Yang, Q., Lin, S. L., Kwok, M. K., Leung, G. M., & Schooling, C. M. (2018). 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. *American Journal of Epidemiology*, *187*(9), 1916-1922. [kwy096]. https://doi.org/10.1093/aje/kwy096

The Role of 27 Human Gut Microbiota Genera in Ischemic Heart Disease, Type 2 Diabetes Mellitus and Their Risk Factors: a Mendelian Randomization Study

Qian Yang, Shi Lin Lin, Man Ki Kwok, Gabriel M. Leung, and C. Mary Schooling

Correspondence to Dr. C. Mary Schooling, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Patrick Manson Building (North Wing), 7 Sassoon Road, Pokfulam, Hong Kong (email: cms1@hku.hk) (Phone: +852 3917 6732; Fax: +852 2855 9528)

Author affiliations: School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (Qian Yang, Shi Lin Lin, Man Ki Kwok, Gabriel M. Leung, and C. Mary Schooling); Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom (Qian Yang); Bristol Medical School: Population Health Sciences, University of Bristol, Bristol, United Kingdom (Qian Yang); and Department of Environmental, Occupational and Geospatial Health Sciences, Graduate School of Public Health and Health Policy, City University of New York, New York, NY (C. Mary Schooling)

Funding information: No specific funding was received for this work.

Conflicts of interest: None declared.

Running head: Gut Microbiota Genera and Cardiovascular Diseases

Abbreviations: BMI, body mass index; GWAS, genome-wide association studies; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment insulin resistance; IHD, ischemic heart disease; LDL-C, low-density lipoprotein cholesterol; MR, Mendelian randomization; RCTs, randomized control trials; SNPs, single nucleotide polymorphisms; T2DM, type 2 diabetes Mellitus.

Abstract

Manipulation of the gut microbiota presents a new opportunity to combat chronic diseases. Randomized controlled trials of probiotics suggest some associations with adiposity, lipids and insulin resistance, but no trials with hard outcomes have been conducted. We used separate-sample Mendelian randomization to obtain estimates of the effects of 27 gut microbiota genera on ischemic heart disease, type 2 diabetes mellitus, adiposity, lipids and insulin resistance, based on summary data from CARDIoGRAAMplusC4D and other consortiums. Among 27 genera, a 1 allele increase in single nucleotide polymorphisms related to higher Bifidobacterium was associated with lower risk of ischemic heart disease (odds ratio 0.977, 95% confidence interval (CI) 0.96, 1.00, *P*=0.04), 0.011 standard deviation lower in body mass index (95% CI -0.017, -0.005) but 0.026 standard deviation higher in low-density lipoprotein cholesterol (95% CI 0.019, 0.033), which, however, were not robust to exclusion of potential pleiotropy. We also identified *Acidaminococcus*, *Aggregatibacter*, *Anaerostipes*, *Blautia*, *Desulfovibrio*, *Dorea* and *Faecalibacterium* as nominally associated with type 2 diabetes mellitus or other risk factors. Results from our study indicate that these 8 genera should be given priority in future search relating the gut microbiome for new means to prevent and treat leading causes of global morbidity and mortality.

Keywords: gut microbiota, Mendelian randomization, ischemic heart disease, type 2 diabetes mellitus

The human intestine is increasingly understood as harboring a complex community of trillions of bacteria having symbiotic relations with their host and thereby potentially affecting risk of major non-communicable diseases. In animals and humans, a microbiota-dependent metabolite, trimethylamine-N-oxide, is a predictor of cardiovascular disease [1, 2], suggesting a potential link between the gut microbiota and cardiovascular disease. Additionally, the gut microbiota may shape host metabolism, affecting the development of type 2 diabetes mellitus (T2DM) and adiposity [3], which are important risk factors for cardiovascular disease.

Observationally some gut microbiota taxa have been associated with cardiovascular disease, its subtypes or risk factors. A small case-control study (n=128) found order *Lactobacillales* positively associated with ischemic heart disease (IHD) and phylum *Bacteroidetes* inversely associated with IHD [4]. A systematic review implicated several species/genera in T2DM, but was only based on four small heterogeneous observational studies (total n=576) [5]. A recent case-control study (n=223) observed lower *Bacteroides thetaiotaomicron* in the obese [6]. *Lactobacillus reuteri* was reported positively associated with body mass index (BMI), and *Bifidobacterium animals*, *Methanobrevibacter smithii* and *Escherichia coli* were negatively associated with BMI in 263 people (51% obese) [7]. In a cohort of 893 adults 34 taxa were associated with BMI and lipids, at a false discovery rate of 0.05 [8]. *Prevotella copri* and *Bacteroides vulgatus* were the main species associated with homeostatic model assessment insulin resistance (HOMA-IR) in 277 people without diabetes (58% obese) [9]. However, these small observational studies are difficult to interpret because they are open to confounding by socially patterned factors, such as diet, which may affect the gut microbiota and health, and to changes in the gut microbiota in response to ill-health.

