetics

We set out to identify human genetic variants to be included as the randomizing layer of MR (**Fig. 1**). The 4D-SZ (multi-omics, with more time points to come, from Shenzhen, China) discovery cohort consisted of high-depth whole-genome sequencing data from 2,002 blood samples (mean depth of 42×, ranged from 21× to 87×, **Supplementary Table 1, Supplementary Fig. 1a**), out of which 1,539 individuals had metagenomic shotgun sequencing data from stool samples (8.56 ± 2.28 GB, **Supplementary Fig. 1b**). Fecal M-GWAS was performed using 10 million common and low-frequency variants (minor allele frequency (MAF) ≥ 0.5%) and 500 unique microbial features (120 from the initial 620 microbial taxonomic or functional features was omitted due to strong association with other microbial features, Spearman's correlation > 0.99). The M-GWAS was adjusted for age, gender, BMI, defecation frequency, stool form, self-reported diet, lifestyle factors, and the first four principal components from the genomic data to account for population stratification.

With this so-far the largest cohort of whole genome and whole metagenome data, we performed M-GWAS analysis and identified a total of 625 associations involving 548 independent loci for one or more of the 500 microbial features at genome-wide significance ($P < 5 \times 10^{-8}$). With a more conservative Bonferroni-corrected study-wide significant *P* value of 1.0×10^{-10} ($= 5 \times 10^{-8} / 500$), we identified 28 associations with fecal microbial features involving 27 genomic loci, of which 5 correlated with gut bacteria and the other 22 associated with gut metabolic pathways (**Supplementary Table 2**).

For MR, it was important for the genetic variants used to be representative of the microbiome features (**Supplementary Fig. 2**), so a more suggestive *P* value of lower than 1×10^{-5} was used (**Supplementary Table 2**), as in previous MR studies[18,19]. Each microbial feature had an average of 44 genetic variants (range: 4-262; sd: 38; **Fig. 2a**, **Supplementary Table 3**). The corresponding genetic variants explained microbial features to a median value of 24.9%, e.g. 45.5% of the microbial metabolic pathway for succinate consumption and 44.6% of *Phascolarctobacterium succinatutens* (an asaccharolytic, succinate-utilizing bacterium), while only 6.8% of genus *Edwardsiella* (**Supplementary Table 3**). The phenotypic (relative abundance) variance of five genera *Bilophila*, *Oscillibacter*, *Faecalibacterium*, *Megasphaera* and

Bacteroides could be explained over 35% by their corresponding independent genetic variants (**Fig. 2b**), and the same is true for species *Bilophila wadsworthia*, *Eubacterium siraeum* and *Faecalibacterium prausnitzii* (a butyrate-producing bacterium that was relatively depleted in metabolic and immune diseases). Thus, although human genetic associations (array data) have been reported to explain only 10% or 1.9% of the gut microbiota[7,10], the suggestive associations from the current M-GWAS study could be highly predictive of certain gut taxa and functions.

For better confidence in these suggestive associations, we sequenced a replication cohort of 1,430 individuals from multiple cities in China (also shotgun metagenomic sequencing for stool samples to an average of 8.65 ± 2.42 GB (**Supplementary Fig. 1d**), but about 8× whole-genome sequencing for human genome, ranged from 6× to 16× (**Supplementary Fig. 1c**)). Among the 22,293 independent associations identified in the discovery cohort with $P < 10^{-5}$, 4,876 variants were not available in the low-depth replication dataset and 87.6% of them were not common variants (MAF < 0.05), which was understandable given the relatively low detection rate of rare genetic variants from 8× sequencing data. For the remaining 17,417 independent associations covered by the low-depth replication dataset, we were able to replicate 2,324 in the same effect direction of minor allele ($P < 0.05$, **Supplementary Table 2**), indicating that the associations were not random false positives. The fraction of associations replicated in the same direction ($P < 0.05$) using the suggestive cut-off of $P < 10^{-5}$ (2,324/17,417) was not lower than the more stringent cut-offs (54/625 of the $P < 5x10^{-8}$, and 2/28 of the $P < 10^{-10}$). Two well replicated signals from the study-wide threshold were chr9:133276163 in the *ABO* blood group associated with module MF0007: lactose and galactose degradation ($P_{discovery} = 2.10 \times 10^{-12}$ and $P_{replication} = 1.09 \times 10^{-10}$; **Supplementary Fig. 3a,b**) and rs142693490 near the *LCORL* gene (implicated in spermatogenesis, body frame and height) associated with MF0034: alanine degradation II ($P_{discovery} = 1.28 \times 10^{-12}$ and $P_{replication} = 0.014$; **Supplementary Fig. 3c,d**). Chr9:133276163 is in strong linkage disequilibrium (LD, $r^2 = 0.99$) with multiple SNPs (rs507666, rs532436, rs651007, rs579459 and rs579459) in the *ABO* gene. These SNPs located in a block were found to be associated with metabolites levels in both this study and previous studies, especially for serum alkaline phosphatase levels (**Supplementary Table 4**). Other fecal microbiome associations confirmed by the low-depth genomes included: *AMIGO1* associated with MF0067:glycolysis (preparatory phase); *RAD51B* associated with MF0019: rhamnose degradation; *IPO8* associated with MF0014: arabinose degradation; *LINC00648* associated with *Streptococcus oralis*; *PLEKHF2* associated with MF0050: threonine degradation II; *IPO8* associated with MF0037: leucine degradation; *RTRAF* associated with *Bacteroides xylanisolvens*; *GNB1* associated with *Megasphaera elsdenii*; *DOCK8* associated with *Actinomyces* etc. (**Supplementary Table 2**).

Besides, 175 loci have been previously[6-10] reported to associate with specific taxa. We were able to replicate 4 of them at nominal significance, including rs147600757 associated with Rikenellaceae and rs62273067 associated with *Acidaminococcus*

reported by Turpin et al.[8], rs10115898 associated with *Streptococcus mutans* and rs78859629 associated with *Lactobacillus acidophilus* reported by Rothschild et al.[10]. To accommodate the differences in taxonomic resolution between amplicon data and our shotgun data, we obtained a minimal *P* value for each SNP across all taxa, and replicated 8 of them at the phylum level ($P < 0.05/134 = 3.7 \times 10^{-4}$, with 134 of the 175 loci available in this study; **Supplementary Table 5**), especially for rs12354611 and *Bacteroides stercoris* ($P = 8.64 \times 10^{-6}$).

Blood metabolic traits associated with human genetics

On the other hand, plasma metabolites are also called on to associate with host genetics (**Fig. 1**). We thus performed whole genome-wide association tests with an additive model on 10 million common and low-frequency variants (MAF ≥ 0.5%) for each of the 112 metabolites, with log-transformed relative abundance. We identified a total of 174 associations involving 158 loci that independently associated with one or more of the 112 metabolites at genome-wide significance ($P < 5 \times 10^{-8}$). With a more conservative Bonferroni-corrected study-wide significant *P* value of 4.5×10^{-10} ($= 5 \times 10^{-8}/112$ metabolites), we identified 39 associations with metabolites involving 28 genomic loci (**Supplementary Table 6**). These included previously well-established associations such as the *UGT1A* family associated with serum total bilirubin[11,20] and *ASPG* associated with asparaginate[11].

According to the suggestive threshold of $P < 10^{-5}$, we identified 6,541 mQTLs, of which 361 were associated with two or more metabolites. Summary statistics for all independent genetic variants associated with metabolic traits with a *P* value lower than 1×10^{-5} are included in **Supplementary Table 6**. The average number of genetic variants was 58 for each metabolic trait (range: 14-240; sd: 36, **Fig. 3a; Supplementary Table 7**). The percentage of variance explained by the corresponding genetic variants ranged from 13.3% (red blood cell distribution) to as high as 48.3% (blood mercury concentration) and 45.9% (blood alpha-fetoprotein value), with a median value of 28.6% (**Fig. 3b**). Among these, 268 variants or their proxy variants ($r^2 > 0.6$; distance < 1MB) have been reported in the GWAS catalog[21] (**Supplementary Table 8**). Some variants were associated with diseases in the GWAS catalog such as chronic kidney disease, Alzheimer's disease, coronary artery disease, Crohn's disease, ovarian cancer, breast cancer and gastric cancer.

Among the 6,541 suggestive mQTLs identified in the 4D-SZ discovery cohort with $P < 10^{-5}$, 5,088 variants were covered by the replication dataset. 717 and 31 were replicated at nominal ($P < 0.05$) and suggestive significance ($P < 10^{-5}$), respectively, in the same effect direction of minor allele (**Supplementary Table 6**). Especially for the 174 genome-wide and 39 study-wide significant associations, we could replicate 51 and 29 associations in the same direction ($P < 0.05$), respectively. The top associations confirmed by the low-depth genomes ($P < 4.5 \times 10^{-10}$ both in discovery and replication cohorts) included: *FECH* associated with manganese; *UGT1A* family

associated with serum total bilirubin as well as direct and indirect (unconjugated) bilirubin; *ASPG* associated with asparagine; *CPS1* associated with glycine; *APOE* associated with low density lipoprotein; *LUC7L* associated with mean corpuscular hemoglobin concentration; *ALAD* associated with lead; *GADL1* associated with beta-alanine; *PRODH* associated with proline; *NPRL3* associated with red blood cell distribution. The association results of the top five traits were shown in **Supplementary Fig. 4**. Overall, the accurate identification of genetic determinants and the high variance explained for both microbial features and blood metabolites are optimal for MR analysis to investigate causality.

From observational correlation to Mendelian randomization

As a prerequisite for strong causality, we investigated the correlation between relative abundances of 500 unique fecal microbial features (taxa and functional modules) and 112 host metabolic traits using multivariate linear regression. After adjustment for gender and age, we observed 457 significant associations (false discovery rate (FDR) corrected $P < 0.05$, **Supplemental Table 9, online methods**). Three metabolites, glutamic acid, 5-methyltetrahydrofolic acid (5-methyl THF, active form of folic acid) and selenium, were associated with the largest number of microbial features (58, 40 and 38, respectively, **Supplementary Fig. 5**). These three metabolites were all associated with the phylum *Proteobacteria* and its constituents, including the family Enterobacteriaceae, genera *Escherichia*, *Methylobacillus* and *Achromobacter*, species *Escherichia coli*, *Pseudomonas stutzeri*, *Achromobacter piechaudii*, *Burkholderia multivorans* and *Methylobacillus flagellates*. Glutamic acid was positively correlated with *Proteobacteria*, whereas 5-methyltetrahydrofolic acid and selenium showed negative correlations with *Proteobacteria*, reminiscent of diverging associations of these metabolites with cardiometabolic diseases and inflammation. In addition to these top three metabolites, *Proteobacteria* also showed the strongest association among gut microbial taxa with another 5 traits (the amino acids hydroxyproline, aspartic acid, cystine, the metal strontium and the hormone aldosterone), suggesting that *Proteobacteria* is an important taxon for this Asian cohort. These associations extend findings from various studies, and suggest quantitative relationships between gut microbial taxa/functions and plasma metabolites.

To reveal the potential causal effects of the fecal microbial features on blood metabolites or the other way around, we conducted bidirectional Mendelian randomization analysis for the 457 observationally significant associations (FDR corrected $P < 0.05$ between metabolites and microbial features, **Supplemental Table 9**). For each trait, we selected independent genetic variants associated with the respective features as instruments ($r^2 < 0.1$ and $P < 1 \times 10^{-5}$). Consistent with previous studies[18,19], the threshold of $P < 1 \times 10^{-5}$ was used to include more variants and maximize the strength of genetic instruments. This threshold ensured that the genetic instruments were not too weak for the low-depth replication cohort (**Methods,**

Supplementary Fig. 6 and Supplemental Table 10). The average F statistic, a measure of the strength of these genetic instruments, was 51.4 (standard deviation (SD): 35.8) for the replication cohort, while an F statistic >10 is considered sufficiently informative for MR analyses[22]. The average microbial variance explained by the genetic instruments was 22.6% for the discovery cohort and 5.09% for the replication cohort (**Supplementary Fig. 2**). These exceeded the commonly reported 1.9%-5% in certain phenotypes due to missing heritability[23].

As we were fortunate to have all the data in the same cohort, we first performed one-sample MR analysis to identify causal relationships for the 457 observational correlations in the discovery cohort consisting of 1,539 individuals with both metabolic and microbiome traits. We found 58 significant causal effects, of which 17 showed causal effects for gut microbial features on blood metabolic traits and the other 41 showed causal effects for blood metabolic traits on gut microbial features ($P < 1.09 \times 10^{-4} = 0.05/457$; **Fig. 4**, **Supplementary Table 11**). Only 4 of these were bidirectional. By applying one-sample MR analyses to the replication dataset of 1,006 low-depth genomes as well as metabolic and microbiome traits from individuals in different cities, we could replicate 43 of the 58 causal relationships (in the same direction and $P < 0.05$; **Supplementary Table 11**), indicating that the effects were not random false positives.

