

## **A multidimensional Mendelian randomization study on the impact of gut dysbiosis on chronic diseases and human longevity**

Éloi Gagnon, MSc<sup>1</sup>, Patricia L. Mitchell, PhD<sup>1</sup>, Hasanga Manikpurage, MSc<sup>1</sup>, Erik Abner, PhD<sup>2</sup>, Nele Taba, MSc<sup>2,3</sup>, Tõnu Esko, PhD<sup>2</sup>, Nooshin Ghodsian, PhD<sup>1</sup>, Sébastien Thériault, MD, MSc<sup>1,4</sup>  
Patrick Mathieu, MD, MSc<sup>1,5</sup>, and Benoit J. Arsenault, PhD<sup>1,6\*</sup>

- 1) Centre de recherche de l'Institut universitaire de cardiologie et de pneumologie de Québec, Québec (QC), Canada
- 2) Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Riia 23b, 51010, Estonia
- 3) Institute of Molecular and Cell Biology, University of Tartu, Tartu, Riia 23, 51010, Estonia
- 4) Department of Molecular Biology, Medical Biochemistry and Pathology, Faculty of Medicine, Université Laval, Québec (QC), Canada
- 5) Department of Surgery, Faculty of Medicine, Université Laval, Québec (QC), Canada
- 6) Department of Medicine, Faculty of Medicine, Université Laval, Québec (QC), Canada

### \*Address for correspondence

Benoit Arsenault, PhD

Centre de recherche de l'Institut universitaire de cardiologie et de pneumologie de Québec –  
Université Laval

Y-3106, Pavillon Marguerite D'Youville, 2725 chemin Ste-Foy

Québec (QC), Canada G1V 4G5

Telephone: 418-656-8711 ext. 3498

Email: [benoit.arsenault@criucpq.ulaval.ca](mailto:benoit.arsenault@criucpq.ulaval.ca)

## **Abstract**

Alterations of the gut microbiota, often referred to as gut dysbiosis, have been associated with several chronic diseases and longevity in pre-clinical models as well as in observational studies. Whether these relationships underlie causal associations in humans remains to be established. We aimed to determine whether gut dysbiosis influences the risk of chronic diseases and longevity using a comprehensive 2-Sample Mendelian randomization (2SMR) approach. We included as exposures inflammatory bowel disease (IBD) as a human model of gut dysbiosis, 11 gut-associated metabolites and pathways and 48 microbial taxa. Study outcomes included eight chronic diseases previously linked with gut dysbiosis using observational studies (Alzheimer's disease, depression, type 2 diabetes, non-alcoholic fatty liver disease, coronary artery disease (CAD), stroke, osteoporosis and chronic kidney disease) as well as parental longevity and life expectancy. Neither IBD, nor gut-associated metabolites were causally associated with chronic disease or lifespan. After multiple testing correction for 582 tests, no microbial taxa-chronic disease associations remained significant. After robustness analyses and multivariate MR to correct for body mass index and alcohol intake on all 42 nominally significant causal relationships, four associations remained. Altogether, results of this multidimensional Mendelian randomization study suggest that gut dysbiosis has little impact on chronic diseases and human longevity and that previous documented associations may not underly causal relationships. Studies with larger sample sizes and more optimal taxonomic discrimination may ultimately be required to determine whether the human gut microbiota plays a causal role in the etiology of chronic diseases and longevity.

## Introduction

The human gut microbiota is the microbial symbiotic organ residing in the gut. It is involved in key metabolic and immunological processes including host immunity, food digestion, intestinal endocrine function and intestinal permeability (Lynch and Pedersen 2016). Several observational studies revealed that alterations of the gut microbiota, often referred to as gut dysbiosis, are associated with a wide range of diseases and metabolic conditions (Hooks and O'Malley 2017). In individuals with gut dysbiosis, dysfunction of the gut mucosal barrier (“leaky gut”) may lead to an increased bacterial translocation through the mucosal cell line and into the blood, thereby affecting the homeostasis of other organs and promoting disease development and progression (Anhê et al. 2020; Konturek et al. 2018).

Inflammatory bowel disease (IBD), comprising ulcerative colitis and Crohn’s disease, is the most frequently reported disease unequivocally associated with gut microbiota changes and gut inflammation (Schippa and Conte 2014). It involves chronic inflammation of the digestive tract due to aberrant immune response, possibly mediated by dysbiosis. Many studies have reported that the composition of microbiota in IBD is altered compared to healthy subjects (Andoh et al. 2007; Fujimoto et al. 2013; Nishino et al. 2018; Takahashi et al. 2016). Gut dysbiosis could contribute to inflammation and immune responses leading to IBD and this inflammatory environment could in turn promote dysbiosis (Sartor and Wu 2017). Results of several prospective observational studies suggested that patients with IBD may be at increased risk of multiple chronic diseases such as cardiovascular disease (Cainzos-Achirica et al. 2020), respiratory disease, arthritis, liver conditions, kidney failure (Xu 2018), osteoporosis and depression (Bernstein et al. 2019), consistent with the view that dysbiosis exerts systemic negative effects promoting human diseases.

The systemic effects of gut dysbiosis are partly mediated through by-products of the microbiome. These microbial metabolites can reach the peripheral circulation via the portal vein (den Besten et al. 2013), or diffuse readily and be taken up by the gut mucosa (Flint 2016), where they can reach organs and act as substrates or signaling molecules. This bidirectional crosstalk between the gut microbiota and different organs occurs via the gut-liver axis (Albillos, Gottardi, and Rescigno 2020), the gut-brain axis (Carabotti et al. 2015), the gut-bone axis (Villa, Ward, and Comelli 2017),

the gut-kidney axis (Evenepoel, Poesen, and Meijers 2017), the gut-lung axis (Marsland, Trompette, and Gollwitzer 2015) and the gut-heart axis (Bartolomaeus, McParland, and Wilck 2020). Specific classes of microbiota-derived metabolites, notably short-chain fatty acids (De Vadder et al. 2014), branched-chain amino acids (Arany and Neinast 2018), trimethylamine N-oxide (Shan et al. 2017) and indole (Beaumont et al. 2018) have been strongly implicated in the pathogenesis of metabolic disorders (Agus, Clément, and Sokol 2020), lifespan (Wilmanski et al. 2021), neurological and cardiovascular diseases (Martinez, Leone, and Chang 2017).

Over the past few years, the gut microbiota emerged as a therapeutic target of great interest to prevent and/or treat chronic diseases and improve human lifespan and healthspan. An overwhelming amount of supportive evidence from pre-clinical models contributed to the widely-accepted view that a large number of diseases and pathological processes could be influenced by the microbiome, from early metabolic perturbations to full-blown diseases and premature mortality. Fecal transplantation studies in rodents have provided promising results for the treatment of obesity (Pérez-Matute et al. 2020), type 2 diabetes (T2D) (H. Wang et al. 2020), depression and chronic stress (Langgartner et al. 2018), liver injury (Y. Liu et al. 2021), myocarditis (Hu et al. 2019) and aging (Chen et al. 2020). Human microbiota-associated (HMA) studies, consisting of the transplantation of feces from human patients into germ-free mice while control mice receive feces from healthy humans, further supported these associations.