Meta-analyses of small randomized controlled trials (RCTs) suggested microbiota manipulation through probiotics, usually of *Lactobacillus* or *Bifidobacterium*, had a protective effect on adiposity [10, 11] but mixed effects on high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C) [12-14], and HOMA-IR [12, 13] with high heterogeneity. No RCT of probiotics with disease end-points has been conducted. Moreover, large meta-analyses of RCTs of antibiotics testing the role of antibiotic therapy in cardiovascular disease, which almost certainly changed the gut microbiome, did not affect cardiovascular disease mortality [15, 16]. No effect of vancomycin on HOMA-IR was found in an RCT of 57 obese, pre-diabetic men [17]. However, the exact effect of the antibiotics used in these RCTs on individual gut microbiota taxa is unknown, so at most they suggest we cannot rule out a role for a specific taxon.

In the absence of definitive studies giving the causal effects of specific gut microbiota taxa on IHD, T2DM, and their risk factors, comparing risk by genetically predicted taxon abundance, i.e. Mendelian randomization (MR), provides an alternative means of assessing the role of the gut microbiota in major non-communicable diseases. Since genetic endowment is randomly allocated at conception, analogous to the randomization in RCTs, MR is less vulnerable to confounding than observational studies [18]. To our knowledge, no MR study of the gut microbiota has been conducted. We conducted a separate-sample MR study based on genome-wide association studies (GWAS) predicting 27 genera applied to large extensively genotyped case-control studies of IHD and T2DM, and cross-sectional studies of adiposity, lipids and HOMA-IR to identify agnostically genera associated with these health outcomes.

METHODS

Genetically predicted gut microbiota genera

Genetic predictors of 27 genera at genome-wide significance ($P < 5\times10^{-8}$) were obtained from all currently available GWAS of stool samples in humans [19-23]. Highly correlated single nucleotide polymorphisms (SNPs) ($r^2 \geq 0.8$) were discarded based on larger *P* with correlations taken from Ensembl [24] (1000 Genomes: phase_3 among Europeans) and SNP Annotation and Proxy Search [25] (1000 Genomes Pilot 1 catalog). If a SNP was not available for an outcome, a highly correlated proxy SNP ($r^2 \geq 0.8$) was used instead, if available. We also replaced rs892244 (*Cadherin 13 (CDH13)*), because of a discrepancy between the major allele given in the GWAS [22] and Ensembl [24], with rs8063330 (*CDH13*), which is highly correlated with rs892244 (r^2 = 0.941) and was associated the same genus ($P = 2.68\times10^{-7}$) in the same GWAS [22]. We checked the phenotypes of selected SNPs using comprehensive genotype-to-phenotype cross-references, i.e. Ensembl [24] and GWAS Catalog [26], and repeated the analysis with potentially pleiotropic SNPs (rs1446585 (*RNA, U6 small nuclear 512, pseudogene*) and rs4988235 (*minichromosome maintenance complex component 6*)) excluded. We calculated SNP-specific F-statistics as a quotient of squared SNP-genus association and its variance [27]. A mean F-statistic for each genus (predicted by uncorrelated SNPs) was approximated as an average of the corresponding quotients [27].

Genetically predicted IHD, T2DM and their risk factors

CARDIoGRAMplusC4D 1000 Genomes is a case (n=60,801)-control (n=123,504) study of IHD, extensively genotyped using the 1000 Genomes phase 1v3 training set, largely of people of European descent (77%) [28]. As sensitivity analysis, we also used CARDIoGRAMplusC4D

Metabochip, (63,746 cases and 130,681 controls) largely of European descent imputed to HapMap 2 [29], which overlaps with 1000 Genomes (57.5% cases, 40.1% controls). If SNPs were not available in CARDIoGRAMplusC4D Metabochip, genetic associations were obtained from the more extensively genotyped subset in CARDIoGRAM, (22,233 cases, 64,762 controls) of European descent [30]. All three studies were age- and sex-adjusted.

Genetic associations with T2DM, adjusted for age and sex, were from DIAbetes Genetics Replication And Meta-analysis case (n=34,380)-control (n=114,981) study [31]. Genetic associations with adiposity were from The Genetic Investigation of Anthropometric Traits with BMI and waist-hip ratio (standard deviation) for 332,154 and 210,222 people of European descent respectively, adjusted for age, age^2, and study-specific covariates [32]. Genetic associations with HDL-C and LDL-C (standard deviation), adjusted for age, age^2 and sex, were obtained from the Global Lipids Genetic Consortium Results, of up to 188,577 participants of European descent and 7,898 participants of non-European descent [33]. Genetic associations with HOMA-IR (log-transformed) were from the Meta-Analyses of Glucose and Insulin-related traits Consortium of 46,186 people of European descent [34].