Moreover, we also used six different two-sample MR methods, which are more commonly performed when only summary statistics are available from two different cohorts, to analyze our data both in the discovery cohort (summary data for 2,002 samples with metabolic traits and 1,539 samples with microbial features) and the replication cohort (summary data for 1,430 samples with metabolic traits and 1,006 samples with microbial features). The one-sample MR and the two-sample MR analyses showed highly consistent results, and the Spearman's correlation for beta coefficients between one-sample and two-sample MR reached 0.767 for the discovery cohort ($P < 2.2 \times 10^{-16}$). The 58 causal associations identified by one-sample MR were also significant in the two-sample MR analyses. An additional 14 causal associations were identified by the two-sample MR analyses (**Supplementary Table 12**), possibly due to the larger cohort size. We also examined the presence of horizontal pleiotropy by using the MR-PRESSO Global test[24]. Only one causal association (the negative effect of selenium on the abundance of *Methylobacillus flagellates*, $P_{\text{MR-PRESSOGlobaltest}} = 0.01$; **Supplementary Table 9**) showed pleiotropy, while all the other 71 causal relationships showed no evidence of pleiotropy ($P > 0.05$). Thus, our MR analyses identified robust causal relationships between blood metabolic traits and specific features of the gut microbiome.

Effects of the gut microbiome on blood metabolic traits

As some of the MR-identified relationships appeared linked, hierarchical clustering was performed for the 12 microbial features and 8 blood metabolites involved in the

17 causal relationships from the gut microbiome to blood metabolites, which formed two clusters. One cluster involved decreasing the plasma levels of triglyceride and alanine by gut microbial taxa or functional modules; and the other involved decreasing the levels of 5-methyltetrahydrofolic acid or progesterone, but increasing serum uric acid or plasma glutamic acid by gut microbial features (**Fig. 4a**). Reassuringly, the species *Mobiluncus curtisii* was clustered next to its corresponding genus *Mobiluncus*, and modules including serine degradation and threonine degradation, sucrose degradation and pectin degradation, were likewise next to one another.

The most significant causal effect of *Oscillibacter* on decreasing blood triglyceride concentration (**Fig. 5a-c**), and to a lesser extent on lowering body-mass index (BMI) and waist-hip ratio (WHR), whereas the effect with plasma alanine was bidirectional. Using 134 genetic variants to construct a polygenic risk score (PRS) (134 genetic variants and the constructed PRS explained 39.3% and 49.6% of the phenotypic variance, respectively, **Fig. 3b and Supplementary Table 11**) for one-sample MR analysis in the discovery cohort, we estimated that each 1-s.d. increase in the abundance of *Oscillibacter* would generate a 0.261 mmol/L decrement in triglyceride concentration ($P = 2.53 \times 10^{-10}$), a 0.161 kg/m2 decrement in BMI ($P = 1.33 \times 10^{-4}$) and 0.126 ratio decrement in WHR ($P = 2.73 \times 10^{-3}$). This causal relationship was robust when four two-sample MR tests were performed ($P_{\text{GCTA-GSMR}} = 4.34 \times 10^{-11}$, $P_{\text{Inverse\_variance\_weighted}} = 2.45 \times 10^{-15}$, $P_{\text{weighted-median}} = 1.22 \times 10^{-7}$ and $P_{\text{MR-Egger}} = 1.35 \times 10^{-5}$) (**Fig. 5c**), and there was no evidence of horizontal pleiotropy ($P_{\text{MR-PRESSOGlobaltest}} = 0.18$; **Supplementary Table 12**). The reverse MR analysis (testing the effect of genetic predictors of triglyceride on *Oscillibacter* abundance) was significant but did not reach the multiple test corrected significance ($10^{-4} < P < 0.05$). *Oscilibacter* is a Gram-negative Clostridial bacteria, phylogenetically close to *Oscillospira*[25] which could produce valerate or butyrate. In addition, higher relative abundance of *Alistipes* was also associated with decreased blood triglyceride concentration ($P = 8.31 \times 10^{-8}$, **Fig. 5d)**. At the species level, both *A. shahii* ($P = 1.37 \times 10^{-6}$) and *unclassified Alistipes sp. HGB5* ($P = 3.36 \times 10^{-5}$) showed negative effects on blood triglyceride. The effect of both *Oscilibacter* and *Alistipes* for lowering blood triglyceride concentration were confirmed in the replication cohort ($P = 3.39 \times 10^{-4}$ and $P = 2.88 \times 10^{-4}$, respectively; **Supplementary Table 11 and 12**). These findings support the decrease in relative abundances of *Oscillibacter* and *Alistipes* in obese individuals compared to individuals with normal BMI reported in previous studies[26-28], suggesting that these bacteria as promising supplementation agents for individuals of a suitable genetic background.

The gut microbiome potential for pectin degradation II (42.6% of the variance explained by GRS) showed a handful of significant MR hits with blood traits (**Fig. 4a**), including positive effects on alanine ($P = 8.57 \times 10^{-5}$) and serum uric acid ($P = 1.34 \times 10^{-6}$), whereas negative effects on progesterone ($P = 6.68 \times 10^{-7}$). *Bacteroidetes* and *Fusobacteria* were the only two phyla that positively correlated with the abundance of pectin degradation II (Spearman rank correlation, $\rho = 0.48$ and 0.15, respectively),

which included the two previously reported pectin-degrading species *Bacteroides thetaiotaomicron* and *Fusobacterium varium*[29,30]. In the 4D-SZ cohort, *F. varium* correlated with pectin degradation II (Spearman's correlation, ρ = 0.12) and increased the blood alanine (P = 0.02) and serum uric acid (P = 0.04); *B. thetaiotaomicron* correlated with pectin degradation II (Spearman's correlation, ρ = 0.21) but showed no detectable effect on alanine or uric acid (P > 0.05; **Supplementary Fig. 7a,b,d**). Instead, *B. dorei,* the bacterial species most strongly correlated with pectin degradation II (Spearman rank correlation, ρ = 0.32, **Supplementary Fig. 7c**), positively contributed to alanine (P = 0.05) and serum uric acid levels ($P = 3.40 \times 10^{-4}$; **Supplementary Fig. 7d**).

Effects of blood metabolites on gut microbial features

For the 41 causal relationships from blood metabolic traits to gut microbial features (one-sample MR, **Supplementary Table 11**), hierarchical clustering revealed two clusters, one mostly involved decreasing abundance of bacteria by plasma alanine or glutamic acid, the other involved decreasing abundance of bacteria by selenium or 5-methyltetrahydrofolic acid (**Fig. 4b**). *F. prausnitzii* showed a negative effect on plasma selenium (**Fig. 4a**), while plasma selenium showed negative effects on gut *Proteobacteria* such as Enterobacteriaceae (e.g. *Escherichia coli*, $P = 3.79 \times 10^{-5}$), *Pseudomonas stutzeri* ($P = 1.06 \times 10^{-6}$), and modules such as arginine degradation II ($P = 2.65 \times 10^{-6}$), succinate conversion to propionate ($P = 3.55 \times 10^{-5}$), and anaerobic fatty acid beta oxidation ($P = 9.71 \times 10^{-5}$) (**Fig. 4b**).

Bacteria from the phylum *Proteobacteria* were negatively affected by not only selenium, but also 5-methyltetrahydrofolic acid (**Fig. 4b**). We directly verified the effect of 5-methyltetrahydrofolic acid on *Escherichia in vitro*. Supplementing 5-methyltetrahydrofolic acid in growth media indeed slowed down the growth of a strain of *Escherichia coli* AM17-9 compared to lower concentrations or absence of 5-methyltetrahydrofolic acid (**Supplementary Fig. 8**).

A handful of bacteria were also affected by glutamic acid. The negative influence of glutamic acid (48 variants with suggestive associations and the constructed PRS explained 24.9% and 25.4% of the phenotypic variance, respectively) on the genus *Oxalobacter* ($P = 1.56 \times 10^{-6}$) may help explain the lower prevalence of *Oxalobacter* in developed countries, besides the lower intake of oxalate and antibiotic use[31]. Whether limiting glutamic acid could raise *Oxalobacter* and prevent kidney stones remains to be tested. Glutamic acid negatively affected melibiose degradation (to glucose, galactose, $P = 2.05 \times 10^{-5}$ from two-sample MR), but showed positive effects on alanine degradation I ($P = 5.46 \times 10^{-5}$), anaerobic fatty acid beta-oxidation ($P = 9.36 \times 10^{-5}$), and bidirectional positive effect on serine degradation ($P = 6.85 \times 10^{-7}$ for serine degradation to glutamic acid and $P = 9.90 \times 10^{-6}$ for glutamic acid to serine degradation, respectively).

Causal relationships with the gut microbiome in the context of diseases

We further investigated the effects of the 72 significant causal relationships (**Supplementary Table 12**) involving 40 microbial features and 12 metabolic traits on diseases, by performing two-sample MR analysis using gut microbiome GWAS summary data in this 4D-SZ cohort, together with blood quantitative traits and diseases GWAS summary statistics from Japan Biobank[32] (**Fig. 1; Supplementary Table 13**), given that Japanese people have a genetic architecture similar to Chinese. Only routine blood parameters but no amino acids, hormones and microelements were included in the Japan Biobank study. Thus, only five of the 72 causal associations, involving triglyceride and serum uric acid were available for further investigation in the Japan Biobank data. The relationship between unclassified Lachnospiraceae bacterium 9_1_43BFAA and uric acid was reciprocal in the 4D-SZ cohort and we could replicate the causal effect of uric acid on increased unclassified Lachnospiraceae bacterium 9_1_43BFAA abundance in the Japanese cohort, whereas the reciprocal effect, i.e. potential effect of unclassified Lachnospiraceae bacterium 9_1_43BFAA on uric acid was not replicated, possibly due to lack of variants in the genotyped Japanese cohort (15 instead of 67, **Supplementary Table 14**). The other three associations were not replicated maybe due to the same reason. For example, genus *Oscillibacter* had 135 variants with $P < 10^{-5}$ in our summary data but only 15 were available in the Japan Biobank summary data.

MR inference using our gut microbiome M-GWAS summary data and diseases GWAS summary statistics from Japan Biobank found that *Alistipes* that showed negative effects on blood triglyceride in the 4D-SZ cohort, lowered the risks of cerebral aneurysm (**Supplementary Table 15**, $P = 4.61 \times 10^{-4}$) and hepatocellular carcinoma ($P = 0.045$) in the Japan Biobank cohort. According to the genetic associations we identified for *Proteobacteria*, we were able to see in Japan Biobank disease data that *Proteobacteria* increased the risk of T2D (**Fig. 6a;** $P = 7.61 \times 10^{-4}$, two-sample MR), congestive heart failure ($P = 0.003$) and colorectal cancer ($P = 0.047$). This is consistent with MWAS findings mainly for Enterobacteriaceae[1], and suggest that the metabolites identified above (5-methyltetrahydrofolic acid, selenium) might help prevent the diseases. Folic acid is indeed recommended for heart diseases[33]. In addition, *Escherichia coli* increased the risk of urolithiasis (**Fig. 6b;** $P = 0.009$) and hepatocellular carcinoma ($P = 0.04$) while decreased the interstitial lung disease risk ($P = 0.007$). Similarly, *Salmonella enterica* increased prostate cancer risk but decreased interstitial lung disease risk. The Pseudomonadales order was the only microbial feature showing a positive effect on pulmonary tuberculosis. The denitrifying bacteria *Achromobacter* increased the risk of atopic dermatitis ($P = 0.005$), gastric cancer ($P = 0.008$), esophageal cancer ($P = 0.027$) and biliary tract cancer ($P = 0.034$). *Bacteroides intestinalis* which was reported to be relatively depleted in patients of atherosclerotic cardiovascular disease[34] was found here to increase with potassium, and *B. intestinalis* showed a negative effect on epilepsy. *Streptococcus parasanguinis* had a positive effect on colorectal cancer and posterior wall thickness

(echocardiography), consistent with MWAS studies[1,34,35]. These results illustrated the potential significance of the gut microbiome-blood metabolite relationships in understanding and preventing cardiometabolic diseases and cancer.