A systematic review conducted in 2019 on the HMA method to study the impact of the microbiota on chronic diseases reported that 95% of such studies (36/38) concluded that fecal transplantation from a sick human donor resulted in at least one worsened symptom compared to healthy controls (Walter et al. 2020). This finding was deemed “implausible” by the authors of this systematic review (Walter et al. 2020). According to Walter et al., in the vast majority of cases, these studies lacked adequate replication and they had statistical and methodological flaws that artificially inflated the odds of obtaining positive findings (Walter et al. 2020). Together with the “file-drawer effect” (whereby positive studies are more likely to be published compared to negative studies), these caveats may distort the odds of translating findings from pre-clinical models into microbiota-targeting therapies to prevent or treat human diseases. Observational studies in humans with various diseases have identified relevant differences in intestinal microbiota composition

(Shreiner, Kao, and Young 2015). However, they are subject to biases such as reverse causality and confounding (through unmeasured confounders) and cannot, by design, assess causality. Obesity, pharmacotherapy, diet, alcohol intake and many other factors appear to be important confounders in the microbiota-health relationships (Vujkovic-Cvijin et al. 2020). Given these limitations, Walter et al. suggested that novel and innovative methods such as Mendelian randomization (MR) should be used to investigate the causal role of the gut microbiota in human disease etiology.

MR is an epidemiological approach that is not subject to many of the biases of observational studies such as reverse causality or confounding. It has the potential to evaluate multiple microbiota-health relationships all at once in a hypothesis-free manner. Briefly, MR uses genetic variants strongly associated with an exposure (features of human gut dysbiosis) to infer causality with an outcome (chronic diseases and human longevity). Twin studies have shown that heritability of the abundance of different bacterial taxa is on average 20%, although some variation exists between taxa (Goodrich et al. 2014; 2016). This is consistent with the view that genes play a non-negligible role in determining gut microbiota composition, making MR a valuable tool to assess the potential causal role of gut dysbiosis in human diseases.

Here, we used a 2-sample MR (2SMR) study design to investigate the potential causal link between three levels of gut dysbiosis-related exposures and eight chronic diseases-related outcomes (coronary artery disease [CAD], T2D, ischemic stroke [IS], nonalcoholic fatty liver disease [NAFLD], chronic kidney disease [CKD], osteoporosis, Alzheimer disease [AD] and depression) and human lifespan (as defined by parental lifespan and living beyond the 90<sup>th</sup> percentile). In our “multidimensional” MR design, we first used IBD as a human model of gut dysbiosis (Schippa and Conte 2014) to determine its association with chronic diseases and human lifespan using multiple MR methods. Second, we investigated the potential causal association of fecal and blood metabolites associated with gut dysbiosis and disease. Third, we leveraged summary statistics from two large genome-wide association studies (GWAS) of gut microbial signature to investigate the association between genetically predicted gut microbes and taxa abundance with chronic diseases and human longevity.

## Results

The conceptual framework of this MR analysis as well as the datasets used to derive the study exposures and outcomes are presented in Figure 1 and Table 1. Briefly, the objective of this multidimensional MR analysis was to test the hypothesis that the gut microbiome causally impacts chronic diseases and longevity and to provide estimates for each exposure/outcome association. We performed 2SMR on exposures associated with gut dysbiosis and relevant outcomes of human diseases and longevity. We used publicly available genome-wide association study (GWAS) summary statistics to extract 10 disease-related outcomes and human longevity (see Methods) and 60 traits related to the microbiome including IBD as a model of human gut inflammation and dysbiosis, fecal and plasma metabolites associated with the gut microbiota as well as microbial abundance of taxa partly under genetic control. Analyses were restricted to participants from European ancestry except for the study by Kurilshikov et al. 2021. The sample from study exposure and outcome did not overlap. We selected only exposures that had at least three independent ( $r^2 < 0.01$ ) genetic instruments at minimum p-value  $< 1e-5$  (the threshold differed between exposures depending of the availability of genetic instruments) with mean F statistics  $> 15$ , resulting in 60 microbiota-related exposures available for MR.

### Effect of inflammatory bowel diseases on chronic diseases and longevity

Gut dysbiosis is a hallmark of IBD (Kostic, Xavier, and Gevers 2014) and the presence of IBD influences the microbial signature (Ni et al. 2017). Using an IBD GWAS of 38,155 cases and 48,485 controls (J. Z. Liu et al. 2015a), we first calculated genetic correlation ( $R_G$ ) with IBD and the 10 outcomes under investigation (Figure 2a). IBD had a weak to modest genetic correlation with all of the outcomes with the majority not significantly different from zero. IBD had a weak positive  $R_G$  with depression and a weak negative  $R_G$  with parental lifespan and CKD. We then selected all independent ( $r^2 \leq 0.001$ ) genome-wide significant single nucleotide polymorphisms (SNPs) (p-value  $< 5e-8$ ) as genetic instruments and performed an Inverse Variance Weighted (IVW)-MR linking IBD with all 10 outcomes. IBD had weak effect on all 10 outcomes under investigation. The only association that reached nominal statistical significance was the one linking IBD to NAFLD (Figure 2b). Each log odds ratio (OR) of IBD increased the risk of NAFLD by 4% (OR=1.04 (1.00-1.08), p=0.02). Robust MR methods were consistent with the absence of causal

association for all associations (Supplementary Table 1). Single SNP MR analyses were consistent with a high degree of heterogeneity among the Wald ratios of individual genetic instruments with the majority of IBD SNPs not being associated with diseases (Supplementary Figure 1A-J). Altogether these analyses do not support a causal role of IBD in the etiology of chronic diseases and longevity.

### Effect of gut microbiome-related metabolites on chronic diseases and longevity

We next sought to determine whether blood metabolites associated with the gut microbiota and fecal microbial metabolites or their functional pathways could influence chronic diseases and longevity. We selected genetic instruments for fecal propionate, a gut metabolite linked with T2D in a recent MR study, from a GWAS of 898 participants (Sanna et al., 2019). We also included 9 plasma metabolites: indole-3-propionate, trimethylamine N-oxide and some of its precursors, carnitine, betaine and choline (Rhee et al. 2013), branched chain amino acids (leucine, isoleucine and valine) as well as the short chain fatty acid acetate (Kettunen et al. 2016). We finally included the microbial pathway involved in 4-aminobutanoate (GABA) degradation (PWY-5022 pathway) acting as a proxy for butyrate production by the gut (Sanna et al., 2019). We selected independent ( $r^2 \leq 0.01$ ) SNPs (for all studies:  $p\text{-value} < 1e\text{-}5$ ; except Kettunen et al.:  $p\text{-value} < 1e\text{-}6$ ) as genetic instruments. We performed IVW-MR for each of the health outcomes under study (Figure 3). A total of 104 exposure-outcomes associations were tested. Nine passed a nominal p-value significance threshold of 0.05 (including the propionate-T2D association) but none of the gut microbiota metabolites were associated with chronic diseases and longevity after multiple testing correction. Figure 3 also reports the association of LDL cholesterol, used as positive control with the outcomes of interest. As expected, LDL cholesterol was strongly associated with cardiovascular diseases (OR = 0.43 95% CI 0.36-0.50,  $P=5.22e\text{-}32$ ) and longevity (Beta = -0.11 95% CI -0.15 -0.075,  $P= 3.94e\text{-}09$ )

### Effect of gut microbial abundance on chronic diseases and longevity

We finally explored the impact of different taxa abundance on health-related outcomes. We first identified 4 recent GWAS on gut microbe abundance with available summary statistics (Kurilshikov et al. 2021; Lopera-Maya et al. 2020; Qin et al. 2020; Rühlemann et al. 2021). We then filtered all microbiota quantitative trait loci (mbQTLs) with  $p\text{-value} < 1.0e\text{-}6$  for all taxa

abundance present in their analysis and kept only exposures with at least three shared mbQTLs with mean F statistics > 15. These criteria were chosen to minimize weak instrument bias and allow the use of sensitivity analysis to assess the validity of the MR assumptions. The study of Lopera-Maya et al., and Qin et al., were removed as none of the exposures satisfied our criterion ( $\geq 3$  independent mbQTL at a p-value <  $1e-6$ , with mean F-statistics > 15). In total, we included available genetic information on 48 microbial taxa abundance from the two recent GWAS studies of Kurilshikov et al., and Rühlemann et al. We then performed IVW-MR on the Wald ratio estimates for each of the 10 health outcomes (Figure 4). Out of 468 exposure-outcome tests, 32 passed a significance threshold of 0.05 and no associations were observed after Benjamini-Hochberg correction for multiple testing.