Statistical analysis

Estimates of the association of each genus with IHD and its risk factors were obtained by combining SNP-specific Wald estimates [35] using inverse variance weighting with fixed effects for uncorrelated SNPs and weighted generalized linear regression, considering correlations between SNPs (Web Appendix 1) [36]. Variance of a Wald estimate was obtained from Fieller's theorem [37] or an approximation if the variance for SNP on exposure was not given [38]. When

different GWAS used incompatible microbiota units for SNPs predicting the same genera, we used SNP-outcome associations (Web Table 1) [39]. If a genus was predicted by >3 uncorrelated SNPs, MR-Egger and weighted median methods were used as sensitivity analyses. MR-Egger checks for unknown horizontal pleiotropy indicated by a non-zero intercept [40], with its "No Measurement Error" assumption tested by I^2 [27]. If I^2 was less than 90%, we performed simulation extrapolation to adjust for this violation [27]. A weighted median estimate is robust to 50% of the SNPs being invalid genetic instruments [40]. Bonferroni correction was used to adjust for multiple comparisons among genera within each outcome, giving a cutoff of 0.00185 for IHD in CARDIoGRAMplusC4D 1000 Genomes and 0.002 for the other outcomes. Given the overlap of participants between the two IHD case-control studies, we also combined their estimates accounting for this overlap using the Lin and Sullivan approach [41]. All statistical analyses were conducted using Stata version 13.1 (StataCorp LP, College Station, TX) and R version 3.2.5 (R Foundation for Statistical Computing, Vienna, Austria). This study used publicly available summary data. Therefore, no ethical approval was required.

RESULTS

Five GWAS of the gut microbiota were identified, giving 94 SNPs related to 27 gut microbiota genera at genome-wide significance. 16S rRNA gene sequencing was used in four studies [19-22] and metagenomics sequencing in one study [23]. In UK Twins (n=2,731, 11% men, age range 19 to 89 years) 13 SNPs predicted 7 genera (Box-Cox transformed relative abundance) [19]. In 1,812 people from Germany (46% men, age range 18 to 83 years) 5 SNPs predicted 4 genera in a generalized linear model with a negative binomial distribution and log link [20]. In 1,561 healthy participants of European descent (45% men, age range 6 to 35 years) 29 SNPs predicted 17

genera (log-transformed relative abundance) [21]. In 127 Hutterites (38% men, age range 6 to 92 years), of European descent, rs2630788 (*zinc finger protein 385D*) and rs892244 (*CDH13*) predicted *Anaerostipes* and *Bifidobacterium* (normalized relative abundance) respectively [22]. Finally, in 1,514 participants (42% men, age range 18 to 84 years) from Dutch cohorts 45 SNPs predicted 5 genera (normalized abundance) [23]. We excluded 37 highly correlated SNPs. The remaining 57 SNPs from 55 genes were used in this study (Web Tables 2, 3) to predict 27 genera: *Acidaminococcus*, *Acinetobacter*, *Aggregatibacter*, *Anaerostipes*, *Atopobium*, *Bacteroides*, *Bifidobacterium*, *Blautia*, *Coprococcus*, *Desulfovibrio*, *Dialister*, *Dorea*, *Eggerthella*, *Escherichia*, *Eubacterium*, *Faecalibacterium*, *Lachnospira*, *Lactobacillus*, *Leuconostoc*, *Megamonas*, *Mogibacterium*, *Oscillibacter*, *Oscillospira*, *Pseudobutyrivibrio*, *Roseburia*, *Slackia* and *Weissella*. All available F-statistics were >10 (Web Table 2).

Bifidobacterium, based on 3 SNPs from different GWAS, was associated with lower IHD in the two CARDIoGRAMplusC4D studies combined, accounting for their overlap (Table 1, Web Figure 1c), although this association was not evident in CARDIoGRAMplusC4D 1000 Genomes (Web Figure 1a). *Bifidobacterium* was also associated with lower BMI (Table 1, Web Figure 1e), higher HDL-C (Table 1, Web Figure 1g), higher LDL-C (Table 1, Web Figure 1h), and lower HOMA-IR (Table 1, Web Figure 1i). Only the associations with BMI and LDL-C were robust to Bonferroni correction (Table 1). However, after the exclusion of pleiotropic SNPs *Bifidobacterium* was not associated with any outcome considered (Web Table 4).

We further identified 7 genera nominally associated with IHD risk factors. *Acidaminococcus*, based on 5 uncorrelated SNPs from the same GWAS, was associated with higher HDL-C (Table 1, Web Figure 1g). Sensitivity analysis using MR-Egger and weighted-median gave similar estimates (Web Table 5). *Aggregatibacter*, based on 1 SNP, was associated with higher HDL-C

(Table 1, Web Figure 1g). *Anaerostipes*, based on 2 SNPs from different GWAS, was associated with lower T2DM (Table 1, Web Figure 1d). *Blautia*, based on 6 SNPs from different SNPs, was associated with lower LDL-C (Table 1, Web Figure 1h). *Desulfovibrio*, based on 2 uncorrelated SNPs from the same GWAS, and *Dorea*, based on 1 SNP, were associated with higher HOMA-IR (Table 1, Web Figure 1). *Faecalibacterium*, based on 4 SNPs from different GWAS, was associated with lower waist-hip ratio (Table 1, Web Figure 1f). However, none of these associations were robust to Bonferroni correction (Table 1).

