American Journal of Epidemiology

Vol. 187, No. 9
DOI: 10.1093/aje/kwy096
Advance Access publication:
May 25, 2018

Original Contribution

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

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, 7 Sassoon Road, Pokfulam, Hong Kong SAR, China (e-mail: cms1@hku.hk).

Initially submitted July 28, 2017; accepted for publication April 23, 2018.

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 to our knowledge no trials with "hard" outcomes have been conducted. We used separate-sample Mendelian randomization to obtain estimates of the associations of 27 genera of gut microbiota with ischemic heart disease, type 2 diabetes mellitus, adiposity, lipid levels, and insulin resistance, based on summary data from CARDIoGRAAMplusC4D and other consortia. Among the 27 genera, a 1-allele increase in single nucleotide polymorphisms related to greater abundance of *Bifidobacterium* was associated with lower risk of ischemic heart disease (odds ratio = 0.985, 95% confidence interval (CI): 0.971, 1.000; $P = 0.04$), a 0.011–standard-deviation lower body mass index (95% CI: −0.017, −0.005), and a 0.026–standard-deviation higher low-density lipoprotein cholesterol level (95% CI: 0.019, 0.033), but the findings were not robust to exclusion of potential pleiotropy. We also identified *Acidaminococcus*, *Aggregatibacter*, *Anaerostipes*, *Blautia*, *Desulfovibrio*, *Dorea*, and *Faecalibacterium* as being nominally associated with type 2 diabetes mellitus or other risk factors. Results from our study indicate that these 8 genera of gut microbiota should be given priority in future research relating the gut microbiome to ischemic heart disease and its risk factors.

gut microbiota; ischemic heart disease; Mendelian randomization; type 2 diabetes mellitus

Abbreviations: BMI, body mass index; CARDIoGRAM, Coronary Artery Disease Genome-Wide Replication and Meta-Analysis; CARDIoGRAMplusC4D, Coronary Artery Disease Genome-Wide Replication and Meta-Analysis (CARDIoGRAM) plus Coronary Artery Disease (C4D) Genetics; *CDH13*, cadherin 13 gene; CVD, cardiovascular disease; GWAS, genome-wide association studies; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment of insulin resistance; IHD, ischemic heart disease; LDL-C, low-density lipoprotein cholesterol; MR, Mendelian randomization; RCT, randomized controlled trial; SNP, single nucleotide polymorphism; T2DM, type 2 diabetes mellitus.

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

Observationally, some taxa of gut microbiota have been associated with CVD, its subtypes, or risk factors. In a small case-control study ($n = 128$), Emoto et al. (4) found the order Lactobacillales to be positively associated with ischemic heart disease (IHD) and the phylum Bacteroidetes to be inversely associated with IHD. A systematic review implicated several species/genera in T2DM but was based on only 4 small heterogeneous observational studies (total $n = 576$) (5). In a recent case-control study ($n = 223$), Liu et al. (6) observed lower abundance of *Bacteroides thetaiotaomicron* in obese persons. In a study of 263 people (51% obese), *Lactobacillus reuteri* was reported to be positively associated with body mass index (BMI; weight (kg)/height $(m)^2$), and *Bifidobacterium animalis*, *Methanobrevibacter smithii*, and *Escherichia coli* were negatively associated with BMI (7). In a cohort of 893 adults, 34 taxa were associated with BMI and lipids, with a false discovery rate of 0.05 (8). In a

study of 277 people without diabetes (58% obese), *Prevotella copri* and *Bacteroides vulgatus* were the main species associated with homeostatic model assessment of insulin resistance (HOMA-IR) (9). However, these small observational studies are difficult to interpret, because they were 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) have suggested that manipulation of the microbiota through probiotics, usually *Lactobacillus* or *Bifidobacterium*, has a protective association with adiposity (10, 11) but mixed associations with high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C) (12–14), and HOMA-IR (12, 13), with high heterogeneity. To our knowledge, no RCT of probiotics with disease endpoints has been conducted. Moreover, large meta-analyses of RCTs of antibiotics, testing the role of antibiotic therapy in CVD—which almost certainly changed the gut microbiome—did not show an association with CVD mortality (15, 16). No association of vancomycin with HOMA-IR was found in an RCT of 57 obese, prediabetic men (17). However, the exact effects of the antibiotics used in those RCTs on individual taxa of gut microbiota are unknown, so at most they suggest that we cannot rule out a role for a specific taxon.