### Power calculations

All primary MR results can be retrieved in supplementary table 2. Overall, for all 582 associations tested (including IBD, microbiome associated metabolites and taxa abundance) the lowest Benjamini-Hochberg false discovery rate is 0.37 for the association between isoleucine and NALFD. Conservatively, there is no robust causal evidence for any of the 582 associations under study after considering multiple testing. Taking a more lenient approach, there are 42 associations nominally significant at  $p < 0.05$ , that may be worth exploring.

In general, pleiotropy is more likely to bias estimates away from the null (Burgess et al. 2020). Therefore, a null result is generally more robust to the horizontal pleiotropy. However, concerns may arise when reporting such null findings as whether they represent a true absence of causality or are attributed to lack of power and need further exploration. To this end, we performed power analyses assessing the robustness of null findings ( $p > 0.05$ ) (Brion et al., 2013). We observed that our analytical framework provided enough power to detect at least moderate effect. For all 582 associations under study, we have on average 80% ( $\pm 29\%$ ) power to detect a true causal effect (at  $\beta = 0.1$ ) (Supplementary Figure 2), meaning that an increase of 1 standard deviation unit in the exposure would change the risk of being a disease case by  $\sim 10\%$ . We chose this  $\beta$  as a conservative estimate to what is generally observed in observational microbiome research (Kazemian et al. 2020; Zhuang et al. 2020).

### Exploration of promising findings and tests for pleiotropy

One of the key assumptions underlying MR is that genetic instruments affecting gut dysbiosis do not affect diseases or longevity by other mechanisms than the one associated with gut dysbiosis (Davies, Holmes, and Davey Smith 2018). This phenomenon is known as horizontal pleiotropy (Lawlor et al. 2019). We performed tests for pleiotropy on all potentially causal relations (primary analysis p-value <0.05) to ensure the robustness of the causal estimates. We performed robust MR analysis to estimate the robustness of our primary causal estimate to pleiotropy. We used six different methods that make different assumptions about the nature of the underlying pleiotropy: MR-Egger, the MR-Robust Adjusted Profile Score (MR-RAPS), the weighted median, the weighted mode, the MR-PRESSO and the contamination mixture approaches. Consistency across the estimates of the methods provides support to causality. Out of the 42 associations tested, eight associations had half (3/6) or more sensitivity analyses consistent with a true causal effect unlikely to be confounded by pleiotropy (p-value <0.05) (Supplementary Table 3). We therefore retained these associations for further sensitivity analysis. The association linking genetically predicted plasma isoleucine levels to NAFLD was particularly strong, where each standard deviation of genetically predicted isoleucine levels was positively associated with NAFLD (OR = 1.75 95%CI [1.18, 2.59]). The effect was consistent across most robust MR methods except MR-Egger, while the weighted mode and MR-PRESSO methods were marginally significant (Supplementary Figure 3).

### Exploration of BMI and alcohol intake as potential confounding factors

Obesity and alcohol intake frequency were recently identified as major confounding factors in the microbiome-disease associations (Vujkovic-Cvijin et al. 2020). We performed multivariable MR on the eight promising associations to determine if the genetic instruments were influenced by obesity and alcohol intake frequency. This analysis provided evidence that 4 out of 8 aforementioned reported associations might, to a certain extent, be confounded by BMI or alcohol consumption (Supplementary Figure 4). For example, the relation between NAFLD and isoleucine plasma level showed evidence of being confounded in large part by alcohol intake frequency and BMI. Overall, four associations were supported at nominal p-value. First, a 1-standard deviation (SD) increase in Actinobacteria Phylum abundance may decrease the risk of CAD OR=0.89 [0.81-0.98]) (Supplementary Figure 5), while a 1-SD increase in the *Actinobacteria* class may increase

the risk of CAD (OR=1.15 [1.06, 1.26]) (Supplementary Figure 6). Second, a 1-SD increase in the abundance of the genus *Bacteroides* decreased the risk of type 2 diabetes (OR=0.93 [0.86-1.00]) (Supplementary Figure 7). Thirdly, a 1-SD increase in the genus *Phascolarctobacterium* increased lifespan (as proxied by parental lifespan) by 0.02 [0.00, 0.03] years (Supplementary Figure 8).

## Discussion

In order to determine whether previously reported studies using pre-clinical models or observational study designs in humans were consistent with causal associations, we assessed the roles of a wide range of microbial factors and eight chronic diseases and human longevity using MR. We found no association between IBD (a model of gut inflammation and dysbiosis), microbiota-associated fecal or blood metabolites and microbial abundance on these diseases and lifespan after accounting for multiple testing. Some associations, such as the relationship between Actinobacteria at Phylum or Class level and CAD and the relationship between the *Bacteroides* genus and T2D were nominally significant before and after sensitivity analyses and after accounting for confounders and may require further validation. Altogether, results of this study suggest that previously reported associations between the human gut microbiome and human disease might have been due to biases such as reverse causality or confounding and that the impact of gut dysbiosis on chronic diseases and human longevity may not be as prominent as previously suggested.

### Comparisons with other studies

Our results generally contrast with those from previous observational studies. First, IBD was reported in observational studies to be a risk factor of several chronic diseases including respiratory disease, arthritis, liver conditions, kidney failure, osteoporosis and depression (Bernstein et al. 2019; Xu 2018). Notably, meta-analyses of cohort studies suggest that IBD is an independent risk factor for CAD. A 2017 meta-analysis found a multivariate-adjusted independent association between IBD and CAD incident (pooled relative risk [RR]: 1.24; 95% confidence interval [CI]: 1.14 to 1.36) (W. Feng et al. 2017). The same was true for a subsequent larger meta-analysis (RR : 1.17; 95% CI[1.07, 1.27]) (Sun and Tian 2018) suggesting IBD as a potential causal factor.

However, this claim is not supported by our MR analysis highlighting the potential bias induced by unmeasured confounders or participant selection bias, where participants diagnosed with IBD are more likely to be diagnosed with another disease, since they conceivably undergo clinical checkups more often.

Second, microbial metabolites have been associated with health and disease such as neurological disease, NAFLD, cardiovascular disease, survival and type 2 diabetes (Agus, Clément, and Sokol 2020; Martinez, Leone, and Chang 2017; Wilmanski et al. 2021). Notably, over 15 microbial metabolites have been identified as predictors of CAD (Q. Feng et al. 2016). Three metabolites of the dietary lipid phosphatidylcholine--choline, trimethylamine N-oxide (TMAO) and betaine--were shown to predict risk for coronary vascular disease in an independent large clinical cohort (Z. Wang et al. 2011). Moreover, increasing through dietary supplementation the level of TMAO accelerated atherosclerosis in mice (Koeth et al. 2013). By contrast, our MR analysis does not support causality between TMAO and CVD, as previously suggested in another MR investigation (Jia et al. 2019). Dietary factors could arguably act as confounding factors, since meat intake increases TMAO levels (Koeth et al. 2013).