Additionally, *Lachnospira*, based on 1 SNP, was associated with higher IHD in CARDIoGRAMplusC4D Metabochip (Table 1, Web Figure 1b), but not in CARDIoGRAMplusC4D 1000 Genomes (Web Figure 1a), or in the two CARDIoGRAMplusC4D studies combined accounting for their overlap (Web Figure 1c). No associations were found for the other 18 genera, namely *Acinetobacter*, *Atopobium*, *Bacteroides*, *Coprococcus*, *Dialister*, *Eggerthella*, *Escherichia*, *Eubacterium*, *Lactobacillus*, *Leuconostoc*, *Megamonas*, *Mogibacterium*, *Oscillibacter*, *Oscillospira*, *Pseudobutyrivibrio*, *Roseburia*, *Slackia* and *Weissella* (Web Figure 1).

DISCUSSION

In what is to our knowledge the first MR study relating gut microbiota to IHD and its risk factors, we found some preliminary indications of beneficial associations of *Bifidobacterium* with BMI, HDL-C and HOMA-IR. We also found some nominal associations of *Acidaminococcus*, *Aggregatibacter*, *Anaerostipes*, *Blautia*, *Desulfovibrio*, *Dorea* and *Faecalibacterium* with modestly lower risk of T2DM, less adiposity, more beneficial lipid

profiles or higher HOMA-IR. Associations of the other genera considered with these outcomes appeared less likely.

Our study has some consistency with an observational study showing no robust association of genera *Bacteroides, Blautia, Coprococcus, Eggerthella* or *Lachnospira* with BMI, HDL-C or LDL-C [8], although we also found *Blautia* nominally associated with lower LDL-C. However, our study is less consistent with a small case-control study showing order *Lactobacillales* positively and phylum *Bacteroidetes* negatively associated with IHD but *Bifidobacterium* unrelated to IHD [4, 42]. In fact, observational studies of the gut microbiota are probably susceptible to unmeasured confounding, by factors such as diet and health status. Our study also has some consistency with meta-analyses of RCTs showing beneficial effects of probiotics, typically including *Bifidobacterium*, on BMI [10, 11], HDL-C [12, 13] and HOMA-IR [13], although associations with HDL-C and HOMA-IR in our study were less evident after correction for multiple comparisons. However, these meta-analyses of RCTs may be vulnerable to biases from small sample sizes (ranging from 234 to 1,931) and/or high heterogeneity (I^2 ranging from 0% to 92%) [10-13]. In addition, some RCTs included in these meta-analyses suggest a role for probiotics including *Lactobacillus* [10-12], but we found no associations for *Lactobacillus*, perhaps because the gut microbiota acts synergistically [43], so that the effect of a particular mix may be different from the effect of its constituent parts. A large well-conducted RCT in a well-characterized population using probiotics capsules containing sole species may provide further clarification. Finally, our study has some consistency with meta-analyses of RCTs showing little association of antibiotics with IHD [15, 16], because these RCTs likely changed the gut microbiota but did not affect cardiovascular disease mortality. RCTs targeting *Bifidobacterium* (or more generally

investigations of exact effects of various antibiotics on specific gut microbiota taxa) might provide further evidence for IHD prevention.

Many potential pathways linking specific gut microbiota to non-communicable diseases have been identified. A possible pathway linking gut microbiota to IHD is via dietary choline (from shrimps and eggs) or dietary carnitine (from meat) to trimethylamine and trimethylamine-N-oxide [44]. However, the role of specific taxa in trimethylamine production is not entirely clear [45] and we did not identify any genus robustly associated with IHD. Host metabolites linking gut microbiota to T2DM/metabolic syndrome may exist. Short-chain fatty acids are generated by many gut microbiota genera, such as *Anaerostipes*, *Bacteroides*, *Bifidobacterium*, *Blautia*, *Clostridium*, *Dialister*, *Prevotella*, *Roseburia*, *Salmonella* and *Streptococcus*, from fermentation of dietary fiber and may have beneficial metabolic effects for the host [46]. Meta-analysis of RCTs showed dietary fiber reduces LDL-C [47], and we further identified that *Blautia*, possibly fueled by dietary fiber [46], might provide the mechanism. Whether any beneficial effect of *Blautia* on LDL-C is mediated by short-chain fatty acids would be informed by RCTs investigating the role of *Blautia* in short-chain fatty acids production. Branched-chain amino acids have essential signaling functions, may be synthesized by *Prevotella copri* and *Bacteroides vulgatus* [9], and were positively associated with T2DM and BMI [48], but not with any marker of glucose metabolism [48, 49]. Correspondingly, we did not find *Bacteroides* associated with HOMA-IR. A recent observational study found several *Bacteroides* species inversely correlated with branched-chain amino acids [6], but the role of these species in the biosynthesis of branched-chain amino acids needs to be further confirmed in humans. Notably, lactase persistence alleles predicting lower Bifidobacterium abundance have been associated with higher milk drinking [50] and with anthropometric traits [24, 26]. Since lactose fuels *Bifidobacterium* in the human intestine [19],