Discussion

In summary, we identified abundant genetic loci to associate with microbial features and metabolic traits, and found 58 causal relationships between the gut microbiome and blood metabolites using one-sample bidirectional MR. 43 out of the 58 one-sample MR signals could be replicated in a low-depth genome cohort also from China. Two-sample MR replicated the 58 causal relationships in the same direction and identified an additional 14 causal relationships. Two-sample MR using summary statistics from Japan Biobank identified effects of gut microbial features on diseases, suggesting potential applications of microbiome intervention in cardiometabolic, kidney and lung diseases and cancer. While mechanistic investigations using germ-free mice and reference bacteria strains have been popular, our data-driven analyses underscore the clinical relevance of gut microbes that have not been extensively cultured and characterized, e.g. *Oscilibacter* and *Alistipes* for lowering triglyceride concentration and a number of disease risks, which may be particularly relevant for East Asian regions undergoing rapid changes in lifestyle and disease profiles.

By applying this MR analysis to explore causality, our results laid further support for several previously reported microbiome-metabolites relationships. For example, *B. thetaiotaomicron* had been reported to inversely correlate with serum glutamate concentration and was lower in obese individuals[36]. Consistently, we confirmed that species from Bacteroides, including *B. thetaiotaomicron*, *B. intestinalis*, *B. helcogenes* and *B. pectinophilus*, reduced plasma glutamic acid concentration ($10^{-4} < P < 0.05$). We also confirmed that *B. thetaiotaomicron* could lower plasma Alpha-aminoadipic acid level, weight and WHR ($P < 0.05$). Besides, we found that cysteine negatively correlated with abundance of *Escherichia coli*, which is consistent with previous finding that cysteine inhibited the growth of *Escherichia coli*[37].

Although associations between the gut microbiome and blood features such as amino acids and vitamins have been known for some time, our MR analyses could inspire more mechanistic and interventional studies. The unique data available from the 4D-SZ cohort allowed appreciation of overlooked features such as selenium. Selenium compounds were deemed essential for human health and development[38]. It is beneficial to an organism only in small amounts, while high concentrations of selenium become toxic[39]. We found that higher amount of blood selenium showed negative effects on some members of the gut microbiome (**Fig. 4b**). Although previous studies reported that *Escherichia coli* had evolved for adaptation to

selenate[40-43], the MR result that blood selenium negatively impacted the relative abundance of *Escherichia coli* suggested that it may still be sensitive to selenium. Increased selenium level had adverse effects on several other bacteria from Gammaproteobacteria, including *Achromobacter piechaudii*, *Methylobacillus flagellatus*, *Pseudomonas stutzeri*, and *Burkholderia multivorans*. *Pseudomonas stutzeri* is a nonfluorescent denitrifying and an opportunistic bacterium[44]. *Burkholderia multivorans* is a prominent *B. cepacia* complex species causing infection in people with cystic fibrosis[45]. Interestingly, *F. prausnitzii* from the *Firmicutes* phylum showed a negative effect on plasma selenium. Further studies on such indirect relationships between opportunistic pathogens and commensal bacteria would be illuminating, and could help to better protect individuals who have a genetic risk.

Nitrogen is a limiting resource for many ecosystems. In the modern human gut microbiome without high intake of nitrite, proteins are probably the major source of nitrogen[46], and the glutamate-glutamine reservoir is a key buffering mechanism for the inflammatory potential of excess amines[36,47-50]. The increase in *Proteobacteria* and decrease in Oxalobacteraceae observed in these Chinese individuals no more than 30 years old on average could potentially explain susceptibility to cardiometabolic and kidney diseases later in life. The bidirectional link between strontium and *Streptococcus parasanguinis* implies an interplay between water source and cardiovascular diseases[34,51].

Metabolism of polysaccharides that cannot be directly digested by the host is an important function of the colonic microbiome. We found degradation of pectin (or sucrose) to negatively affect progesterone level. This is an interesting possibility to provide scientific support for traditional dietary advice for pregnant women to ensure full-term pregnancy. Hyperuricemia and gout is a growing epidemic in East Asia, and soft drinks containing fructose is a strong factor that is no less important than beer and meat[52]. Gut microbial (*Bacteroides*, *Fusobacterium*) pectin degradation module positively contributed to circulating levels of alanine and uric acid. Further studies on the trans-kingdom metabolic flux of carbon and nitrogen would be necessary for personalized management of uric acid and alanine levels.

For the nascent field of M-GWAS and microbiome MR, there is also a lot of opportunities for methodological development by statistical experts. Low-frequency microbes are common in an individual's gut and could play physiological or pathological roles[1,53]. Our MR results for gut microbial species were supported by MR for higher taxonomic units such as genus or phylum (**Fig. 4**, **Supplementary Table 11**). Yet, the *P* values were sometimes more significant for the larger taxa, suggesting similar functions contributed by other species. Functional redundancy in the microbiome has been discussed ever since the beginning of the microbiome field[54,55], and here we identified study-wide significant host genetic associations with gut microbial functional modules, and causal effects of other gut microbial functional modules on host levels of circulating metabolites. Distribution of the microbiome

taxonomic or functional data constitutes another layer of consideration, in addition to the human allele frequencies. Gathering a more homogenous cohort could enable identification of signals in a relatively small cohort, while corrections for comparing different populations might involve host-microbiome interactions. As the gut microbiome can be influenced by medication[56], and heritability for most traits is higher in younger individuals[57], healthy young adults are probably preferable for M-GWAS studies, while microbiome-drug interactions in older individuals could be an important direction for MR studies.

In short, our data-driven approach underscores the great potential of M-GWAS and MR for a full picture of the microbiome, which can be mechanistically illuminating and are poised to help focus intervention efforts to mitigate inflammation and prevent or alleviate complex diseases.

Methods

Study subjects

All the adult Chinese individuals were recruited for a multi-omic study, with some volunteers having samples from as early as 2015, which would constitute the time dimension in '4D'. The discovery cohort was recruited during a physical examination from March to May in 2017 in the city of Shenzhen, including 2,002 individuals with blood samples and of which 1,539 had fecal samples. All these individuals were enlisted for high-depth whole genome and whole metagenomic sequencing. As for replication, blood samples were collected from 1,430 individuals, out of which 1,006 had fecal samples. The replication cohort was designed in the same manner but organized at smaller scales in multiple cities (Wuhan, Qingdao, etc.) in China. The protocols for blood and stool collection, as well as the whole genome and metagenomic sequencing were similar to our previous literature[5,48]. For blood sample, buffy coat was isolated and DNA was extracted using HiPure Blood DNA Mini Kit (Magen, Cat. no. D3111) according to the manufacturer's protocol. Feces were collected with MGIEasy kit and stool DNA was extracted in accordance with the MetaHIT protocol as described previously[58]. The DNA concentrations from blood and stool samples were estimated by Qubit (Invitrogen). 200 ng of input DNA from blood and stool samples were used for library preparation and then processed for paired-end 100bp and single-end 100bp sequencing, respectively, using BGISEQ-500 platform[59].

The study was approved by the Institutional Review Boards (IRB) at BGI-Shenzhen, and all participants provided written informed consent at enrolment.

High-depth whole genome sequence for discovery cohort

2,002 individuals in discovery cohort were sequenced to a mean of 42x for whole genome. The reads were aligned to the latest reference human genome GRCh38/hg38 with BWA[60] (version 0.7.15) with default parameters. The reads consisting of base quality <5 or containing adaptor sequences were filtered out. The alignments were indexed in the BAM format using Samtools[61] (version 0.1.18) and PCR duplicates were marked for downstream filtering using Picardtools (version 1.62). The Genome Analysis Toolkit's (GATK[62], version 3.8) BaseRecalibrator created recalibration tables to screen known SNPs and INDELs in the BAM files from dbSNP (version 150). GATKlite (v2.2.15) was used for subsequent base quality recalibration and removal of read pairs with improperly aligned segments as determined by Stampy. GATK's HaplotypeCaller were used for variant discovery. GVCFs containing SNVs and INDELs from GATK HaplotypeCaller were combined (CombineGVCFs), genotyped (GenotypeGVCFs), variant score recalibrated (VariantRecalibrator) and filtered (ApplyRecalibration). During the GATK VariantRecalibrator process, we took our variants as inputs and used four standard SNP sets to train the model: (1) HapMap3.3 SNPs; (2) dbSNP build 150 SNPs; (3) 1000 Genomes Project SNPs from Omni 2.5 chip; and (4) 1000G phase1 high confidence SNPs. The sensitivity threshold of 99.9% to SNPs and 99% to INDELs were applied for variant selection

after optimizing for Transition to Transversion (TiTv) ratios using the GATK ApplyRecalibration command. After applying the recalibration, there were 60,978,451 raw variants left, including 55 million SNPs, and 6 million INDELs.

We applied a conservative inclusion threshold for variants: (i) mean depth >8×; (ii) Hardy-Weinberg equilibrium (HWE) $P > 10^{-5}$; and (iii) genotype calling rate > 98%. We demanded samples to meet these criteria: (i) mean sequencing depth > 20×; (ii) variant calling rate > 98%; (iii) no population stratification by performing principal components analysis (PCA) analysis implemented in PLINK[63] (version 1.07) and (iv) excluding related individuals by calculating pairwise identity by descent (IBD, Pi-hat threshold of 0.1875) in PLINK. Only 10 samples were removed in quality control filtering. After variant and sample quality control, 1,992 individuals with 6.12 million common (MAF ≥ 5%) and 3.90 million low-frequency (0.5% ≤ MAF < 5%) variants from discovery cohort were left for subsequent analyses.

Low-depth whole genome sequence for replicate cohort

1,430 individuals in replication cohort were sequenced to a mean of 8x for whole genome. We used BWA to align the whole genome reads to GRCh38/hg38 and used GATK to perform variants calling by applying the same pipelines as for the high-depth WGS data. After completing the joint calling process with CombineGVCFs and GenotypeGVCFs options, we obtained 43,402,368 raw variants. A more stringent process in the GATK VariantRecalibrator stage compared with the high-depth WGS was then used, the sensitivity threshold of 98.0% to both SNPs and INDELs was applied for variant selection after optimizing for Transition to Transversion (TiTv) ratios using the GATK ApplyRecalibration command. Further, we kept variants with less than 10% missing genotype frequency and minor allele count more than 5. All these high-quality variants were then imputed using BEAGLE 5[64] with the 1,992 high-depth WGS dataset as reference panel. We retained only variants with imputation info. > 0.7 and obtained 10,905,418 imputed variants. We further filtered this dataset to keep variants with Hardy-Weinberg equilibrium $P > 10^{-5}$ and genotype calling rate > 90%. Similar to what we have done for discovery cohort, samples were demanded to have mean sequencing depth > 6×, variant call rate > 98%, no population stratification and no kinship. Finally, 1,430 individuals with 5,884,439 high-quality common and low-frequency variants (MAF ≥ 0.5%) from replication cohort were left for subsequent analysis.

To assess the data quality, we sequenced 27 samples with both high-depth and low-depth WGS data and then compared the 5,318,809 variants between them for each individual. The average genotype concordance was 98.66% (**Supplementary Table 16**).

Metagenomic sequencing and profiling

High-quality whole metagenomic sequencing was performed for 1,539 samples from discovery cohort and 1,004 samples from replication cohort with fecal samples available. The reads were aligned to hg38 using SOAP2[65] (version 2.22; identity ≥ 0.9) to remove human reads. The gene profiles were generated by aligning high-quality

sequencing reads to the integrated gene catalog (IGC) by using SOAP2 (identity ≥ 0.95) as previously described[53]. The relative abundance profiles of phylum, order, family, class, genera and species were determined from the gene abundances. To eliminate the influence of sequencing depth in comparative analyses, we downsized the unique IGC mapped reads to 20 million for each sample. The relative abundance profiles of gene, phylum, order, family, class, genus and species were determined accordingly using the downsized mapped reads per sample.

GMMs (gut metabolic modules) reflect bacterial and archaeal metabolism specific to the human gut, with a focus on anaerobic fermentation processes[66]. The current set of GMMs was built through an extensive review of the literature and metabolic databases, inclusive of MetaCyc[67] and KEGG, followed by expert curation and delineation of modules and alternative pathways. Finally, we identified 620 common microbial taxa and GMMs present in 50% or more of the samples.