Third, several differences in the microbial composition of diseased and healthy individuals have been identified, but causality remains to be elucidated. RCT of fecal microbiome transfer (FMT) in humans are currently employed to establish causality between microbiome and health, but few have been attempted, and even fewer have been conclusive (Depommier et al. 2019; Mocanu et al. 2021). To date, most successful randomized control study of FMT on humans has been applied to the treatment of recurrent or refractory *Clostridioides difficile* infections (Wortelboer, Nieuwdorp, and Herrema 2019) and some to ulcerative colitis (Costello et al. 2017). Mice FMTs are a valuable exploratory tool, but inference to human subjects is hazardous. Particularly, a substantial proportion of species in the human gut are not present in mice (Zheng, Baird, et al. 2017). For example, several FMT in mice from lean to obese mice resulted in improved cardiometabolic profile (Lai et al. 2018; Zoll et al. 2020), but these findings failed to replicate in humans. A systematic review of all three randomized placebo-controlled studies to treat obesity published to date found no impact of FMT on obesity, fasting plasma glucose, hepatic insulin sensitivity, or cholesterol markers across all included studies (Zhang et al. 2019). For human

observational studies, multiple confounding factors could create spurious correlation between microbiome and chronic diseases, including antibiotic use, age, sex, diet, geography, BMI and alcohol intake (Kim et al. 2017). Moreover, dysbiosis could potentially be a consequence of disease states rather than a causal factor (Cani 2018).

In sum, the large proportion of “negative” findings (i.e., human gut dysbiosis may not cause chronic disease) is in line with recent literature showing an overwhelming positive publication bias in the microbiome literature (Walter et al. 2020). The publication bias can occur if studies are not adequately corrected for multiple testing and can be identified with attempt and failure to replicate (i.e. winners curse bias) (Schloss 2018). For example, it was originally published that individuals with obesity were more likely to have lower bacterial diversity and relative abundances of the phylum Bacteroidetes (Turnbaugh et al. 2009), but this result failed to replicate in 9 independent cohorts (Finucane et al. 2014; Sze and Schloss 2016; Walters, Xu, and Knight 2014). Every method has its own strengths and weaknesses. Therefore, it is crucial to triangulate with different methods such as MR to address a causal research question, combining their strengths to overcome their individual weaknesses (Schloss 2018).

### *Strengths and limitations*

An important strength of this study is the use of the MR design with the largest publicly available GWAS datasets. Because alleles are randomly assigned and fixed at conception, biases due to confounding and reverse causality are mitigated in an MR analysis (Larsson et al. 2020). A further strength is that sample was majoritally restricted to individuals of European ancestry to reduce bias due to population stratification. However, it also restricts the generalizability of the results to this ethnic group. Lastly, power calculations were performed to ensure that results reported here were null because of a clear lack of association and not due to inadequate power to detect associations (Burgess et al. 2020).

Our study, however, has limitations. Robust genetic instruments for microbial species are challenging to find. First, microbiome heterogeneity and interindividual variability are high, substantially reducing the statistical power of microbiome GWAS analyses. Second, the phenotype is distal from individual genes making it a complex polymorphic trait with many variants of small

effect size which could be prone to pleiotropy. Lastly, the twin heritability for gut microbiota taxa abundance is only on average 20% (Goodrich et al. 2014; 2016). This heritability estimate represents the upper ceiling that variance explained by genetic instruments can attain, reducing power. These factors all contributed to the fact that the number of mbQTLs identified to date is rather modest. For these reasons, we used a less stringent p-value cut-off to include a greater number of genetic variants to allow the use of sensitivity analysis and increase power. However, a less stringent p-value cut-off has the trade-off to potentially increase the chance of including false-positive effect variants which induce biases. The most important bias it introduces, the winner's curse bias, refers to the fact that the genetic hits in discovery samples are more likely to be false positive, adding noise to the analysis which will typically bias MR results toward the null. Second in importance, the weak instrument bias occurs when the variance explained by the instrument and the sample size are low (Burgess, Thompson, and CRP CHD Genetics Collaboration 2011). In the two sample MR setting, it will bias towards the null (Burgess, Davies, and Thompson 2016). Third in importance, invalid instruments are pleiotropic variants that affect the outcome via another pathway than the one going through the exposure. Pleiotropic variants can invalidate the MR estimate and usually bias away from the null (Slob and Burgess 2020). We minimized weak instrument biases by including only exposures with genetic instruments with an average F-statistic above 15. We minimized the propensity of the results to be biased by pleiotropic variants by including only exposures with three genetic instruments minimum and performed sensitivity analysis. A second limitation is that the microbiome GWAS included in the current analysis did not target the entire 16S gene, which greatly diminished their ability to achieve a sufficient taxonomic resolution to identify potential therapeutic targets. The meta-analysis by Ruhlemann targeted the V1-V2 subregion while the meta-analysis by Kurilshikov et al. included mostly the V4 subregion and to a lesser extent the V1-V3 subregion. Targeting only 16S subregions such as V4 leads to lower taxonomic resolution achieved compared to sequencing the full V1-V9 16S gene (Johnson et al. 2019). Indeed, using a variable region as a surrogate for the entire 16S gene only allows for the identification of taxa at the genus level or above (Johnson et al. 2019). Being confident at the genus level provides little information to shape disease treatment. Indeed, within a high-level taxon such as a phylum, some species may have a positive correlation with a disease, but some neutral or negative. For example, in our study, the phylum Actinobacteria was potentially protective for CAD, while the subsequent level, the class Actinobacteria was a potential risk factor

for the same disease. Employing third-generation technologies has the potential to allow the sequencing of the full 16S gene in a high throughput manner, and improve taxonomic discrimination.

### Conclusions

Using MR, an approach less subject to reverse causality and confounding factors in comparison to traditional methods, we showed that several features of human gut dysbiosis including IBD, plasma and fecal metabolites as well as a large proportion of associations between 48 microbial taxa and 8 chronic diseases and longevity were not significant. While further study is needed, these results do not support causal impact of microbiome and microbial metabolites on human chronic diseases and longevity. As the microbiome field matures, the use of more populous microbiome GWAS study taking advantage of discriminatory potential of the full 16S gene is warranted.

## **Methods**

### Study exposures

We derived our 60 exposures of interest from six different publicly available data sources (supplementary table 4). Genetic instruments for IBD were obtained from a meta-analysis of 38,155 cases and 48,485 controls. Data originated from a meta-analysis of genome-wide association from seven Crohn's disease and eight ulcerative colitis GWAS collections with genome-wide genotyping or immunochip data from individuals of European descent from 15 countries in Europe, North America and Oceania (Liu et al., 2015a). Diagnosis of IBD was based on radiological, endoscopic or histopathological evaluation. We included independent genetic variants ( $r^2 \leq 0.001$ ) strongly associated with IBD ( $p$ -value  $< 5e-8$ ).

Genetic instruments for fecal propionate and PWY-5022 were obtained from a GWAS on 952 normoglycemic participants of the LifeLines-DEEP cohort, a population-based cohort from northern Netherlands (age ranges 18–84 years) (Sanna et al. 2019). Fecal propionate levels were measured by gas chromatography-mass spectrometry (GCMS). The functional pathway PWY-5022 was obtained with HUMAnN2 (v 0.4.0) (Franzosa et al., 2018) and MetaCyc metabolic-

pathway database (Vatanen et al. 2016). Genotyping was carried out with two Illumina arrays, HumanCytoSNP-12 BeadChip and ImmunoChip.

Genetic instruments for plasma TMAO, carnitine, betaine, choline and indole-3-propionate were extracted from a GWAS conducted in 2076 participants from European ancestry from the Framingham Heart Study (FHS) offspring cohort (Rhee et al. 2013). The FHS offspring cohort is a prospective community-based cohort from Framingham, Massachusetts, USA. Children of the spouse of the FHS study were recruited in 1971. Metabolites profiling was performed by liquid chromatography-mass spectrometry (T. J. Wang et al. 2011). Genotyping was conducted using the Affymetrix 500K mapping array and the Affymetrix 50K gene-focused MIP array. The participants all provided their informed consent and the study was approved by the Boston University Medical Center.