Bifidobacterium may have more of an effect in populations who drink milk despite lactose intolerance. Given the role of *Bifidobacterium* is difficult to distinguish from that of lactase persistence in people of European descent, replication in a population without lactase persistence, such as East Asians, would be helpful. Bidirectional MR studies to assess whether IHD and its risk factors influence the gut microbiota might also be informative. More generally, this study raises the question as to whether the search for a healthy diet should focus on the effect of foods and their constituents on health or their many mechanisms, including the gut microbiota.

In the era of "big data", taking advantage of GWAS and large publicly available data with extensive genotyping enables a cost-efficient MR study [36]. Nevertheless, limitations regarding MR and gut microbiota exist. First, MR has stringent assumptions. Although we selected SNPs uniquely associated with 27 genera at genome-wide significance, few of them achieved study-wide significance, and thus we could not fully rule out the possibility of weak instrument bias. However, our F-statistics suggest little evidence of that [51]. A post-hoc power calculation [52] assuming a statistical confidence level of 0.05, an R^2 equaling genus heritability and an effect size shown in Table 1 suggested power of greater than 80% for the associations of Bifidobacterium with BMI and LDL-C, but less than 80% for weaker associations. As such larger MR studies are necessary, to distinguish associations with small effect sizes from null associations. In addition, some SNPs identified in one GWAS were not replicated in others due to low variance in the corresponding genera or different SNP selections. Publicly releasing all available individual GWAS, or their summary, would be helpful, as would further GWAS in larger more homogenous samples. More generally, our study did not consider associations between the 27 genera or all bacterial taxa. For example, family Bifidobacteria is inversely associated with species Escherichia coli [53]. Cross-phenotype association analysis [54] combining

GWAS may help identify more accurate genetic instruments and clarify our MR estimates, when data is available. Residual pleiotropy is difficult to exclude, as functions of most SNPs have not been comprehensively identified; use of MR-Egger and a weighted median to identify pleiotropy statistically was restricted by the limited number of genetic instruments. Confounding by population stratification is possible. However, all five GWAS concerned participants of European descent [19-23] and the genetic associations with IHD and its risk factors are all from studies conducted largely in people of European descent with genomic control [28-34]. Second, canalization may buffer the genetic effects of gut microbiota, so its manipulation might not have the same effect as that genetically predicted. However, whether the relevance of canalization exists is unknown. Third, winner's curse may bias our MR estimates, but its direction is ambiguous [51]. Finally, selection bias may influence our MR estimates, where genetic associations are obtained from studies in older people [55] or otherwise condition on genetic make-up and exposure or outcome. However, they did not condition one phenotype on another, reducing the risk of bias [56].

In terms of specific limitations of applying MR to gut microbiota, the studies used to identify genetic predictors of *Bacteroides*, *Bifidobacterium*, *Coprococcus*, *Dorea*, *Eggerthella* and *Faecalibacterium* and to identify their associations with adiposity and lipids overlapped slightly because of the participants in the TwinsUK study [57]. However, they only form a very small proportion of these studies which is unlikely to create a bias [58] and separate-sample MR reduces the risk of chance associations generated by the underlying data structure in a one-sample MR [59]. Use of separate samples also means that possible non-linear associations, subgroup analysis by age and sex, and diet-microbiome interactions could not be tested [60], but causal effects should be generally consistent. Second, the 16S rRNA gene sequencing used by most microbiota GWAS

usually only permits resolution at genus level rather than at a more specific level, so we cannot rule out the possibility that some specific species or strains are associated with IHD or its risk factors. Additionally, we cannot rule out the possibility that a ratio of two taxa or dysbiosis of gut microbiota contributes to cardiovascular disease or its risk factors as suggested by some observational studies [8, 61, 62], although the ratio of *Bacteroidetes* to *Firmicutes* is not consistently associated with adiposity in humans [63]. Fourth, gut microbiota may also be influenced by other factors, such as the time/season of stool sampling, which may decrease the variance explained by genetics. However, gut microbiota is thought to have temporal stability especially after early childhood, and the dominant force in determining its composition is long-term dietary habits [64]. As such, our findings may be more relevant to the effects of gut microbiota from adolescence or adulthood. Our study is also limited by the current understanding of the gut microbiota. A hypothesis driven study testing epidemiologically established associations would have been preferable, but was precluded by the lack of knowledge as to the function of each constituent of the microbiome and by the lack of large epidemiological studies. In addition, differences in statistical methods between gut microbiota GWAS made the units hard to interpret. As such, we presented results per allele for *Bifidobacterium*, *Blautia*, *Anaerostipes*, *Bacteroides*, *Dialister* and *Faecalibacterium*, so these estimates are best understood as providing direction and we could not completely rule in/out their causal effects on the outcomes considered [65]. Finally, our findings mainly concern Europeans. Gut microbiota may vary between populations [66], so replication in different populations are needed. Replication with functionally relevant genetic prediction of gut microbiota would also be helpful.