Metabolic traits profiling

Measurements of metabolic traits (anthropometric characteristics and blood metabolites) were performed for all the 3,432 individuals during the physical examination in this study. The clinical tests, including blood tests and urinalysis, were performed in licensed physical examination organization. The anthropometric measurements such as height, weight, waistline and hipline were measured by nurses. Age and gender were self-reported. The metabolites, i.e. vitamins, hormones, amino acids and trace elements including heavy metals, were chosen from a health management perspective. Measurements of blood metabolites were performed as described in detail by Jie et al[39], blood amino acids were measured by ultra high pressure liquid chromatography (UHPLC) coupled to an AB Sciex Qtrap 5500 mass spectrometry (AB Sciex, US) with the electrospray ionization (ESI) source in positive ion mode using 40 µl plasma; blood hormones were measured by UHPLC coupled to an AB Sciex Qtrap 5500 mass spectrometry (AB Sciex, US) with the atmospheric pressure chemical ionization (APCI) source in positive ion mode using 250 µl plasma; blood trace elements were measured by an Agilent 7700x ICP-MS (Agilent Technologies, Tokyo, Japan) equipped with an octupole reaction system (ORS) collision/reaction cell technology to minimize spectral interferences using 200 µl whole blood; Water-soluble vitamins were measured by UPLC coupled to a Waters Xevo TQ-S Triple Quad mass spectrometry (Waters, USA) with the electrospray ionization (ESI) source in positive ion mode using 200 µl plasma; Fat-soluble vitamins were measured by UPLC coupled to an AB Sciex Qtrap 4500 mass spectrometry (AB Sciex, USA) with the atmospheric pressure chemical ionization (APCI) source in positive ion mode using 250 µl plasma.

Observational correlation of microbial features with metabolic traits

As many microbial features (taxonomies and pathways) are highly correlated, we first performed a number of Spearman correlation tests and kept only one member of pairs of bacteria or GMMs showing >0.99 correlation coefficient. This filtering resulted in a final set of 500 unique features (99 GMMs and 401 gut taxa) that were used for

analyses. We correlated these 500 microbial features with 112 measured metabolic traits, including 9 anthropometric measurements (BMI, WHR, etc.) and 103 blood metabolites (amino acids, vitamins, microelements, etc.) in the 3,432 individuals. All metabolic traits and microbial features were transformed using natural logarithmic function to reduce skewness of distributions. For each phenotype, we excluded outlier individuals with more than four standard deviations away from the mean. The metabolite measures were then centered and scaled to mean of 0 and standard deviation of 1.

The relationship between metabolic traits and microbial features were evaluated by multivariable linear regression analysis while adjusted for age and gender. After achieving the raw *P* value, we used the *p.adjust()* function in R (v3.2.5)) to perform the multiple test correction and calculated adjusted *P* values with the Benjamini–Hochberg procedure. The results were considered significant when FDR adjusted *P* value was <0.05. The correlated microbial features and metabolic traits, raw *P* and FDR adjusted *P* values, were included in the **Supplementary Table 9**.

Clustering of microbiome-metabolites associations

To assess the association clusters of 58 identified causal relationships involving the effects of 12 microbial features on 8 metabolic traits and the effects of 7 metabolic traits on 33 microbial features, we performed a hierarchical clustering analysis. Beta coefficients of associations between the microbial features and metabolic traits from one-sample MR analysis were used to construct distance matrices. Complete-linkage hierarchical clustering was used to cluster the metabolites and microbiome traits from the distance matrices using the 'hclust' function in R, and the results were visualized as a heatmap.

Genome-wide Association analysis for microbial features

We tested the associations between host genetics and gut bacteria using linear or logistic model based on the abundance of gut bacteria. The abundance of bacteria with occurrence rate over 95% in the cohort was transformed by the natural logarithm and the outlier individual who was located away from the mean by more than four standard deviations was removed, so that the abundance of bacteria could be treated as a quantitative trait. Otherwise, we dichotomized bacteria into presence/absence patterns to prevent zero inflation, then the abundance of bacteria could be treated as a dichotomous trait. Next, for 10 million common and low-frequency variants (0.5% ≤ MAF < 5%) identified in the discovery cohort and 5.9 million common and low-frequency variants identified in replication cohort, we performed a standard single variant (SNP/INDEL)-based M-GWAS analysis via PLINK using a linear model for quantitative trait or a logistic model for dichotomous trait. Given the effects of diet and lifestyles on microbial features, we included age, gender, BMI, defecation frequency, stool form, 12 diet and lifestyle factors, as well as the top four principal components (PCs) as covariates for M-GWAS analysis in both the discovery and the replication cohort.

Genome-wide Association analysis for anthropometric and metabolic traits

For each of the 112 anthropometric and metabolic traits, the log10-transformed of the median-normalized values was used as a quantitative trait. Samples with missing values and values beyond 4 s.d. from the mean were excluded from association analysis. Each of the 10 million common and low-frequency variants identified in the discovery cohort and the 5.9 million common and low-frequency variants identified in the replication cohort was tested independently using a linear model for quantitative trait implemented in PLINK. Age, gender and the top four PCs were included as covariates.

Independent predictor and explained phenotypic variance

For each whole-genome wide association result of microbial features and metabolic traits, we first selected genetic variants that showed association at $P < 1 \times 10^{-5}$ and then performed the linkage disequilibrium (LD) estimation with a threshold of LD $r^2 < 0.1$ for clumping analysis to get independent genetic predictors. The P-value threshold of 1×10^{-5} was used for selection of genetic predictors associated with microbial features by maximizing the strength of genetic instruments and the amount of the average genetic variance explained by the genetic predictors in an independent sample. For each microbial feature, we got genetic instruments in discovery cohort using different *P* thresholds, including 5×10^{-8}, 1×10^{-7}, 1×10^{-6} and 1×10^{-5}. We tested the strength of these instruments under different *P* thresholds by checking whether they predicted corresponding microbial features in an independent sample (**Supplementary Table 10** and **Supplementary Fig. 6**), we observed that the mean value of instrumental F statistics is 3.57 and on average only 0.28% phenotype variance could be explained by instruments on microbial features when using 5×10^{-8} as instrumental cut-off. Therefore, we used a more liberal threshold of $P < 1 \times 10^{-5}$ to select the instruments for microbial features, and the instrumental mean F statistics reached 51.4 (greater than 10) that indicates a strong instrument. The average phenotypic variance explained by instruments on microbial features was 22.6% for the discovery cohort and 5.09% for the replication cohort (**Supplementary Fig. 2**). For consistency, we used the same threshold and procedure for selecting genetic predictors of metabolic traits in both the discovery and the replication cohort. The LD estimation between variants was calculated in 2,002 samples for the discovery cohort and in 1,430 samples for the replication cohort, respectively. For each phenotype, the variance explained by the corresponding independent genetic predictors was estimated using a restricted the maximum likelihood (REML) model as implemented in the GCTA software[68]. We adjusted for age, gender and the top four PCs in the REML analysis.

One-sample MR analysis

To investigate the causal effects between microbial features and metabolic traits available from the same cohort, we first performed one-sample bidirectional MR analysis in discovery cohort, which included 1,539 individuals with both metabolites and microbiome traits. We specified a threshold of $P < 1 \times 10^{-5}$ to select SNP

instruments and LD r^2 < 0.1 threshold for clumping analysis to get independent genetic variants for MR analysis. Then, an unweighted polygenic risk score (PRS) was calculated for each individual using independent genetic variants from GWAS data. Each SNP was recoded as 0, 1 and 2, depending on the number of trait-specific risk increasing alleles carried by an individual. We performed Instrumental variable (IV) analyses employing two-stage least square regression (TSLS) method. In the first stage, for each exposure trait, association between the GRS and observational phenotype value was assessed using linear regression and predicted fitted values based on the instrument were obtained. In the second stage, linear regression was performed with outcome trait and genetically predicted exposure level from the first stage. In both stages, analyses were adjusted for age, gender and top four principal components of population structure. For each trait, TSLS was performed using 'ivreg' command from the AER package in R. We attempted to replicate the causal effects between traits in replication cohort with 1,004 individuals.

Two-sample MR analysis

To maximize the sample size in MR analysis and confirmed the causal effect between microbial features and metabolic traits, we also performed two-sample bidirectional MR analysis using six different methods, including genome-wide complex trait analysis-generalized summary Mendelian randomization (GCTA-GSMR) approach[69] and the other five methods implemented in "TwoSampleMR" R package as a robust validation. A consistent effect across the six methods is less likely to be a false positive. If the genetic variants have horizontally pleiotropic effects but are independent of the effects of the genetic variants on the exposure, this is known as balanced pleiotropy. If all the pleiotropic effects are biasing the estimate in the same direction (directional pleiotropy), this will bias the results (with the exception of the MR-Egger method). We used the MR-PRESSO (mendelian randomization pleiotropy residual sum and outlier) Global test to estimate for the presence of directional pleiotropy.

We first performed GWAS analysis for every trait and used summary statistics data for MR analysis. Genetic variants with $P < 1 \times 10^{-5}$ and LD r^2<0.1 were selected as instrumental variables.

The six two-sample MR methods were described as following:

GCTA-GSMR. GSMR tackled pleiotropy using HEIDI test which assumes that most SNPs are not strongly affected by horizontal pleiotropy and attempt to control SNP-heterogeneity by removing SNP-outliers. The p-value default threshold of 0.01 was specified for the HEIDI-outlier analysis to remove horizontal pleiotropic SNPs. After pruning for LD by a clump analysis and filtered for horizontal pleiotropy by the HEIDI-outlier analysis, we got the final independent predictors required for the GSMR analysis.

Inverse-variance weighting (IVW). The simplest way to obtain a MR estimate using multiple SNPs is to perform an inverse variance weighted (IVW) meta-analysis of each Wald ratio[70,71], effectively treating each SNP as a valid natural experiment. We used a multiplicative random effects version of the method, which incorporates

between instrument heterogeneity in the confidence intervals (allowing each SNP to have different mean effects).

MR–Egger regression. This method was adapted from the IVW analysis by allowing a non-zero intercept, which allows the nethorizontal pleiotropic effect across all SNPs to be unbalanced, or directional[72,73]. Horizontal pleiotropy refers to the effects of the SNPs on the outcome not mediated by the exposure.

Weighted median. This method allows for consistent causal effect estimation even if the InSIDE assumption is violated, which allows stronger SNPs to contribute more towards the estimate, and can be obtained by weighting the contribution of each SNP by the inverse variance of its association with the outcome[74].

Mode-based estimate (MBE). The mode-based estimator clusters the SNPs into groups based on similarity of causal effects, and returns the causal effect estimate based on the cluster that has the largest number of SNPs[75]. This procedure allows for consistent causal effect estimation even if most instruments are invalid. The weighted mode introduces an extra element similar to IVW and the weighted median, weighting each SNP's contribution to the clustering by the inverse variance of its outcome effect.We tested **Simple mode** and **Weighted mode** method in "TwoSampleMR" R packages.

In vitro growth of *Escherichia coli* with 5-methyltetrahydrofolic acid supplementation

To directly test the interactions between *Escherichia coli* and 5-methyltetrahydrofolic acid, the anaerobic growth of a strain *Escherichia coli* AM17-9 was characterized at different concentrations of 5-methyltetrahydrofolic acid. The *Escherichia coli* AM17-9, isolated from feces of a male, was routinely grown in Luria-Bertani (LB) broth while supplementing 5-methyltetrahydrofolic acid with concentrations of 0, 1 and 2 ng/ml, respectively. The normal concentration of 5-methyltetrahydrofolic acid in human blood ranged from 4.4 ng/ml to 32.8 ng/ml. The growth of *Escherichia coli* AM17-9 was inhibited when supplementing 5-methyltetrahydrofolic acid from 0 to 2 ng/ml. The optical density at 600 nm (OD600) was measured at intervals of two hours using a microplate reader.

MR analyses for diseases in Japan Biobank

We downloaded summary statistics data for 42 diseases and 59 blood quantitative traits in 212,453 Japanese individuals[32] (http://jenger.riken.jp/en/result, **Supplementary Table 13**). More specifically, the 42 diseases encompassed a wide-range of disease categories; 13 neoplastic diseases, five cardiovascular diseases, four allergic diseases, three infectious diseases, two autoimmune diseases, one metabolic disease, and 14 uncategorized diseases. The 59 quantitative traits were comprised of common blood parameters. By combining these data and the gut microbiome GWAS summary data from discovery cohort with high-depth WGS, we performed the two-sample bidirectional MR analysis to investigate the causal effect between the exposure (40 microbial features and 12 metabolic traits that were involved in the 72 significant causal relationships (**Fig. 4**)) and the outcome (42

diseases from BioBank Japan), by applying the GSMR method and the other five MR tests as described in the previous paragraph. For consistency, genetic variants with $P < 1 \times 10^{-5}$ and LD r^2 <0.1 were also selected as instrumental variables for phenotypes in the Japan Biobank study.

DATA AVAILABILITY

All summary statistics such as associations are available as Supplementary files (Supplementary Table 2 and 6). Individual data are protected at CNGBdb (https://db.cngb.org/search/project/CNP0000794).

Acknowledgments

We sincerely thank the support provided by China National GeneBank. We thank all the volunteers for their time and for self-collecting the fecal samples using our kit.