Genetic instruments for plasma branched-chain amino acids (leucine, isoleucine and valine) and acetate were extracted from a meta-analysis of GWAS conducted on 10 European cohorts totaling 24,925 individuals (Kettunen et al. 2016). Human blood metabolites were quantified with quantitative high-throughput NMR metabolomics platform.

Genetic instruments for bacterial taxon were extracted from a GWAS of bacterial taxon abundances of 8,956 German individuals from the PopGen (population-based cohort), the FoCUS (population registry based), the KORA FF4 (population-based adult cohort initiated in 1984) and the SHIP cohort (longitudinal population-based cohort) (Rühlemann et al. 2021). Human Genotyping and fecal microbial 16S rRNA gene surveys were performed using multiple arrays. Additionally, other genetic instruments for bacterial taxon were extracted from a meta-analysis conducted by the MiBioGen consortium on 16S fecal microbiome data from 18,340 individuals (24 cohorts) (Kurilshikov et al. 2021). All cohorts implemented the standardized 16S processing pipeline that uses SILVA as a reference database, with truncation of the taxonomic resolution of the database to genus level. Cohorts were Middle Eastern, East Asian, American Hispanic/Latin, African American and admixed, although the majority of the sample (more than 72%) came from European descent. Because the effect sizes of mbQTLs were not available in the summary data, we estimated beta from the  $z$  statistic using the following equation.  $b = z * SE$  where  $b$  is the beta,

$z$  is the  $z$ -statistics and  $SE$  is the standard error. We estimated  $SE$  with the following equation  $SE = 1/\sqrt{2p(1-p)(n+z^2)}$  where  $p$  is minor allele frequency and  $n$  is sample size. Allele frequency was also not available in the GWAS summary statistics. It was estimated from the reference panel used by 23 out of 24 cohorts, the Haplotype Reference Consortium (HRC 1.1 reference panel) (McCarthy et al. 2016).

### Study Outcomes

We used publicly available GWAS summary statistics from the largest studies of chronic disease that have been previously linked to the human gut microbiota (with the exception of NAFLD which was performed for the purpose of this study) (Supplementary Table 5). **CAD.** GWAS summary statistics for CAD were obtained from a GWAS on 88,192 cases and 162,544 controls from CARDIoGRAMplusC4D and UK Biobank (van der Harst and Verweij 2018). **Ischemic stroke.** GWAS summary statistics for ischemic stroke were obtained in 67,162 cases/454,450 controls from the MEGASTROKE consortium (Malik et al. 2018). **Alzheimer's disease.** GWAS summary statistics for Alzheimer's Disease were obtained from a meta-analysis of the PGC-ALZ, IGAP, ADSP consortium and UK Biobank comprising 71,880 cases/383,378 controls (Jansen et al. 2019). **Depression.** GWAS summary statistics were obtained for depression from a meta analysis of PGC/UKB (excluding 23andme) including 500,199 (170,756 cases and 329,443 controls) individuals (Howard et al. 2019). **Osteoporosis.** Genetic association estimates for osteoporosis were obtained in the UK Biobank (PheCode\_743.1) comprising of 7547 cases and 455,386 controls (Zhou et al. 2018). **Type 2 diabetes.** GWAS summary statistics for type 2 diabetes were obtained from the DIAGRAM consortium (74,124 cases/824,006 controls) (Mahajan et al. 2018). **Chronic kidney disease.** GWAS summary statistics for chronic kidney disease were obtained from the CKDGen consortium including 41,395 cases and 439,303 controls (Wuttke et al. 2019). **Parental lifespan and longevity.** We also used publicly available GWAS summary statistics for human longevity-related outcomes namely parental survival (from a meta-analysis of the UK Biobank and the LifeGen consortium of 26 population cohorts [n=1,012,240; all of European ancestry]) (Timmers et al. 2019) and human longevity from a meta-analysis of 18 European cohorts (Deelen et al. 2019). Longevity was defined as individuals surviving at or beyond the age corresponding to the 90th survival percentile and controls whose age at death or at last contact was at or below the age corresponding to the 60th survival percentile (11,262 cases/25,483 controls).

**Non-alcoholic fatty liver disease.** To obtain a comprehensive set of NAFLD GWAS summary statistics, we performed a GWAS meta-analysis of four cohorts as previously described (Ghodsian et al. 2021). Data sources included the Electronic Medical Records and Genomics (eMERGE) network, the UK Biobank, the Estonian Biobank and FinnGen. The NAFLD GWAS (1106 cases and 8571 controls) in the eMERGE network has previously been published and summary statistics were made available (Namjou et al. 2019). We performed a new GWAS for NAFLD in the UK Biobank that included 2558 cases and 395,241 controls (data application number 25205). UK Biobank has approval from the North West Multi-Centre Research Ethics Committee. We also performed a GWAS for NAFLD using SAIGE in the Estonian Biobank (4119 cases and 190,120 controls). This was approved by the Research Ethics Committee of the University of Tartu (Approval number 288/M-18). Finally, we used publicly available NAFLD GWAS summary statistics of the FinnGen cohort (651 cases and 176,248 controls). The total study sample for the GWAS meta-analysis included 8434 cases and 770,180 controls. This project was approved by the Institutional Review Board of the Quebec Heart and Lung Institute.

#### Selection of Genetic Variants and Variants Harmonization

We first identified all SNPs associated with exposures. Summary parameters for genetic instrument selection can be found at supplementary table 6. These SNPs were then clumped using the 1,000-Genome Project Phase 3 European LD reference panel to make sure instrumental variables were independent (Auton et al., 2015) with a 10 Mb window and pairwise linkage disequilibrium (LD)  $r^2 < 0.01$ . This step was implemented with the *TwoSampleMR* package in R (Hemani et al. 2018). Because of their known association with pleiotropic pathways, we excluded from the analysis all SNPs of the *HLA*, *ABO* and *APOE* genetic regions. We then performed Steiger filtering to remove variants with evidence of a stronger association with the outcome than its association with the exposure (Hemani, Tilling, and Davey Smith 2017). Instrument strength was quantified using the F-statistic (Burgess, Thompson, and CRP CHD Genetics Collaboration 2011), and the variance explained was quantified using the  $R^2$  (Pierce, Ahsan, and VanderWeele 2011).

Variant harmonization was performed by aligning the betas of different studies on the same effect allele with the *TwoSampleMR* package (Hemani et al. 2018). We inferred positive strand alleles, using allele frequencies for palindromes whenever possible. This method is effective when effect

allele frequency difference is large, but may not, however, be a reliable indicator of reference strands when it is close to 0.5 (Hartwig et al. 2016). Consequently, palindromic SNPs with MAF > 0.3 were removed from the analysis. When a particular exposure SNP was not present in the outcome dataset, we used proxy SNPs instead ( $r^2 > 0.8$ ). We obtained the LD matrix of the 1000 genomes project - European sample of the Utah residents from North and West Europe via the LDlink API (Machiela and Chanock 2015). We interrogated the LDlink API with the *LDlinkR* package (Myers, Chanock, and Machiela 2020). We kept only the results based on at least three independent shared SNPs with mean F statistics > 15 to reduce weak instrument bias and allow for sensitivity analysis.

### Statistical Analyses

Genetic correlation was performed using linkage disequilibrium score regression (LDSC). LDSC computes the genetic covariance between traits based on genome wide summary statistics obtained from GWAS (Bulik-Sullivan et al. 2015). The complete summary statistics of IBD and all outcomes were used. The LD Score regression function was fit using the GenomicSEM R package (Grotzinger et al. 2019) except for osteoporosis, where the genetic correlation was computed with the LD Hub database (Zheng, Erzurumluoglu, et al. 2017). LD Hub includes GWAS publicly available summary statistics on hundreds of human traits and enables the assessment of LD score regression among those traits.