Our study generates the hypothesis that *Acinetobacter*, *Atopobium*, *Bacteroides*, *Coprococcus*, *Dialister*, *Eggerthella*, *Escherichia*, *Eubacterium*, *Lachnospira*, *Lactobacillus*, *Leuconostoc*,

Megamonas, *Mogibacterium*, *Oscillibacter*, *Oscillospira*, *Pseudobutyrivibrio*, *Roseburia*, *Slackia* and *Weissella* are unlikely to have a major causal association with IHD or T2DM, and so might not warrant extensive testing. Our study also raises the possibility of a beneficial association of *Bifidobacterium* with IHD, adiposity, HDL-C and HOMA-IR, as well as associations of *Acidaminococcus*, *Aggregatibacter*, *Anaerostipes*, *Blautia*, *Desulfovibrio*, *Dorea* and *Faecalibacterium* with cardiovascular disease risk factors, suggesting these might be the focus of future investigation. Further MR studies using multiple robust instruments are needed to confirm these results given our study was limited by single genetic instruments for some genera.

ACKNOWLEDGMENTS

The authors thank CARDIoGRAMplusC4D, DIAbetes Genetics Replication And Meta-analysis, The Genetic Investigation of Anthropometric Traits, Global Lipids Genetic Consortium and Meta-Analyses of Glucose and Insulin-related traits Consortium for their publicly available summary data.

References

1. Kitai T, Kirsop J, Tang WH. Exploring the microbiome in heart failure. *Curr Heart Fail Rep*. 2016;13(2):103-109.

2. Zhu W, Gregory JC, Org E, *et al*. Gut microbial metabolite TMAO enhances platelet hyperreactivity and thrombosis risk. *Cell*. 2016;165(1):111-124.

3. Bouter KE, van Raalte DH, Groen AK, *et al*. Role of the gut microbiome in the pathogenesis of obesity and obesity-related metabolic dysfunction. *Gastroenterology*. 2017;152(7):1671-1678.

4. Emoto T, Yamashita T, Sasaki N, *et al*. Analysis of gut microbiota in coronary artery disease patients: a possible link between gut microbiota and coronary artery disease. *J Atheroscler Thromb*. 2016;23(8):908-921.

5. He C, Shan Y, Song W. Targeting gut microbiota as a possible therapy for diabetes. *Nutr Res*. 2015;35(5):361-367.

6. Liu R, Hong J, Xu X, *et al*. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat Med*. 2017;23(7):859-868.

7. Million M, Angelakis E, Maraninchi M, *et al*. Correlation between body mass index and gut concentrations of Lactobacillus reuteri, Bifidobacterium animalis, Methanobrevibacter smithii and Escherichia coli. *Int J Obes (Lond)*. 2013;37(11):1460-1466.

8. Fu J, Bonder MJ, Cenit MC, *et al*. The gut microbiome contributes to a substantial proportion of the variation in blood lipids. *Circ Res*. 2015;117(9):817-824.

9. Pedersen HK, Gudmundsdottir V, Nielsen HB, *et al*. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature*. 2016;535(7612):376-381.

10. Zhang Q, Wu Y, Fei X. Effect of probiotics on body weight and body-mass index: a systematic review and meta-analysis of randomized, controlled trials. *Int J Food Sci Nutr*. 2015;67(5):571-580.

11. Sun J, Buys N. Effects of probiotics consumption on lowering lipids and CVD risk factors: a systematic review and meta-analysis of randomized controlled trials. *Ann Med*. 2015;47(6):430-440.

12. Li C, Li X, Han H, *et al*. Effect of probiotics on metabolic profiles in type 2 diabetes mellitus: A meta-analysis of randomized, controlled trials. *Medicine (Baltimore)*. 2016;95(26):e4088.

13. Hu YM, Zhou F, Yuan Y, *et al*. Effects of probiotics supplement in patients with type 2 diabetes mellitus: A meta-analysis of randomized trials. *Med Clin (Barc)*. 2017;148(8):362-370.

14. Hendijani F, Akbari V. Probiotic supplementation for management of cardiovascular risk factors in adults with type II diabetes: a systematic review and meta-analysis. *Clin Nutr*. 2018;37(2):532-541.

15. Almalki ZS, Guo JJ. Cardiovascular events and safety outcomes associated with azithromycin therapy: a meta-analysis of randomized controlled trials. *Am Health Drug Benefits*. 2014;7(6):318-328.