Author contributions

H.J. and T.Z. conceived and organized this study. J.W. initiated the overall health project. X.X., H.Y. and S.Z. performed the sample collection and questionnaire collection. X.Liu, T.Z., X.T., H.L., X.Q., J.Z., R.W. and Y.H. generated and processed the whole genome data. Y.Z., X.Lin, Z.Z., H.Z., L.T., Q.W., Z.J., and L.X. generated and processed the metagenome data. X.Liu, X.T., H.Z. and L.T. performed the bioinformatic analyses. K.K. joined in the discussion. X.Liu and H.J. wrote the manuscript. All authors contributed to data and texts in this manuscript.

Declaration of interests

The authors declare no competing financial interest.

Reference

1 Wang, J. & Jia, H. Metagenome-wide association studies: fine-mining the microbiome. *Nature reviews. Microbiology* **14**, 508-522, doi:10.1038/nrmicro.2016.83 (2016).

2 Moschen, A. R. *et al.* Lipocalin 2 Protects from Inflammation and Tumorigenesis Associated with Gut Microbiota Alterations. *Cell host & microbe* **19**, 455-469, doi:10.1016/j.chom.2016.03.007 (2016).

3 Long, X. *et al.* Peptostreptococcus anaerobius promotes colorectal carcinogenesis and modulates tumour immunity. *Nature microbiology* **4**, 2319-2330,

883 doi:10.1038/s41564-019-0541-3 (2019).

884 4 Zhu, F. *et al.* Transplantation of microbiota from drug-free patients with schizophrenia
885 causes schizophrenia-like abnormal behaviors and dysregulated kynurenine
886 metabolism in mice. *Mol Psychiatry*, doi:10.1038/s41380-019-0475-4 (2019).

887 5 Liu, X. *et al.* M-GWAS for the gut microbiome in Chinese adults illuminates on
888 complex diseases. *bioRxiv* (2019).

889 6 Bonder, M. J. *et al.* The effect of host genetics on the gut microbiome. *Nature genetics*
890 **48**, 1407-1412, doi:10.1038/ng.3663 (2016).

891 7 Wang, J. *et al.* Genome-wide association analysis identifies variation in vitamin D
892 receptor and other host factors influencing the gut microbiota. *Nature genetics* **48**,
893 1396-1406, doi:10.1038/ng.3695 (2016).

894 8 Turpin, W. *et al.* Association of host genome with intestinal microbial composition in a
895 large healthy cohort. *Nature genetics* **48**, 1413-1417, doi:10.1038/ng.3693 (2016).

896 9 Blekhman, R. *et al.* Host genetic variation impacts microbiome composition across
897 human body sites. *Genome biology* **16**, 191, doi:10.1186/s13059-015-0759-1 (2015).

898 10 Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut
899 microbiota. *Nature* **555**, 210-215, doi:10.1038/nature25973 (2018).

900 11 Shin, S. Y. *et al.* An atlas of genetic influences on human blood metabolites. *Nature*
901 *genetics* **46**, 543-550, doi:10.1038/ng.2982 (2014).

902 12 Draisma, H. H. M. *et al.* Genome-wide association study identifies novel genetic
903 variants contributing to variation in blood metabolite levels. *Nature communications* **6**,
904 7208, doi:10.1038/ncomms8208 (2015).

13 Kettunen, J. *et al.* Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nature communications* **7**, 11122, doi:10.1038/ncomms11122 (2016).

14 Long, T. *et al.* Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nature genetics* **49**, 568-578, doi:10.1038/ng.3809 (2017).

15 Yousri, N. A. *et al.* Whole-exome sequencing identifies common and rare variant metabolic QTLs in a Middle Eastern population. *Nature communications* **9**, 333, doi:10.1038/s41467-017-01972-9 (2018).

16 Burgess, S., Timpson, N. J., Ebrahim, S. & Davey Smith, G. Mendelian randomization: where are we now and where are we going? *Int J Epidemiol* **44**, 379-388, doi:10.1093/ije/dyv108 (2015).

17 Yang, Q., Lin, S. L., Kwok, M. K., Leung, G. M. & Schooling, C. M. The Roles of 27 Genera of Human Gut Microbiota in Ischemic Heart Disease, Type 2 Diabetes Mellitus, and Their Risk Factors: A Mendelian Randomization Study. *Am J Epidemiol* **187**, 1916-1922, doi:10.1093/aje/kwy096 (2018).

18 Sanna, S. *et al.* Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. *Nature genetics* **51**, 600-605, doi:10.1038/s41588-019-0350-x (2019).

19 Sakaue, S. *et al.* Trans-biobank analysis with 676,000 individuals elucidates the association of polygenic risk scores of complex traits with human lifespan. *Nature Medicine*, doi:10.1038/s41591-020-0785-8 (2020).

20 Cox, A. J. *et al.* Association of SNPs in the UGT1A gene cluster with total bilirubin and mortality in the Diabetes Heart Study. *Atherosclerosis* **229**, 155-160, doi:10.1016/j.atherosclerosis.2013.04.008 (2013).

21 MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* **45**, D896-D901, doi:10.1093/nar/gkw1133 (2017).

22 Teumer, A. Common Methods for Performing Mendelian Randomization. *Front Cardiovasc Med* **5**, 51, doi:10.3389/fcvm.2018.00051 (2018).

23 Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747-753, doi:10.1038/nature08494 (2009).

24 Verbanck, M., Chen, C. Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature genetics* **50**, 693-698, doi:10.1038/s41588-018-0099-7 (2018).

25 Katano, Y. *et al.* Complete genome sequence of Oscillibacter valericigenes Sjm18-20(T) (=NBRC 101213(T)). *Stand Genomic Sci* **6**, 406-414, doi:10.4056/sigs.2826118 (2012).

26 Thingholm, L. B. *et al.* Obese Individuals with and without Type 2 Diabetes Show Different Gut Microbial Functional Capacity and Composition. *Cell host & microbe* **26**, 252-264 e210, doi:10.1016/j.chom.2019.07.004 (2019).

27 Hu, H. J. *et al.* Obesity Alters the Microbial Community Profile in Korean Adolescents. *PLoS One* **10**, e0134333, doi:10.1371/journal.pone.0134333 (2015).

28 Tims, S. *et al.* Microbiota conservation and BMI signatures in adult monozygotic twins. *ISME J* **7**, 707-717, doi:10.1038/ismej.2012.146 (2013).

29 Noack, J., Dongowski, G., Hartmann, L. & Blaut, M. The human gut bacteria Bacteroides thetaiotaomicron and Fusobacterium varium produce putrescine and spermidine in cecum of pectin-fed gnotobiotic rats. *J Nutr* **130**, 1225-1231, doi:10.1093/jn/130.5.1225 (2000).

30 Luis, A. S. *et al.* Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides. *Nature microbiology* **3**, 210-219, doi:10.1038/s41564-017-0079-1 (2018).

31 PeBenito, A. *et al.* Comparative prevalence of Oxalobacter formigenes in three human populations. *Sci Rep* **9**, 574, doi:10.1038/s41598-018-36670-z (2019).

32 Ishigaki, K. *et al.* Large scale genome-wide association study in a Japanese population identified 45 novel susceptibility loci for 22 diseases. *bioRxiv*, 795948, doi:10.1101/795948 (2019).

33 Huo, Y. *et al.* Efficacy of folic acid therapy in primary prevention of stroke among adults with hypertension in China: the CSPPT randomized clinical trial. *JAMA* **313**, 1325-1335, doi:10.1001/jama.2015.2274 (2015).

34 Jie, Z. *et al.* The gut microbiome in atherosclerotic cardiovascular disease. *Nature communications* **8**, 845, doi:10.1038/s41467-017-00900-1 (2017).

35 Yachida, S. *et al.* Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nature Medicine* **25**, 968-976, doi:10.1038/s41591-019-0458-7 (2019).

971 36 Liu, R. *et al.* Gut microbiome and serum metabolome alterations in obesity and after
972 weight-loss intervention. *Nat Med* **23**, 859-868, doi:10.1038/nm.4358 (2017).
973 37 Harris, C. L. Cysteine and growth inhibition of Escherichia coli: threonine deaminase
974 as the target enzyme. *J Bacteriol* **145**, 1031-1035 (1981).
975 38 Rayman, M. P. Selenium and human health. *Lancet* **379**, 1256-1268,
976 doi:10.1016/S0140-6736(11)61452-9 (2012).
977 39 MacFarquhar, J. K. *et al.* Acute selenium toxicity associated with a dietary supplement.
978 *Arch Intern Med* **170**, 256-261, doi:10.1001/archinternmed.2009.495 (2010).
979 40 Shrift, A. & Kelly, E. Adaptation of Escherichia coli to selenate. *Nature* **195**, 732-733,
980 doi:10.1038/195732a0 (1962).
981 41 Huber, R. E., Segel, I. H. & Criddle, R. S. Growth of Escherichia coli on selenate.
982 *Biochim Biophys Acta* **141**, 573-586, doi:10.1016/0304-4165(67)90186-9 (1967).
983 42 Guymer, D., Maillard, J. & Sargent, F. A genetic analysis of in vivo selenate reduction
984 by Salmonella enterica serovar Typhimurium LT2 and Escherichia coli K12. *Arch*
985 *Microbiol* **191**, 519-528, doi:10.1007/s00203-009-0478-7 (2009).
986 43 Yee, N. *et al.* Selenate reductase activity in Escherichia coli requires Isc iron-sulfur
987 cluster biosynthesis genes. *FEMS Microbiol Lett* **361**, 138-143,
988 doi:10.1111/1574-6968.12623 (2014).
989 44 Lalucat, J., Bennasar, A., Bosch, R., Garcia-Valdes, E. & Palleroni, N. J. Biology of
990 Pseudomonas stutzeri. *Microbiol Mol Biol Rev* **70**, 510-547,
991 doi:10.1128/MMBR.00047-05 (2006).
992 45 Lipuma, J. J. The changing microbial epidemiology in cystic fibrosis. *Clin Microbiol*

Rev **23**, 299-323, doi:10.1128/CMR.00068-09 (2010).

46 Reese, A. T. *et al.* Microbial nitrogen limitation in the mammalian large intestine. *Nature microbiology* **3**, 1441-1450, doi:10.1038/s41564-018-0267-7 (2018).

47 Petrus, P. *et al.* Glutamine Links Obesity to Inflammation in Human White Adipose Tissue. *Cell Metab*, doi:10.1016/j.cmet.2019.11.019 (2019).

48 Jie, Z. *et al.* A multi-omic cohort as a reference point for promoting a healthy human gut microbiome. *bioRxiv*, 585893, doi:10.1101/585893 (2019).

49 Choi, W. M. *et al.* Glutamate Signaling in Hepatic Stellate Cells Drives Alcoholic Steatosis. *Cell Metab* **30**, 877-889 e877, doi:10.1016/j.cmet.2019.08.001 (2019).

50 Kang, D. J. *et al.* Gut microbiota drive the development of neuroinflammatory response in cirrhosis in mice. *Hepatology* **64**, 1232-1248, doi:10.1002/hep.28696 (2016).

51 Long, T. *et al.* Plasma metals and cardiovascular disease in patients with type 2 diabetes. *Environ Int* **129**, 497-506, doi:10.1016/j.envint.2019.05.038 (2019).

52 Kuo, C. F., Grainge, M. J., Zhang, W. & Doherty, M. Global epidemiology of gout: prevalence, incidence and risk factors. *Nat Rev Rheumatol* **11**, 649-662, doi:10.1038/nrrheum.2015.91 (2015).

53 Li, J. *et al.* An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* **32**, 834-841, doi:10.1038/nbt.2942 (2014).

54 Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65, doi:10.1038/nature08821 (2010).

55 Human Microbiome Project, C. Structure, function and diversity of the healthy human

microbiome. *Nature* **486**, 207-214, doi:10.1038/nature11234 (2012).

56 Maier, L. *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**, 623-628, doi:10.1038/nature25979 (2018).

57 Polderman, T. J. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature genetics* **47**, 702-709, doi:10.1038/ng.3285 (2015).

58 Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55-60, doi:10.1038/nature11450 (2012).

59 Fang, C. *et al.* Assessment of the cPAS-based BGISEQ-500 platform for metagenomic sequencing. *GigaScience* **7**, 1-8, doi:10.1093/gigascience/gix133 (2018).

60 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

61 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

62 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303, doi:10.1101/gr.107524.110 (2010).

63 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).

64 Browning, B. L., Zhou, Y. & Browning, S. R. A One-Penny Imputed Genome from Next-Generation Reference Panels. *Am J Hum Genet* **103**, 338-348, doi:10.1016/j.ajhg.2018.07.015 (2018).