We conducted primary MR analysis on each outcome and exposure association. As primary method for causal inference, we performed the IVW method with multiplicative random effects with a standard error correction for under dispersion as recommended by recent MR guidelines (Burgess et al. 2020). The IVW-MR combines the ratio estimates from each genetic instrument in a meta-analysis model by giving more weight to the ratio estimates with lower variance (Burgess, Small, and Thompson 2017). A total of 582 primary analyses were performed: 60 exposures \* 10 outcomes—18 exposures/outcomes with fewer than three overlapping genetic instruments. We applied a Benjamini Hochberg correction for multiple testing using a false discovery rate of 5% to reduce the propensity of false positive finding. All continuous exposure estimates are normalized and reported on a standard deviation scale. For dichotomous traits (all outcomes except longevity

and IBD as exposure), we transformed ORs and CIs to effect sizes and standard error when it was not already done. Longevity estimates were reported by additional year of lifespan.

For nominally significant associations (IVW-MR p-value <0.05) sensitivity analyses to estimate the robustness of our primary causal estimate were performed. We used 6 different robust methods that make different assumptions about the nature of the underlying pleiotropy. As a general test to the presence of pleiotropy, we used Egger intercept regression (MR-Egger), the MR-Robust Adjusted Profile Score (MR-RAPS) approaches and the contamination mixture approach. MR-Egger is similar to the IVW method except the regression model estimates an intercept (Bowden et al., 2015). An intercept significantly different from zero gives indication of pleiotropy (Burgess & Thompson, 2017). The MR-Egger method gives consistent estimates of the causal effect under the InSIDE assumption, which states that pleiotropic effects of genetic variants must be uncorrelated with genetic variants—exposure association (Slob and Burgess 2020). MR-RAPS models the pleiotropic effects of genetic variants directly using a linear random effects distribution. This method provides an unbiased estimate when there is no pleiotropy (Zhao et al. 2018). The contamination mixture provides consistent estimate under the plurality valid assumption (Slob and Burgess 2020). It assumes that not all genetic variants are valid IVs and runs a likelihood function to categorize genetic instruments as valid or invalid. As a general test of robustness of the IVW-MR estimates, we used the weighted median and the weighted mode methods. The weighted median allows an unbiased estimate to be obtained if the “majority valid” assumption is upheld, that is if up to 50% of the weight comes from variants that are valid IVs (Bowden et al. 2016). The weighted mode estimates the causal effect from the value taken by the largest number of genetic variants. This method allows the true causal estimates to be unbiasedly estimated if the “plurality valid” assumption is upheld, that is only 40% of the genetic variants are valid instruments and no larger group with the same estimate exists (Hartwig, Davey Smith, and Bowden 2017). Finally, as a general test to the presence of outliers, we used the outlier-robust method MR-PRESSO, which is a simulation approach where genetic variants are removed based on their contributions to heterogeneity (Verbanck et al. 2018). This method provides consistent estimates under the same assumptions as the IVW-MR method for the set of genetic variants that are not identified as outliers (Slob and Burgess 2020). We also provide a measure of the heterogeneity between variant-specific causal estimates, with Cochran’s Q statistic (Bowden et al. 2017).

Obesity and alcohol intake frequency were recently identified as major confounding factors in the gut-disease associations (Vujkovic-Cvijin et al. 2020) because they are both to some extent associated with the health outcome under study (Blüher 2019; Jani et al. 2021) while potentially simultaneously influencing microbiome composition (Dubinkina et al. 2017). In observational studies, adding BMI and alcohol intake frequency are the most important predictors of microbiota composition and health and adding them as covariates in linear mixed-effect models reduced the numbers of spurious microbiome health associations that are identified as significant (Vujkovic et al. 2021). To account for this, we performed multivariable MR as a sensitivity analysis to correct for measured confounders (Gormley et al. 2020). BMI and alcohol intake frequency GWAS were obtained from accessing the MR-BASE database from the *TwoSampleMR* package (Walker et al. 2019). Both GWAS were derived from the UK biobank. Multivariate MR IVW estimates were computed using the *MendelianRandomization* package (Yavorska and Burgess 2017).

#### *Institutional Review Board Approval*

All GWAS summary statistics that were used we in the public domain with the exception of the NAFLD GWAS meta-analysis that is based on two new analysis from the UK Biobank and the Estonian Biobank. The UK Biobank was performed using data application number 25205. UK Biobank has approval from the North West Multi-Centre Research Ethics Committee. This analysis of the Estonian Biobank was approved by the Research Ethics Committee of the University of Tartu (approval number 288/M-18). For the other publicly available GWAS summary statistics, all participants provided informed consent and study protocols were approved by their respective local ethical committee (Deelen et al. 2019; Howard et al. 2019; Jansen et al. 2019; Kettunen et al. 2016; Kurilshikov et al. 2021; J. Z. Liu et al. 2015b; Mahajan et al. 2018; Rhee et al. 2013; Rühlemann et al. 2021; Sanna et al. 2019; Timmers et al. 2019; van der Harst and Verweij 2018; Wuttke et al. 2019). This project was approved by the Institutional Review Board of the Quebec Heart and Lung Institute.

#### *Data Availability*

All data used in this study are in the public domain (or will be shortly). Supplementary tables 4 and 5 describe the data used and relevant information to retrieve the summary statistics.

### Code Availability

Code was performed in the R V.4.0.0 computing environment using publicly accessible functions from the *TwoSampleMR* V.0.5.5 <https://github.com/MRCIEU/TwoSampleMR>, the *MendelianRandomization* V.0.5.1 <https://cran.rproject.org/web/packages/MendelianRandomization/index.html> and the *data.table* V.1.14.0 <https://github.com/Rdatatable/data.table> packages. The *tidyverse* V.1.3.1 collection of R packages was also used (Wickham et al. 2019).

### References

- Agus, Allison, Karine Clément, and Harry Sokol. 2020. ‘Gut Microbiota-Derived Metabolites as Central Regulators in Metabolic Disorders’. *Gut*, December, gutjnl-2020-323071. <https://doi.org/10.1136/gutjnl-2020-323071>.
- Albillos, Agustín, Andrea de Gottardi, and María Rescigno. 2020. ‘The Gut-Liver Axis in Liver Disease: Pathophysiological Basis for Therapy’. *Journal of Hepatology* 72 (3): 558–77. <https://doi.org/10.1016/j.jhep.2019.10.003>.
- Andoh, Akira, Shinji Sakata, Yuhsuke Koizumi, Keiichi Mitsuyama, Yoshihide Fujiyama, and Yoshimi Benno. 2007. ‘Terminal Restriction Fragment Length Polymorphism Analysis of the Diversity of Fecal Microbiota in Patients with Ulcerative Colitis’. *Inflammatory Bowel Diseases* 13 (8): 955–62. <https://doi.org/10.1002/ibd.20151>.
- Anhê, Fernando F., Benjamin Anderschou Holbech Jensen, Thibault V. Varin, Florence Servant, Sebastian Van Blerk, Denis Richard, Simon Marceau, et al. 2020. ‘Type 2 Diabetes Influences Bacterial Tissue Compartmentalisation in Human Obesity’. *Nature Metabolism* 2 (3): 233–42. <https://doi.org/10.1038/s42255-020-0178-9>.
- Arany, Zoltan, and Michael Neinast. 2018. ‘Branched Chain Amino Acids in Metabolic Disease’. *Current Diabetes Reports* 18 (10): 76. <https://doi.org/10.1007/s11892-018-1048-7>.
- Auton, Adam, Gonçalo R. Abecasis, David M. Altshuler, Richard M. Durbin, Gonçalo R. Abecasis, David R. Bentley, Aravinda Chakravarti, et al. 2015. ‘A Global Reference for Human Genetic Variation’. *Nature* 526 (7571): 68–74. <https://doi.org/10.1038/nature15393>.
- Bartolomaeus, Hendrik, Victoria McParland, and Nicola Wilck. 2020. ‘[Gut-heart axis : How gut bacteria influence cardiovascular diseases]’. *Herz* 45 (2): 134–41. <https://doi.org/10.1007/s00059-020-04897-0>.
- Beaumont, Martin, Audrey M. Neyrinck, Marta Olivares, Julie Rodriguez, Audrey de Rocca Serra, Martin Roumain, Laure B. Bindels, et al. 2018. ‘The Gut Microbiota Metabolite Indole Alleviates Liver Inflammation in Mice’. *The FASEB Journal* 32 (12): 6681–93. <https://doi.org/10.1096/fj.201800544>.
- Bernstein, Charles N, Eric I Benchimol, Alain Bitton, Sanjay K Murthy, Geoffrey C Nguyen, Kate Lee, Jane Cooke-Lauder, and Gilaad G Kaplan. 2019. ‘The Impact of Inflammatory