16. Andraws R, Berger JS, Brown DL. Effects of antibiotic therapy on outcomes of patients with coronary artery disease: a meta-analysis of randomized controlled trials. *JAMA*. 2005;293(21):2641-2647.

17. Reijnders D, Goossens GH, Hermes GD, *et al*. Effects of gut microbiota manipulation by antibiotics on host metabolism in obese humans: a randomized double-blind placebo-controlled Trial. *Cell Metab*. 2016;24(1):63-74.

18. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1-22.

19. Goodrich JK, Davenport ER, Beaumont M, *et al*. Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe*. 2016;19(5):731-743.

20. Wang J, Thingholm LB, Skieceviciene J, *et al*. Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat Genet*. 2016;48(11):1396-1406.

21. Turpin W, Espin-Garcia O, Xu W, *et al*. Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat Genet*. 2016;48(11):1413-1417.

22. Davenport ER, Cusanovich DA, Michelini K, *et al*. Genome-wide association studies of the human gut microbiota. *PLoS One*. 2015;10(11):e0140301.

23. Bonder MJ, Kurilshikov A, Tigchelaar EF, *et al*. The effect of host genetics on the gut microbiome. *Nat Genet*. 2016;48(11):1407-1412.

24. Aken BL, Achuthan P, Akanni W, *et al*. Ensembl 2017. *Nucleic Acids Res*. 2017;45(D1):D635-D642.

25. Johnson AD, Handsaker RE, Pulit SL, *et al*. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008;24(24):2938-2939.

26. MacArthur J, Bowler E, Cerezo M, *et al*. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res*. 2017;45(D1):D896-D901.

27. Bowden J, Del Greco MF, Minelli C, *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*. 2016;45(6):1961-1974.

28. Nikpay M, Goel A, Won HH, *et al*. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47(10):1121-1130.

29. Deloukas P, Kanoni S, Willenborg C, *et al*. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. 2013;45(1):25-33.

30. Schunkert H, Konig IR, Kathiresan S, *et al*. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011;43(4):333-338.

31. Morris AP, Voight BF, Teslovich TM, *et al*. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*. 2012;44(9):981-990.

32. Locke AE, Kahali B, Berndt SI, *et al*. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518(7538):197-206.

33. Willer CJ, Schmidt EM, Sengupta S, *et al*. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013;45(11):1274-1283.

34. Dupuis J, Langenberg C, Prokopenko I, *et al*. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. 2010;42(2):105-116.

35. Wald A. The fitting of straight lines if both variables are subject to error. *Stat Med*. 1940;35(11):284-300.

36. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat Med*. 2016;35(11):1880-1906.

37. Fieller EC. Some problems in intervel estimation. *J R Stat Soc Series B Stat Methodol*. 1954;16(2):175-185.

38. Burgess S, Scott RA, Timpson NJ, *et al*. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol*. 2015;30(7):543-552.

39. Katan MB. Apolipoprotein E isoforms, serum cholesterol, and cancer. 1986 [letter]. *Int J Epidemiol*. 2004;33(1):9.

40. Burgess S, Bowden J, Fall T, *et al*. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology*. 2017;28(1):30-42.

41. Lin DY, Sullivan PF. Meta-analysis of genome-wide association studies with overlapping subjects. *Am J Hum Genet*. 2009;85(6):862-872.

42. Emoto T, Yamashita T, Kobayashi T, *et al*. Characterization of gut microbiota profiles in coronary artery disease patients using data mining analysis of terminal restriction fragment length polymorphism: gut microbiota could be a diagnostic marker of coronary artery disease. *Heart Vessels*. 2017;32(1):39-46.

43. Vieira AT, Teixeira MM, Martins FS. The role of probiotics and prebiotics in inducing gut immunity. *Front Immunol*. 2013;4:445.

44. Yamashita T, Emoto T, Sasaki N, *et al*. Gut microbiota and coronary artery disease. *Int Heart J*. 2016;57(6):663-671.

45. Zhang C, Zhao L. Strain-level dissection of the contribution of the gut microbiome to human metabolic disease. *Genome Med*. 2016;8(1):41.

46. Koh A, De Vadder F, Kovatcheva-Datchary P, *et al*. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell*. 2016;165(6):1332-1345.

47. Hartley L, May MD, Loveman E, *et al*. Dietary fibre for the primary prevention of cardiovascular disease. *Cochrane Database Syst Rev*. 2016(1):CD011472.

48. Lotta LA, Scott RA, Sharp SJ, *et al*. Genetic predisposition to an impaired metabolism of the branched-chain amino acids and risk of type 2 diabetes: a Mendelian randomisation analysis. *PLoS Med*. 2016;13(11):e1002179.

49. Mahendran Y, Jonsson A, Have CT, *et al*. Genetic evidence of a causal effect of insulin resistance on branched-chain amino acid levels. *Diabetologia*. 2017;60(5):873-878.