1037 65 Li, R. *et al.* SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics*
1038 **25**, 1966-1967, doi:10.1093/bioinformatics/btp336 (2009).
1039 66 Vieira-Silva, S. *et al.* Species-function relationships shape ecological properties of the
1040 human gut microbiome. *Nature microbiology* **1**, 16088,
1041 doi:10.1038/nmicrobiol.2016.88 (2016).
1042 67 Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the
1043 BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* **42**, D459-471,
1044 doi:10.1093/nar/gkt1103 (2014).
1045 68 Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide
1046 complex trait analysis. *Am J Hum Genet* **88**, 76-82, doi:10.1016/j.ajhg.2010.11.011
1047 (2011).
1048 69 Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred
1049 from GWAS summary data. *Nature communications* **9**, 224,
1050 doi:10.1038/s41467-017-02317-2 (2018).
1051 70 Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis
1052 with multiple genetic variants using summarized data. *Genet Epidemiol* **37**, 658-665,
1053 doi:10.1002/gepi.21758 (2013).
1054 71 Bowden, J. *et al.* A framework for the investigation of pleiotropy in two-sample
1055 summary data Mendelian randomization. *Stat Med* **36**, 1783-1802,
1056 doi:10.1002/sim.7221 (2017).
1057 72 Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid
1058 instruments: effect estimation and bias detection through Egger regression. *Int J*

Epidemiol **44**, 512-525, doi:10.1093/ije/dyv080 (2015).

73 Bowden, J. *et al.* Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. *Int J Epidemiol* **45**, 1961-1974, doi:10.1093/ije/dyw220 (2016).

74 Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* **40**, 304-314, doi:10.1002/gepi.21965 (2016).

75 Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* **46**, 1985-1998, doi:10.1093/ije/dyx102 (2017).

Figure legends

Figure 1. The design and workflow of this study. The schematic representation of our study highlights, for each step, the research question that we sought to answer, the analysis workflow, the used data and the generalized result. We first performed metagenome and metabolome GWAS to detect genetic variants associated with microbial features and metabolic traits, respectively, both in discovery and replication cohorts (Step 1). We then performed observational analysis to identify which microbial feature (taxa, GMM) correlated with metabolic traits in this cohort (Step 2). We used 2,545 samples with information of both microbial features and metabolic traits; We observed 457 significant associations between 500 unique microbial features and 112 anthropometric and blood metabolic traits at a FDR adjusted $P < 0.05$. We then estimated causal relationships for the 457 observational associations through bidirectional MR analysis in discovery cohort (Step 3). One-sample BMR detected 58 causal associations between microbial features and blood metabolites after multiple test correction ($P < 1.09 \times 10^{-4}$); two-sample BMR detected the same associations and an additional 14 associations. As a validation, we replicated the discovered causal relationships by using the same MR analysis in an independent replication cohort (Step 4). Over half (43) of the 58 causal associations were replicated in the same direction ($P < 0.05$). Finally, we used two-sample MR analysis to investigate the effects of the identified 72 causal relationships on diseases from Japan Biobank study (Step 5).

Figure 2. Independent genetic variants and their explained variance of microbial features. (a) The density plot showed the distribution of number of independent genetic variants for 500 unique microbial features at $P < 10^{-5}$. The X-axis indicates the number of independent genetic variants for each microbial feature (taxon or GMM). The Y-axis indicates the number of microbial features under a given number of independent predictors. **(b)** Variance explained by the corresponding independent genetic variants for each microbial feature was shown. The polar bar plot indicates how much the independent genetic variants of each common genus (appeared at least 50% of samples) explained for their phenotypic variance (relative abundance of each genus). Genera were classified according to their respective phylum which were marked with different colors. The h^2 was calculated using REML method in GCTA tools.

Figure 3. Independent genetic variants and their explained variance of metabolic traits. (a) The density plot showed the distribution of number of independent genetic variants for 112 metabolic traits at $P < 10^{-5}$. The X-axis indicates the number of independent genetic variants for each metabolic trait. The Y-axis indicates the number of metabolic traits under a given number of independent predictors. **(b)** Variance explained by the corresponding independent genetic variants for each metabolic trait was shown. The polar bar plot indicates how much the independent genetic variants of each metabolic trait explained for their phenotypic

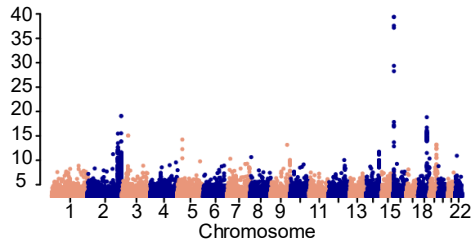
variance. Each metabolic trait was classified into different catalogs which were marked with different colors. The h^2 was calculated using REML method in GCTA tools.

Figure 4. Identifying 58 causal relationships for the microbial features and metabolic traits. (a) showed the causal effects of 12 specific microbial features on 8 metabolic traits involved in the 17 causal associations from gut microbiome to blood metabolites. **(b)** showed the causal effects of 7 blood metabolites on 33 microbial features involved in 41 causal associations from blood metabolites to gut microbiome. The cells marked with "**" represented 43 of the 58 associations that identified in discovery cohort were also replicated in replication cohort, while "*" represented the other 15 only significant in discovery cohort. The cell was colored according to the beta coefficients from one-sample MR analysis, with red and blue corresponding to positive and negative associations, respectively.

Figure 5. Causal effects of genus *Oscillibacter* and *Alistipes* on decreasing blood triglyceride concentration. (a) Schematic representation of the MR analysis results: genetic predisposition to higher abundance of *Oscillibacter* is associated with decreased blood triglyceride concentration, to a lesser extent for lowering body mass index (BMI) and waist-hip ratio (WHR). **(b)** Forest plot represented the effect of per 1-s.d. increase in *Oscillibacter* abundance on blood triglyceride, BMI and WHR, as estimated using observational and Mendelian randomization (MR) analysis, respectively. Observational correlation analysis was performed in a total of 2,545 samples (purple). One-sample MR analysis was carried out by using a PGS constructed by up to 134 genetic predictors as an instrumental variable, as estimated in discovery cohort (blue) and replication cohort (red), respectively. Corresponding *P* values from both the observational and MR analysis were shown. CI, confidence interval. **(c-d)** Forest plots represented the MR estimates and 95% CI values of the causal effects of *Oscillibacter* **(c)** and *Alistipes* **(d)** on triglyceride level, respectively. The MR analyses were performed using an one-sample MR and six different two-sample MR methods both in discovery cohort (blue) and replication cohort (red), respectively.

Figure 6. Causal effects of *Proteobacteria* and *Escherichia coli* on diseases. Forest plots represented the MR estimates and 95% CI values of the causal effects of *Proteobacteria* **(a)** and *Escherichia coli* **(b)** on diseases. The diseases' summary statistics data was from Japan Biobank study. The gut microbiome GWAS summary data from this discovery cohort with high-depth WGS was used. Six different two-sample MR approaches were used. GSMR, generalized summary Mendelian randomization implemented in GCTA toolbox. IVW, inverse variance weighted. The corresponding *P* values and β values were shown.

Step 1. What are the genetic variants associated with metabolic traits and microbial features?

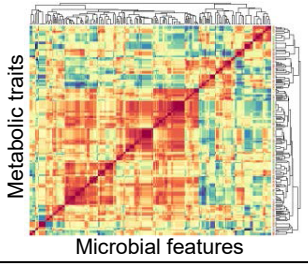


Discovery cohort: 2,002 samples with anthropometric and blood metabolites and high-depth whole genome, 1,539 of which with gut metagenome; 10M common and low-frequency variants (MAF > 0.005; HWE $P > 10^{-5}$; variants calling rate > 0.98)

Replication cohort: 1,430 samples with anthropometric and blood metabolites and low-depth whole genome, 1,006 of which with gut metagenome; 5.9M common and low-frequency variants (MAF > 0.005; HWE $P > 10^{-5}$; variants calling rate > 0.98)

Metagenome and metabolome GWAS detected genetic variants associated with microbial features and metabolic traits, respectively.

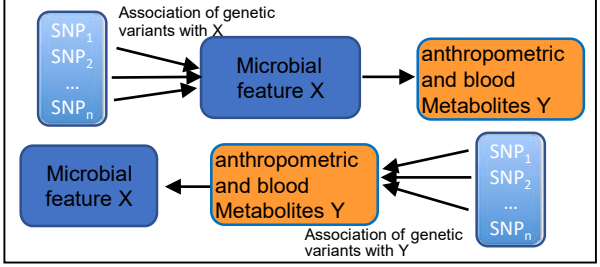
Step 2. Which microbial features correlate with metabolic traits?



A total of 3,432 samples with anthropometric and blood metabolic traits and whole genome data, of which 2,545 samples with gut metagenome

Identifying 457 observationally significant associations between 500 unique microbial features and 112 anthropometric and blood metabolites at a FDR adjusted $P < 0.05$ using a multivariable linear regression analysis

Step 3. Do changes in microbial features causally affect metabolic traits or vice versa?

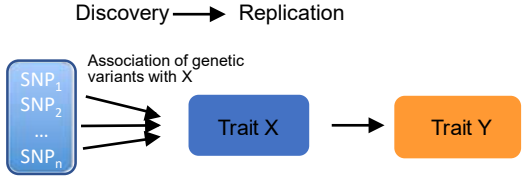


One sample BMR detected 58 causal associations between microbial features and blood metabolites after multiple test correction ($P < 1.09 \times 10^{-4}$); two sample BMR detected the same associations and an additional 14 associations

17 causal associations from gut microbiome to blood metabolites involving12 microbial features and 8 blood metabolites
Two taxa, genus *Oscillibacter* and *Alistipes* decreased triglyceride level.

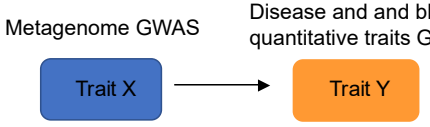
41 causal associations from blood metabolites to gut microbiome involving 7 blood metabolites and 33 microbial features

Step 4. Can we replicate causal relationships?



43 of the 58 causal associations were replicated in the same directions and $P < 0.05$ in an independent sample of replication cohort, for example genus *Oscillibacter* and *Alistipes* consistently decreased triglyceride level.

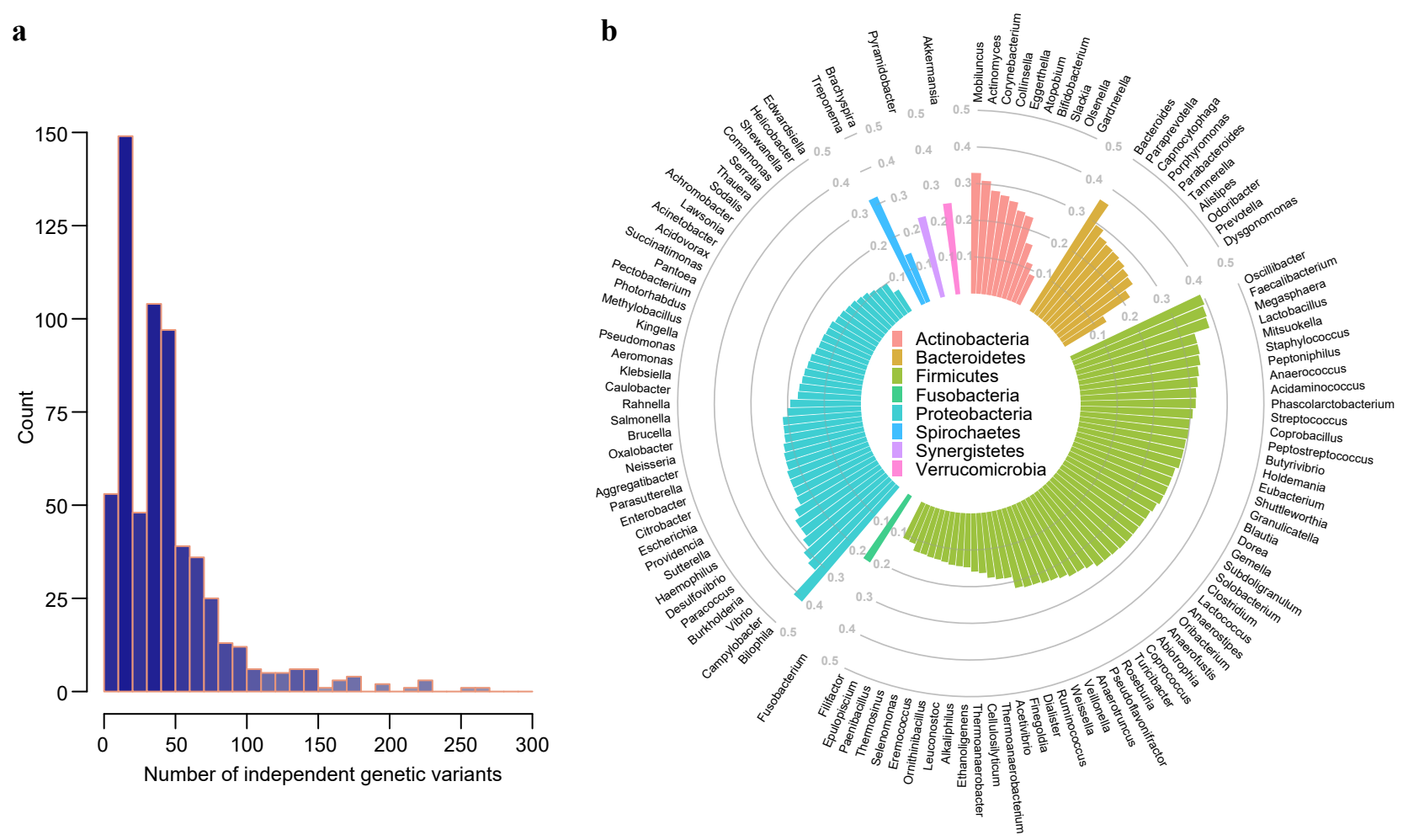
Step 5. Whether the identified causal relationships have potential link to disease?