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

METHODS

Genetically predicted gut microbiota genera

Genetic predictors of the abundance of 27 genera of gut microbiota at the level of genome-wide significance ($P < 5 \times 10^{-8}$) 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 on the basis of larger P values, 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 gene (*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 with the same genus ($P = 2.68 \times 10^{-7}$) 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 (*RNU6-6P*)) and rs4988235 (minichromosome maintenance complex component 6 gene (*MCM6*))) 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 and T2DM and their risk factors

The Coronary Artery Disease Genome-Wide Replication and Meta-Analysis (CARDIoGRAM) plus Coronary Artery Disease (C4D) Genetics (CARDIoGRAMplusC4D) Consortium combines data from multiple large genetic studies to identify risk loci for coronary artery disease and myocardial infarction (28–30). CARDIoGRAMplusC4D 1000 Genomes is a case-control study (60,801 cases, 123,504 controls) of IHD, extensively genotyped using the 1000 Genomes, phase 1, version 3 training set, comprising largely people of European descent (77%) (28). As a sensitivity analysis, we also used data from the CARDIoGRAMplusC4D Metabochip Study (63,746 cases and 130,681 controls), a study of persons 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 the CARDIoGRAMplusC4D Metabochip Study, genetic associations were obtained from the more extensively genotyped subset of European-descent individuals in CARDIoGRAM (22,233 cases, 64,762 controls) (30).

Genetic associations with T2DM, adjusted for age and sex, were obtained from the Diabetes Genetics Replication and Meta-Analysis (DIAGRAM) case-control study (34,380 cases, 114,981 controls) (31). Genetic associations with the adiposity measures BMI and waist:hip ratio (per standard deviation increment) were obtained from the Genetic Investigation of Anthropometric Traits (GIANT) Consortium 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 (per standard deviation increment), adjusted for age, age^2, and sex, were obtained from the Global Lipids Genetic Consortium Results, which contain data on 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 obtained from the Meta-Analyses of Glucose and Insulin-Related Traits (MAGIC) Consortium, comprising 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 (see Web Table 1, available at https://academic.oup.com/aje) (36). The variance of a Wald estimate was obtained from Fieller's theorem (37) or an approximation if the variance for the association of the SNP with exposure was not given (38). When different GWAS used incompatible microbiota units for SNPs predicting the same genera, we used SNP

allele-outcome associations (Web Table 2) (39). If the abundance of a genus was predicted by more than 3 uncorrelated SNPs, MR-Egger and weighted median methods were used as sensitivity analyses. The MR-Egger method checks for unknown horizontal pleiotropy as indicated by a nonzero intercept (40), with its "no measurement error" assumption tested by the I^2 statistic (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 $P = 0.0019$ for IHD in CARDIoGRAMplusC4D 1000 Genomes and a cutoff of $P = 0.002$ for the other outcomes. Given the overlap of participants between the 2 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 LLC, College Station, Texas) 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, producing 94 SNPs related to 27 genera of gut microbiota at the level of genome-wide significance. 16S rRNA gene sequencing was used in 4 studies (19–22) and metagenomic sequencing in 1 study (23). In a United Kingdom study of twins (TwinsUK) ($n = 2,731$; 11% men; age range, 19–89 years), 13 SNPs predicted the abundance of 7 genera (Box-Cox–transformed relative abundance) (19). Among 1,812 people from Germany (46% men; age range, 18–83 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, 6–35 years), 29 SNPs predicted 17 genera (log-transformed relative abundance) (21). In 127 Hutterites (38% men; age range, 6–92 years) of European descent, rs2630788 (zinc finger protein 385D gene (*ZNF385D*)) and rs892244 (*CDH13*) predicted the abundance of *Anaerostipes* and *Bifidobacterium* (normalized relative abundance), respectively (22). Finally, in 1,514 participants (42% men; age range, 18–84 years) from Dutch cohorts, 45 SNPs predicted the abundance of 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 3 and 4) to predict the abundance of 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 greater than 10 (Web Table 3).

On the basis of 3 SNPs from different GWAS, *Bifidobacterium* was associated with lower risk of IHD in the 2 CARDIoGRAMplusC4D studies combined, accounting for their overlap (Table 1, Web Figure 1A), although this association was not evident in CARDIoGRAMplusC4D 1000 Genomes (Web Figure 1B). *Bifidobacterium* was also associated with lower BMI (Table 1, Web Figure 1C), higher HDL-C (Table 1, Web Figure 1D), higher LDL-C (Table 1, Web Figure 1E), and lower HOMA-IR (Table 1, Web Figure 1F). 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 5).