50. Yang Q, Lin SL, Au Yeung SL, *et al*. Genetically preicted milk consumption and bone health, ischemic heart disease, and type 2 diabetes: a Mendelian randomization study. *Eur J Clin Nutr*. 2017;71(8):1008-1012.

51. Lawlor DA, Harbord RM, Sterne JA, *et al*. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8):1133-1163.

52. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol*. 2013;42(5):1497-1501.

53. Lee JH, O'Sullivan DJ. Genomic insights into bifidobacteria. *Microbiol Mol Biol Rev*. 2010;74(3):378-416.

54. Cichonska A, Rousu J, Marttinen P, *et al*. metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. *Bioinformatics*. 2016;32(13):1981-1989.

55. Zhao J, Jiang C, Lam TH, *et al*. Genetically predicted testosterone and cardiovascular risk factors in men: a Mendelian randomization analysis in the Guangzhou Biobank Cohort Study. *Int J Epidemiol*. 2014;43(1):140-148.

56. Zheng J, Baird D, Borges MC, *et al*. Recent developments in Mendelian randomization studies. *Curr Epidemiol Rep*. 2017;4(4):330-345.

57. Moayyeri A, Hammond CJ, Valdes AM, *et al*. Cohort profile: TwinsUK and healthy ageing twin study. *Int J Epidemiol*. 2013;42(1):76-85.

58. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol*. 2016;40(7):597-608.

59. Taylor AE, Davies NM, Ware JJ, *et al*. Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. *Econ Hum Biol*. 2014;13:99-106.

60. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol*. 2016;45(3):908-915.

61. Pevsner-Fischer M, Blacher E, Tatirovsky E, *et al*. The gut microbiome and hypertension. *Curr Opin Nephrol Hypertens*. 2017;26(1):1-8.

62. Yin J, Liao SX, He Y, *et al*. Dysbiosis of gut microbiota with reduced trimethylamine-N-oxide level in patients with large-artery atherosclerotic stroke or transient ischemic attack. *J Am Heart Assoc*. 2015;4(11):e002699.

63. Sze MA, Schloss PD. Looking for a signal in the noise: revisiting obesity and the microbiome [published correction appears in *MBio*. 2017;8(6):e01995-17]. *MBio*. 2016;7(4):e01018-16.

64. Sonnenburg JL, Backhed F. Diet-microbiota interactions as moderators of human metabolism. *Nature*. 2016;535(7610):56-64.

65. Tobin MD, Minelli C, Burton PR, *et al*. Commentary: development of Mendelian randomization: from hypothesis test to 'Mendelian deconfounding'. *Int J Epidemiol*. 2004;33(1):26-29.

66. Lozupone CA, Stombaugh JI, Gordon JI, *et al*. Diversity, stability and resilience of the human gut microbiota. *Nature*. 2012;489(7415):220-230.

Table 1. Associations of Selected Genetically Predicted Gut Microbiota Genera with IHD, T2DM, Adiposity, Lipids and HOMA-IR

Genus	Unit of exposure	Outcome	Combined Estimate[a]	95% confidence interval	*P*
Acidaminococcus	per relative abundance (log_{10})	HDL-C (SD)	0.001[b]	0.0003, 0.002	0.006
Aggregatibacter	per relative abundance (log_{10})	HDL-C (SD)	0.039	0.002, 0.075	0.038
Anaerostipes	per allele	T2DM	0.960	0.926, 0.996	0.032
Bifidobacterium	per allele	IHD Metabochip	0.959	0.943, 0.976	1.7×10^{-6}
		IHD two studies combined	0.985	0.971, 1.000	0.043
		Body mass index (SD)	-0.011	-0.017, -0.005	1.6×10^{-4}
		HDL-C (SD)	0.010	0.003, 0.017	0.004
		LDL-C (SD)	0.026	0.019, 0.033	4.3×10^{-12}
		HOMA-IR (log-transformed)	-0.008	-0.015, -0.001	0.022
Blautia	per allele	LDL-C (SD)	-0.008	-0.014, -0.002	0.011
Desulfovibrio	per relative abundance (log_{10})	HOMA-IR (log-transformed)	0.007	0.0001, 0.014	0.046
Dorea	per relative abundance (Box-Cox transformed)	HOMA-IR (log-transformed)	0.024	0.005, 0.043	0.013
Faecalibacterium	per allele	Waist-hip ratio (SD)	-0.009	-0.016, -0.003	0.008
Lachnospira	per relative abundance (log_{10})	IHD Metabochip	1.095	1.001, 1.197	0.046

Abbreviations: HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment insulin resistance; IHD, ischemic heart disease; LDL-C, low-density lipoprotein cholesterol; SD, standard deviation; T2DM, type 2 diabetes mellitus.

[a] Odds ratio for IHD and T2DM; β for other outcomes.

[b] 0.001 SD higher in HDL-C per relative abundance (log_{10}) increase in *Acidaminococcus*.