- Bowel Disease in Canada 2018: Extra-Intestinal Diseases in IBD'. *Journal of the Canadian Association of Gastroenterology* 2 (Supplement\_1): S73–80. <https://doi.org/10.1093/jcag/gwy053>.
- Besten, Gijs den, Karen van Eunen, Albert K. Groen, Koen Venema, Dirk-Jan Reijngoud, and Barbara M. Bakker. 2013. 'The Role of Short-Chain Fatty Acids in the Interplay between Diet, Gut Microbiota, and Host Energy Metabolism'. *Journal of Lipid Research* 54 (9): 2325–40. <https://doi.org/10.1194/jlr.R036012>.
- Blüher, Matthias. 2019. 'Obesity: Global Epidemiology and Pathogenesis'. *Nature Reviews Endocrinology* 15 (5): 288–98. <https://doi.org/10.1038/s41574-019-0176-8>.
- Bowden, Jack, George Davey Smith, and Stephen Burgess. 2015. 'Mendelian Randomization with Invalid Instruments: Effect Estimation and Bias Detection through Egger Regression'. *International Journal of Epidemiology* 44 (2): 512–25. <https://doi.org/10.1093/ije/dyv080>.
- Bowden, Jack, George Davey Smith, Philip C. Haycock, and Stephen Burgess. 2016. 'Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator'. *Genetic Epidemiology* 40 (4): 304–14. <https://doi.org/10.1002/gepi.21965>.
- Bowden, Jack, Fabiola Del Greco M, Cosetta Minelli, George Davey Smith, Nuala Sheehan, and John Thompson. 2017. 'A Framework for the Investigation of Pleiotropy in Two-Sample Summary Data Mendelian Randomization'. *Statistics in Medicine* 36 (11): 1783–1802. <https://doi.org/10.1002/sim.7221>.
- Brion, Marie-Jo A, Konstantin Shakhbazov, and Peter M Visscher. 2013. 'Calculating Statistical Power in Mendelian Randomization Studies'. *International Journal of Epidemiology* 42 (5): 1497–1501. <https://doi.org/10.1093/ije/dyt179>.
- Bulik-Sullivan, Brendan K., Po-Ru Loh, Hilary K. Finucane, Stephan Ripke, Jian Yang, Nick Patterson, Mark J. Daly, Alkes L. Price, and Benjamin M. Neale. 2015. 'LD Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies'. *Nature Genetics* 47 (3): 291–95. <https://doi.org/10.1038/ng.3211>.
- Burgess, Stephen, George Davey Smith, Neil M. Davies, Frank Dudbridge, Dipender Gill, M. Maria Glymour, Fernando P. Hartwig, et al. 2020. 'Guidelines for Performing Mendelian Randomization Investigations'. *Wellcome Open Research* 4 (April). <https://doi.org/10.12688/wellcomeopenres.15555.2>.
- Burgess, Stephen, Neil M. Davies, and Simon G. Thompson. 2016. 'Bias Due to Participant Overlap in Two-sample Mendelian Randomization'. *Genetic Epidemiology* 40 (7): 597–608. <https://doi.org/10.1002/gepi.21998>.
- Burgess, Stephen, Dylan S Small, and Simon G Thompson. 2017. 'A Review of Instrumental Variable Estimators for Mendelian Randomization'. *Statistical Methods in Medical Research* 26 (5): 2333–55. <https://doi.org/10.1177/0962280215597579>.
- Burgess, Stephen, and Simon G. Thompson. 2017. 'Interpreting Findings from Mendelian Randomization Using the MR-Egger Method'. *European Journal of Epidemiology* 32 (5): 377–89. <https://doi.org/10.1007/s10654-017-0255-x>.
- Burgess, Stephen, Simon G Thompson, and CRP CHD Genetics Collaboration. 2011. 'Avoiding Bias from Weak Instruments in Mendelian Randomization Studies'. *International Journal of Epidemiology* 40 (3): 755–64. <https://doi.org/10.1093/ije/dyr036>.
- Cainzos-Achirica, Miguel, Kerri Glassner, Hassan Syed Zawahir, Amit K. Dey, Tanushree Agrawal, Eamonn M.M. Quigley, Bincy P. Abraham, et al. 2020. 'Inflammatory Bowel