We further identified 7 genera nominally associated with IHD risk factors. On the basis of 5 uncorrelated SNPs from the same GWAS, *Acidaminococcus* was associated with higher HDL-C (Table 1, Web Figure 1D). Sensitivity analysis using the MR-Egger and weighted-median methods produced similar estimates (Web Table 6). *Aggregatibacter*, based on 1 SNP, was associated with higher HDL-C (Table 1, Web Figure 1D). *Anaerostipes*, based on 2 SNPs from different GWAS, was associated with lower risk of T2DM (Table 1, Web Figure 1G). *Blautia*, based on 6 SNPs from different SNPs, was associated with lower LDL-C (Table 1, Web Figure 1E). *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 1F). *Faecalibacterium*, based on 4 SNPs from different GWAS, was associated with lower waist:hip ratio (Table 1, Web Figure 1H). However, none of these associations were robust to Bonferroni correction (Table 1).

Additionally, *Lachnospira*, on the basis of 1 SNP, was associated with higher risk of IHD in the CARDIoGRAMplusC4D Metabochip Study (Table 1, Web Figure 1I) but not in the CARDIoGRAMplusC4D 1000 Genomes Study (Web Figure 1B) or in the 2 CARDIoGRAMplusC4D studies combined, accounting for their overlap (Web Figure 1A). No associations were found for any of 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 was, to our knowledge, the first MR study relating gut microbiota to IHD and its risk factors, we found 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 findings had some consistency with those of an observational study showing no robust association of the genera *Bacteroides*, *Blautia*, *Coprococcus*, *Eggerthella*, and *Lachnospira* with BMI, HDL-C, or LDL-C (8), although we also found *Blautia* to be nominally associated with lower LDL-C. However, our study results were less consistent with those of a small case-control study showing the order Lactobacillales to be positively associated with IHD and the phylum Bacteroidetes to be negatively associated with IHD but *Bifidobacterium* to be 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.

Table 1. Associations of Selected Genetically Predicted Genera of Gut Microbiota With Ischemic Heart Disease, Type 2 Diabetes Mellitus, and Their Risk Factors in a Mendelian Randomization Analysis Based on Summary Statistics From Genome-Wide Association Studies

Genus	Unit of Exposure	Outcome	Combined Estimate[a]	95% Confidence Interval	*P* Value
Acidaminococcus	Per relative abundance (log_{10})	HDL-C (per SD increment)	0.001[b]	0.000, 0.002	0.006
Aggregatibacter	Per relative abundance (log_{10})	HDL-C (per SD increment)	0.039	0.002, 0.075	0.038
Anaerostipes	Per allele	Type 2 diabetes mellitus	0.960	0.926, 0.996	0.032
Bifidobacterium	Per allele	IHD (CARDIoGRAMplusC4D Metabochip Study)	0.959	0.943, 0.976	1.7×10^{-6}
		IHD (2 studies combined)	0.985	0.971, 1.000	0.043
		Body mass index[c] (per SD increment)	−0.011	−0.017, −0.005	1.6×10^{-4}
		HDL-C (per SD increment)	0.010	0.003, 0.017	0.004
		LDL-C (per SD increment)	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 (per SD increment)	−0.008	−0.014, −0.002	0.011
Desulfovibrio	Per relative abundance (log_{10})	HOMA-IR (log-transformed)	0.007	0.000, 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 (per SD increment)	−0.009	−0.016, −0.003	0.008
Lachnospira	Per relative abundance (log_{10})	IHD (CARDIoGRAMplusC4D Metabochip Study)	1.095	1.001, 1.197	0.046

Abbreviations: CARDIoGRAMplusC4D, Coronary Artery Disease Genome-Wide Replication and Meta-Analysis (CARDIoGRAM) plus Coronary Artery Disease (C4D) Genetics; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostatic model assessment of insulin resistance; IHD, ischemic heart disease; LDL-C, low-density lipoprotein cholesterol; SD, standard deviation.

[a] Odds ratio for IHD and type 2 diabetes mellitus; β coefficient for other outcomes.

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

[c] Body mass index is calculated as weight (kg)/height $(m)^2$, but SD is the unit because the data were standardized.

Our study also had some consistency with meta-analyses of RCTs showing beneficial associations of probiotics, typically including *Bifidobacterium*, with 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 values 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 act 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 probiotic capsules containing sole species might provide further clarification. Finally, our study had some consistency with meta-analyses of RCTs showing little association between antibiotics and IHD (15, 16), because these RCTs probably changed the gut microbiota but did not affect CVD mortality. RCTs targeting *Bifidobacterium* (or, more generally, investigations of exact relationships between various antibiotics and specific taxa of gut microbiota) might provide further evidence for IHD prevention.