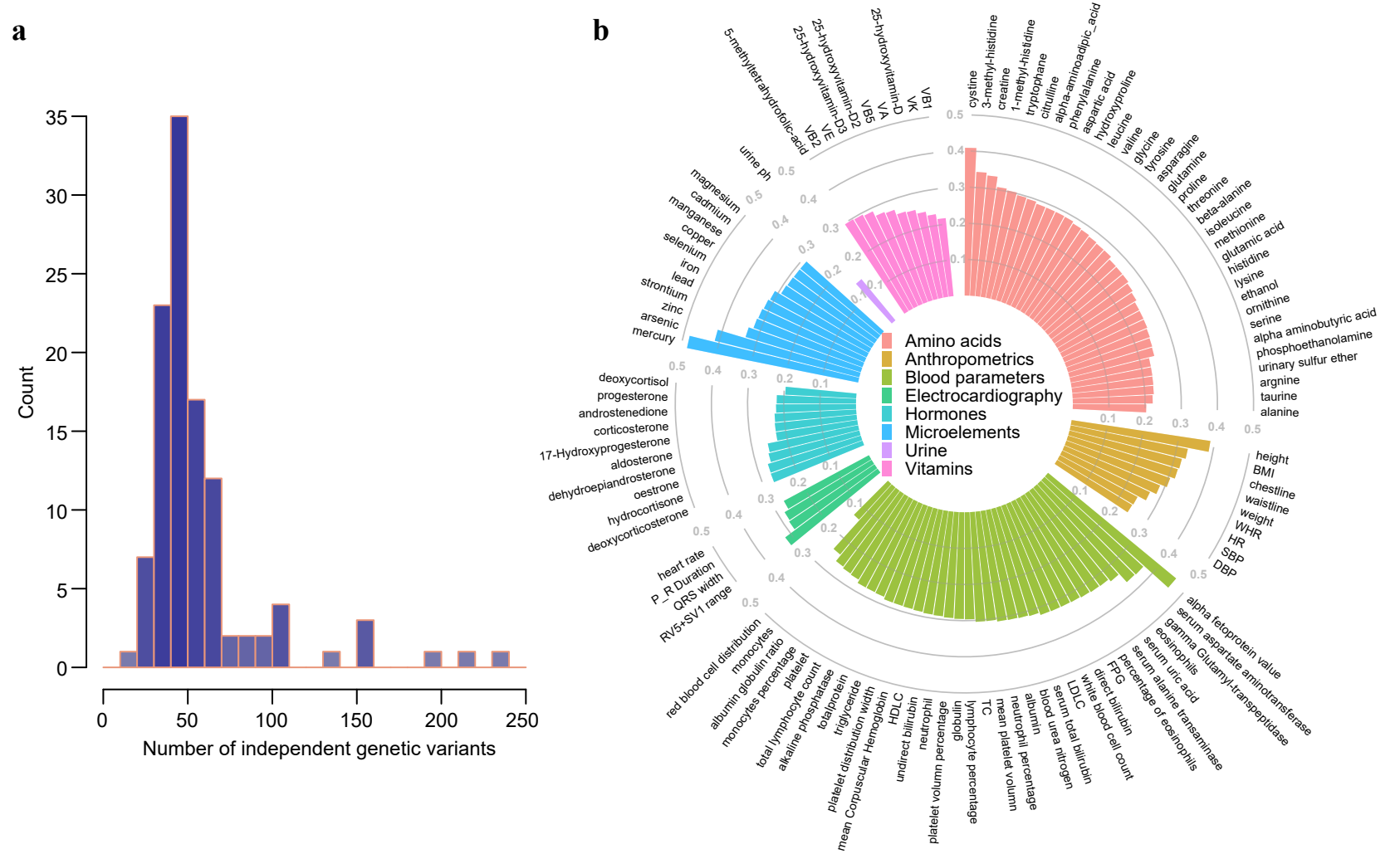


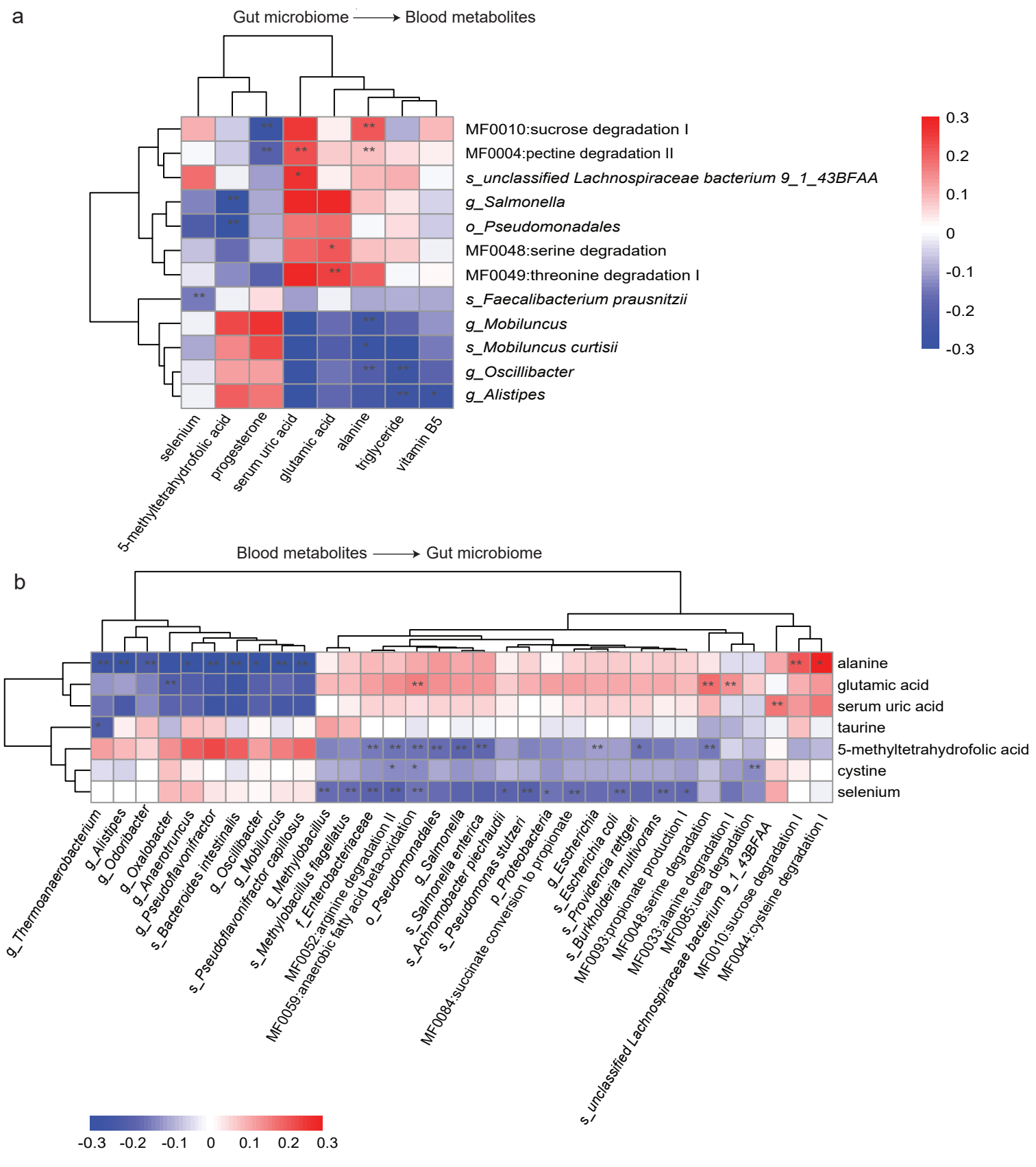
Two sample MR using metagenome GWAS summary data in this study, together with disease and blood quantitative traits GWAS summary data from Biobank Japan study

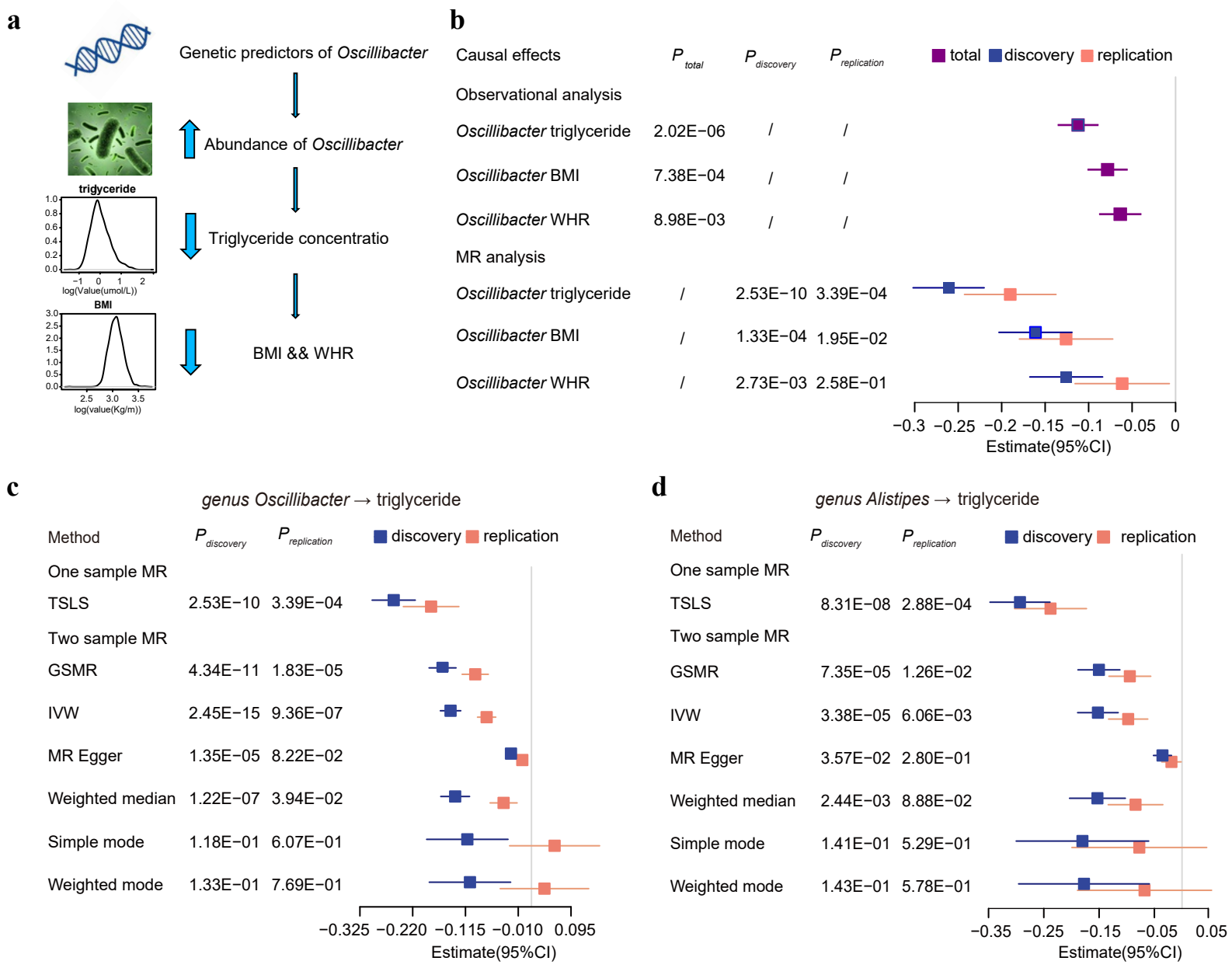
We replicate the causal effect of uric acid on increased species unclassified Lachnospiraceae bacterium 9_1_43BFAA abundance