Many potential pathways linking specific gut microbiota to noncommunicable diseases have been identified. A possible pathway linking gut microbiota to IHD is the pathway from 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 that was robustly associated with IHD. Host metabolites linking gut microbiota to T2DM/the metabolic syndrome may exist. Short-chain fatty acids are generated by many genera of gut microbiota, 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 that dietary fiber reduces levels of LDL-C (47), and we further identified that *Blautia*, possibly fueled by dietary fiber (46), might provide the mechanism. The question of 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 acid production. Branched-chain amino acids have essential signaling functions, may be synthesized by *Prevotella copri* and *Bacteroides vulgatus* (9), and have been found to be positively associated with T2DM and BMI (48) but not with any marker of glucose metabolism (48, 49). Correspondingly, we did not find *Bacteroides* to be associated with HOMA-IR. In a recent observational study, Liu et al. (6) found several species of *Bacteroides* to be inversely correlated with levels of branched-chain amino acids, 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 that 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 designed 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 those related to gut microbiota.

In the era of "big data," taking advantage of GWAS and large groups of publicly available data with extensive genotyping enables one to conduct 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 the level of 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 suggested little evidence of that (51). A post-hoc power calculation (52) assuming a statistical confidence level of 0.05, an R^2 value equaling genus heritability, and an effect size shown in Table 1 suggested statistical power greater than 80% for the associations of *Bifidobacterium* with BMI and LDL-C but power 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 findings or a summary of them would be helpful, as would carrying out further GWAS in larger, more homogenous samples. More generally, our study did not consider associations between each 2 genera of the 27 genera or all bacterial taxa. For example, the family Bifidobacteriaceae is inversely associated with the 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 are available. The possibility of residual pleiotropy is difficult to exclude, as the functions of most SNPs have not been comprehensively identified; use of the MR-Egger method 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 5 GWAS concerned people of European descent (19–23), and the genetic associations with IHD and its risk factors were all from studies conducted largely in European-descent participants 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, but whether canalization is relevant is unknown. Third, the "winner's curse" may have biased our MR estimates, but its direction is ambiguous (51). Finally, selection bias may have influenced our MR estimates, where genetic associations are obtained from studies in older people (55) or otherwise condition on genetic makeup 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 the abundance 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, the United Kingdom twins formed only a very small proportion of these studies, which is unlikely to have created a bias (58), and separate-sample MR reduces the risk of chance associations being generated by the underlying data structure in a 1-sample MR (59). Use of separate samples also meant that we could not test for possible nonlinear associations, conduct subgroup analysis by age and sex, or evaluate diet-microbiome interactions (60), but causal effects should be generally consistent. Second, the 16S rRNA gene sequencing used by investigators in most microbiota GWAS usually only permits resolution at the 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. Third, we cannot rule out the possibility that a ratio of 2 taxa or dysbiosis of gut microbiota contributes to CVD 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, the abundance of 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 are thought to have temporal stability, especially after early childhood, and the dominant force in determining their 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 was also limited by the current lack of understanding of the gut microbiome. A hypothesis-driven study testing epidemiologically established associations would have been preferable, but it was precluded by the lack of knowledge as to the function of each constituent of the microbiome and by the lack of large epidemiologic studies. In addition, differences in statistical methods between gut microbiota GWAS made the units hard to interpret. As such, we have 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 concerned persons of European descent. Gut microbiota may vary between populations (66), so replication in different populations is 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 CVD risk factors, suggesting that these species might be the focus of future

investigation. Further MR studies using multiple robust instruments are needed to confirm these results, given that our study was limited by the use of single genetic instruments for some genera.

ACKNOWLEDGMENTS

Author affiliations: School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China (Qian Yang, Shi Lin Lin, Man Ki Kwok, Gabriel M. Leung, C. Mary Schooling); MRC Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom (Qian Yang); Population Health Sciences, Bristol Medical School, 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, New York (C. Mary Schooling).

No specific funding was received for this work.

We thank the following consortia for providing their publicly available summary data: Coronary Artery Disease Genome-Wide Replication and Meta-Analysis (CARDIoGRAM) plus Coronary Artery Disease (C4D) Genetics (CARDIoGRAMplusC4D), Diabetes Genetics Replication and Meta-Analysis (DIAGRAM), Genetic Investigation of Anthropometric Traits (GIANT), Global Lipids Genetics Consortium, and Meta-Analyses of Glucose and Insulin-Related Traits (MAGIC).

Conflict of interest: none declared.

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, Skiecevičienė 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 I^2 statistic. *Int J Epidemiol*. 2016;45(6):1961–1974.

28. Nikpay M, Goel A, Won HH, et al. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47(10):1121–1130.
29. CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. 2013;45:25–33.
30. Schunkert H, König 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. *Ann Math Statist*. 1940;11(3):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 interval 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 [letter; reprinted]. *Int J Epidemiol*. 2004;33(1):9. (First published in *Lancet*. 1986;i(8479):507–508).
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;8:41.
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;7(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, Bäckhed 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.