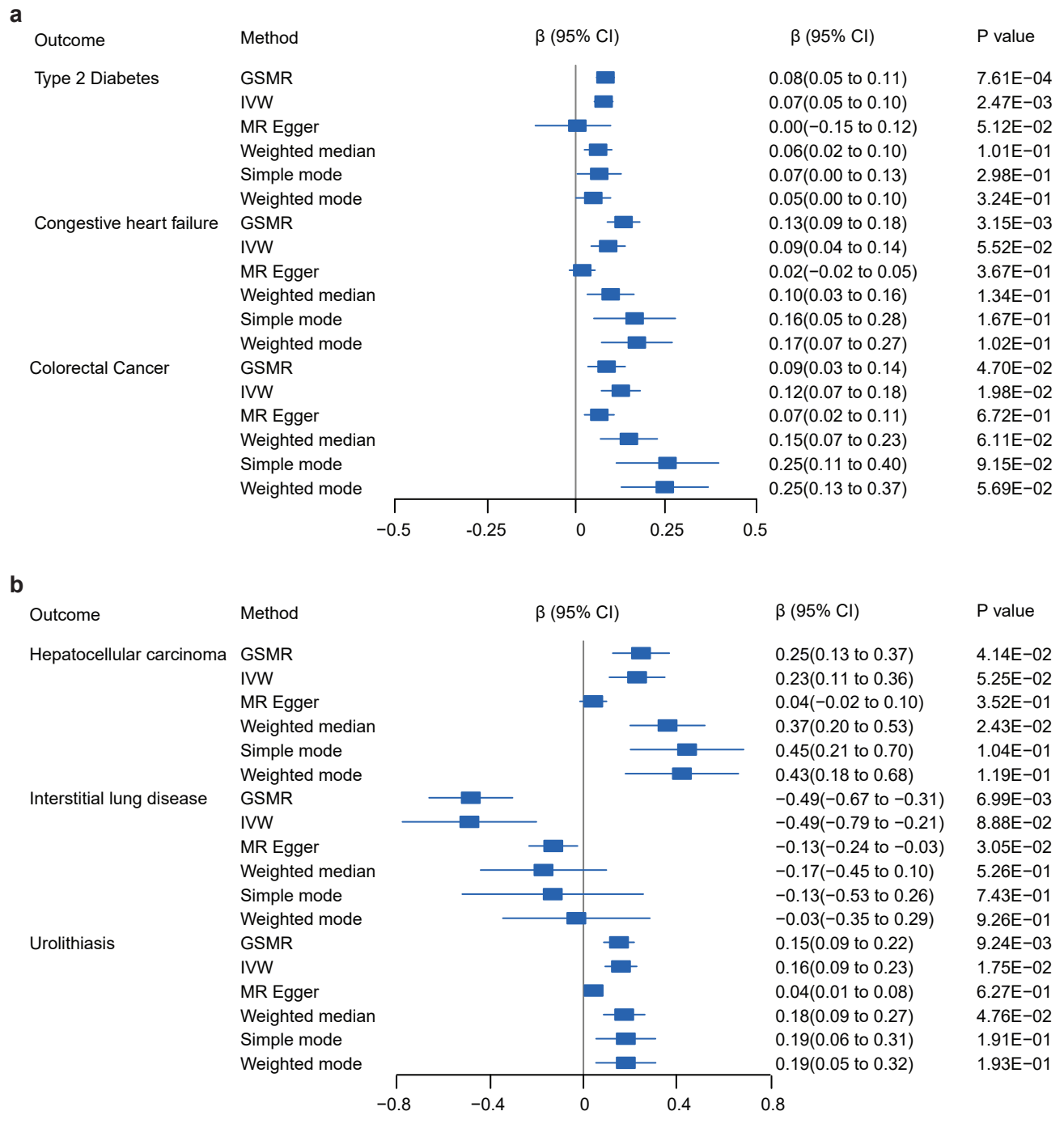
Some gut bacteria potentially linked to disease risk


a
Count
Number of independent genetic variants
b
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Proteobacteria
Spirochaetes
Synergistetes
Verrucomicrobia


a
Count
Number of independent genetic variants
0
5
10
15
20
25
30
35
0
50
100
150
200
250
b
Amino acids
Anthropometrics
Blood parameters
Electrocardiography
Hormones
Microelements
Urine
Vitamins
cystine
3-methyl-histidine
creatine
1-methyl-histidine
tryptophane
citrulline
alpha-aminoadipic_acid
phenylalanine
asparic acid
hydroxyproline
leucine
valine
glycine
tyrosine
asparagine
glutamine
proline
threonine
beta-alanine
isoleucine
methionine
glutamic acid
histidine
lysine
ethanol
ornithine
serine
alpha aminobutyric acid
phosphoethanolamine
urinary sulfur ether
argnine
taurine
alanine
height
BMI
chestline
waistline
weight
WHR
HR
SBP
DBP
alpha fetoprotein value
serum aspartate aminotransferase
gamma Glutamyl-transpeptidase
eosinophils
serum uric acid
serum alanine transaminase
percentage of eosinophils
FPG
direct bilirubin
white blood cell count
LDLC
serum total bilirubin
blood urea nitrogen
albumin
neutrophil percentage
mean platelet volumn
TC
lymphocyte percentage
globulin
platelet volumn percentage
neutrophil
undirect bilirubin
HDLC
mean Corpuscular Hemoglobin
platelet distribution width
triglyceride
totalprotein
alkaline phosphatase
total lymphocyte count
platelet
monocytes percentage
albumin globulin ratio
monocytes
red blood cell distribution
RV5+SV1 range
QRS width
P_R Duration
heart rate
deoxycorticosterone
hydrocortisone
oestrone
dehydroepiandrosterone
aldosterone
17-Hydroxyprogesterone
corticosterone
androstenedione
progesterone
deoxycortisol
mercury
arsenic
zinc
strontium
lead
iron
selenium
copper
manganese
cadmium
magnesium
urine ph
5-methyltetrahydrofolic-acid
VB2
VE
25-hydroxyvitamin-D3
25-hydroxyvitamin-D2
VB5
VA
VK
25-hydroxyvitamin-D
VB1


a
Gut microbiome → Blood metabolites
MF0010:sucrose degradation I
MF0004:pectine degradation II
s_unclassified Lachnospiraceae bacterium 9_1_43BFAA
g_Salmonella
o_Pseudomonadales
MF0048:serine degradation
MF0049:threonine degradation I
s_Faecalibacterium prausnitzii
g_Mobiluncus
s_Mobiluncus curtisii
g_Oscillibacter
g_Alistipes
selenium
5-methyltetrahydrofolic acid
progesterone
serum uric acid
glutamic acid
alanine
triglyceride
vitamin B5
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
b
Blood metabolites → Gut microbiome
alanine
glutamic acid
serum uric acid
taurine
5-methyltetrahydrofolic acid
cystine
selenium
g_Thermoanaerobacterium
g_Alistipes
g_Odoribacter
g_Oxalobacter
g_Anaerotruncus
g_Pseudoflavonifractor
s_Bacteroides intestinalis
g_Oscillibacter
g_Mobiluncus
s_Pseudoflavonifractor capillosus
g_Methylobacillus
s_Methylobacillus flagellatus
f_Enterobacteriaceae
MF0052:arginine degradation II
MF0059:anaerobic fatty acid beta-oxidation
o_Pseudomonadales
g_Salmonella
s_Salmonella enterica
s_Achromobacter piechaudii
s_Pseudomonas stutzeri
p_Proteobacteria
MF0084:succinate conversion to propionate
g_Escherichia
s_Escherichia coli
s_Providencia rettgeri
s_Burkholderia multivorans
MF0093:propionate production I
MF0048:serine degradation I
MF0033:alanine degradation I
MF0085:urea degradation I
s_unclassified Lachnospiraceae bacterium 9_1_43BFAA
MF0010:sucrose degradation I
MF0044:cysteine degradation I
-0.3 -0.2 -0.1 0 0.1 0.2 0.3


a
Genetic predictors of Oscillibacter
Abundance of Oscillibacter
Triglyceride concentratio
BMI && WHR
triglyceride
log(Value(umol/L))
BMI
log(value(Kg/m))
b
Causal effects
P total
P discovery
P replication
total
discovery
replication
Observational analysis
Oscillibacter triglyceride 2.02E−06 / /
Oscillibacter BMI 7.38E−04 / /
Oscillibacter WHR 8.98E−03 / /
MR analysis
Oscillibacter triglyceride / 2.53E−10 3.39E−04
Oscillibacter BMI / 1.33E−04 1.95E−02
Oscillibacter WHR / 2.73E−03 2.58E−01
−0.3 −0.25 −0.2 −0.15 −0.1 −0.05 0
Estimate(95%CI)
c
genus Oscillibacter → triglyceride
Method
P discovery
P replication
discovery
replication
One sample MR
TSLS 2.53E−10 3.39E−04
Two sample MR
GSMR 4.34E−11 1.83E−05
IVW 2.45E−15 9.36E−07
MR Egger 1.35E−05 8.22E−02
Weighted median 1.22E−07 3.94E−02
Simple mode 1.18E−01 6.07E−01
Weighted mode 1.33E−01 7.69E−01
−0.325 −0.220 −0.115 −0.010 0.095
Estimate(95%CI)
d
genus Alistipes → triglyceride
Method
P discovery
P replication
discovery
replication
One sample MR
TSLS 8.31E−08 2.88E−04
Two sample MR
GSMR 7.35E−05 1.26E−02
IVW 3.38E−05 6.06E−03
MR Egger 3.57E−02 2.80E−01
Weighted median 2.44E−03 8.88E−02
Simple mode 1.41E−01 5.29E−01
Weighted mode 1.43E−01 5.78E−01
−0.35 −0.25 −0.15 −0.05 0.05
Estimate(95%CI)


a
Outcome
Method
β (95% CI)
β (95% CI)
P value
Type 2 Diabetes
GSMR
0.08(0.05 to 0.11)
7.61E−04
IVW
0.07(0.05 to 0.10)
2.47E−03
MR Egger
0.00(−0.15 to 0.12)
5.12E−02
Weighted median
0.06(0.02 to 0.10)
1.01E−01
Simple mode
0.07(0.00 to 0.13)
2.98E−01
Weighted mode
0.05(0.00 to 0.10)
3.24E−01
Congestive heart failure
GSMR
0.13(0.09 to 0.18)
3.15E−03
IVW
0.09(0.04 to 0.14)
5.52E−02
MR Egger
0.02(−0.02 to 0.05)
3.67E−01
Weighted median
0.10(0.03 to 0.16)
1.34E−01
Simple mode
0.16(0.05 to 0.28)
1.67E−01
Weighted mode
0.17(0.07 to 0.27)
1.02E−01
Colorectal Cancer
GSMR
0.09(0.03 to 0.14)
4.70E−02
IVW
0.12(0.07 to 0.18)
1.98E−02
MR Egger
0.07(0.02 to 0.11)
6.72E−01
Weighted median
0.15(0.07 to 0.23)
6.11E−02
Simple mode
0.25(0.11 to 0.40)
9.15E−02
Weighted mode
0.25(0.13 to 0.37)
5.69E−02
−0.5
-0.25
0
0.25
0.5
b
Outcome
Method
β (95% CI)
β (95% CI)
P value
Hepatocellular carcinoma
GSMR
0.25(0.13 to 0.37)
4.14E−02
IVW
0.23(0.11 to 0.36)
5.25E−02
MR Egger
0.04(−0.02 to 0.10)
3.52E−01
Weighted median
0.37(0.20 to 0.53)
2.43E−02
Simple mode
0.45(0.21 to 0.70)
1.04E−01
Weighted mode
0.43(0.18 to 0.68)
1.19E−01
Interstitial lung disease
GSMR
−0.49(−0.67 to −0.31)
6.99E−03
IVW
−0.49(−0.79 to −0.21)
8.88E−02
MR Egger
−0.13(−0.24 to −0.03)
3.05E−02
Weighted median
−0.17(−0.45 to 0.10)
5.26E−01
Simple mode
−0.13(−0.53 to 0.26)
7.43E−01
Weighted mode
−0.03(−0.35 to 0.29)
9.26E−01
Urolithiasis
GSMR
0.15(0.09 to 0.22)
9.24E−03
IVW
0.16(0.09 to 0.23)
1.75E−02
MR Egger
0.04(0.01 to 0.08)
6.27E−01
Weighted median
0.18(0.09 to 0.27)
4.76E−02
Simple mode
0.19(0.06 to 0.31)
1.91E−01
Weighted mode
0.19(0.05 to 0.32)
1.93E−01
−0.8
−0.4
0
0.4
0.8