- Disease and Atherosclerotic Cardiovascular Disease’. *Journal of the American College of Cardiology* 76 (24): 2895–2905. <https://doi.org/10.1016/j.jacc.2020.10.027>.
- Cani, Patrice D. 2018. ‘Human Gut Microbiome: Hopes, Threats and Promises’. *Gut* 67 (9): 1716–25. <https://doi.org/10.1136/gutjnl-2018-316723>.
- Carabotti, Marilia, Annunziata Scirocco, Maria Antonietta Maselli, and Carola Severi. 2015. ‘The Gut-Brain Axis: Interactions between Enteric Microbiota, Central and Enteric Nervous Systems’. *Annals of Gastroenterology: Quarterly Publication of the Hellenic Society of Gastroenterology* 28 (2): 203–9.
- Chen, Yinfeng, Siyuan Zhang, Bo Zeng, Jiangchao Zhao, Mingyao Yang, Mingwang Zhang, Yan Li, Qingyong Ni, De Wu, and Ying Li. 2020. ‘Transplant of Microbiota from Long-Living People to Mice Reduces Aging-Related Indices and Transfers Beneficial Bacteria’. *Aging (Albany NY)* 12 (6): 4778–93. <https://doi.org/10.18632/aging.102872>.
- Costello, Samuel P., Oliver Waters, Robert V. Bryant, Rosa Katsikeros, Jesica Makanyanga, Mark Schoeman, Reme E. Mountifield, et al. 2017. ‘Short Duration, Low Intensity, Pooled Fecal Microbiota Transplantation Induces Remission in Patients with Mild-Moderately Active Ulcerative Colitis: A Randomised Controlled Trial’. *Gastroenterology* 152 (5): S198–99. [https://doi.org/10.1016/S0016-5085\(17\)30969-1](https://doi.org/10.1016/S0016-5085(17)30969-1).
- Davies, Neil M, Michael V Holmes, and George Davey Smith. 2018. ‘Reading Mendelian Randomisation Studies: A Guide, Glossary, and Checklist for Clinicians’. *BMJ*, July, k601. <https://doi.org/10.1136/bmj.k601>.
- De Vadder, Filipe, Petia Kovatcheva-Datchary, Daisy Goncalves, Jennifer Vinera, Carine Zitoun, Adeline Duchamp, Fredrik Bäckhed, and Gilles Mithieux. 2014. ‘Microbiota-Generated Metabolites Promote Metabolic Benefits via Gut-Brain Neural Circuits’. *Cell* 156 (1–2): 84–96. <https://doi.org/10.1016/j.cell.2013.12.016>.
- Deelen, Joris, Daniel S. Evans, Dan E. Arking, Niccolò Tesi, Marianne Nygaard, Xiaomin Liu, Mary K. Wojczynski, et al. 2019. ‘A Meta-Analysis of Genome-Wide Association Studies Identifies Multiple Longevity Genes’. *Nature Communications* 10 (1): 3669. <https://doi.org/10.1038/s41467-019-11558-2>.
- Depommier, Clara, Amandine Everard, Céline Druart, Hubert Plovier, Matthias Van Hul, Sara Vieira-Silva, Gwen Falony, et al. 2019. ‘Supplementation with *Akkermansia muciniphila* in Overweight and Obese Human Volunteers: A Proof-of-Concept Exploratory Study’. *Nature Medicine* 25 (7): 1096–1103. <https://doi.org/10.1038/s41591-019-0495-2>.
- Dubinkina, Veronika B., Alexander V. Tyakht, Vera Y. Odintsova, Konstantin S. Yarygin, Boris A. Kovarsky, Alexander V. Pavlenko, Dmitry S. Ischenko, et al. 2017. ‘Links of Gut Microbiota Composition with Alcohol Dependence Syndrome and Alcoholic Liver Disease’. *Microbiome* 5 (1): 141. <https://doi.org/10.1186/s40168-017-0359-2>.
- Evenepoel, Pieter, Ruben Poesen, and Björn Meijers. 2017. ‘The Gut-Kidney Axis’. *Pediatric Nephrology (Berlin, Germany)* 32 (11): 2005–14. <https://doi.org/10.1007/s00467-016-3527-x>.
- Feng, Qiang, Zhipeng Liu, Shilong Zhong, Ruijun Li, Huihua Xia, Zhuye Jie, Bo Wen, et al. 2016. ‘Integrated Metabolomics and Metagenomics Analysis of Plasma and Urine Identified Microbial Metabolites Associated with Coronary Heart Disease’. *Scientific Reports* 6 (March): 22525. <https://doi.org/10.1038/srep22525>.
- Feng, Wan, Guangxia Chen, Dawei Cai, Song Zhao, Jiafei Cheng, and Hong Shen. 2017. ‘Inflammatory Bowel Disease and Risk of Ischemic Heart Disease: An Updated Meta-

- Analysis of Cohort Studies'. *Journal of the American Heart Association* 6 (8). <https://doi.org/10.1161/JAHA.117.005892>.
- Finucane, Mariel M., Thomas J. Sharpton, Timothy J. Laurent, and Katherine S. Pollard. 2014. 'A Taxonomic Signature of Obesity in the Microbiome? Getting to the Guts of the Matter'. *PLOS ONE* 9 (1): e84689. <https://doi.org/10.1371/journal.pone.0084689>.
- Flint, Harry J. 2016. 'Gut Microbial Metabolites in Health and Disease'. *Gut Microbes* 7 (3): 187–88. <https://doi.org/10.1080/19490976.2016.1182295>.
- Franzosa, Eric A., Lauren J. McIver, Gholamali Rahnavard, Luke R. Thompson, Melanie Schirmer, George Weingart, Karen Schwarzberg Lipson, et al. 2018. 'Species-Level Functional Profiling of Metagenomes and Metatranscriptomes'. *Nature Methods* 15 (11): 962–68. <https://doi.org/10.1038/s41592-018-0176-y>.
- Fujimoto, Takehide, Hirotugu Imaeda, Kenichiro Takahashi, Eiji Kasumi, Shigeki Bamba, Yoshihide Fujiyama, and Akira Andoh. 2013. 'Decreased Abundance of Faecalibacterium Prausnitzii in the Gut Microbiota of Crohn's Disease'. *Journal of Gastroenterology and Hepatology* 28 (4): 613–19. <https://doi.org/10.1111/jgh.12073>.
- Goodrich, Julia K., Emily R. Davenport, Michelle Beaumont, Matthew A. Jackson, Rob Knight, Carole Ober, Tim D. Spector, Jordana T. Bell, Andrew G. Clark, and Ruth E. Ley. 2016. 'Genetic Determinants of the Gut Microbiome in UK Twins'. *Cell Host & Microbe* 19 (5): 731–43. <https://doi.org/10.1016/j.chom.2016.04.017>.
- Goodrich, Julia K., Jillian L. Waters, Angela C. Poole, Jessica L. Sutter, Omry Koren, Ran Blekhman, Michelle Beaumont, et al. 2014. 'Human Genetics Shape the Gut Microbiome'. *Cell* 159 (4): 789–99. <https://doi.org/10.1016/j.cell.2014.09.053>.
- Gormley, Mark, Tom Dudding, Eleanor Sanderson, Richard M. Martin, Steven Thomas, Jessica Tyrrell, Andrew R. Ness, et al. 2020. 'A Multivariable Mendelian Randomization Analysis Investigating Smoking and Alcohol Consumption in Oral and Oropharyngeal Cancer'. *Nature Communications* 11 (1): 6071. <https://doi.org/10.1038/s41467-020-19822-6>.
- Grotzinger, Andrew D., Mijke Rhemtulla, Ronald de Vlaming, Stuart J. Ritchie, Travis T. Mallard, W. David Hill, Hill F. Ip, et al. 2019. 'Genomic Structural Equation Modelling Provides Insights into the Multivariate Genetic Architecture of Complex Traits'. *Nature Human Behaviour* 3 (5): 513–25. <https://doi.org/10.1038/s41562-019-0566-x>.
- Harst, Pim van der, and Niek Verweij. 2018. 'Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease'. *Circulation Research* 122 (3): 433–43. <https://doi.org/10.1161/CIRCRESAHA.117.312086>.
- Hartwig, Fernando Pires, George Davey Smith, and Jack Bowden. 2017. 'Robust Inference in Summary Data Mendelian Randomization via the Zero Modal Pleiotropy Assumption'. *International Journal of Epidemiology* 46 (6): 1985–98. <https://doi.org/10.1093/ije/dyx102>.
- Hartwig, Fernando Pires, Neil Martin Davies, Gibran Hemani, and George Davey Smith. 2016. 'Two-Sample Mendelian Randomization: Avoiding the Downsides of a Powerful, Widely Applicable but Potentially Fallible Technique'. *International Journal of Epidemiology* 45 (6): 1717–26. <https://doi.org/10.1093/ije/dyx028>.
- Hemani, Gibran, Kate Tilling, and George Davey Smith. 2017. 'Orienting the Causal Relationship between Imprecisely Measured Traits Using GWAS Summary Data'. *PLoS Genetics* 13 (11): e1007081. <https://doi.org/10.1371/journal.pgen.1007081>.

- Hemani, Gibran, Jie Zheng, Benjamin Elsworth, Kaitlin H Wade, Valeriia Haberland, Denis Baird, Charles Laurin, et al. 2018. 'The MR-Base Platform Supports Systematic Causal Inference across the Human Phenome'. *ELife* 7 (May): e34408. <https://doi.org/10.7554/eLife.34408>.
- Hooks, Katarzyna B., and Maureen A. O'Malley. 2017. 'Dysbiosis and Its Discontents'. *MBio* 8 (5). <https://doi.org/10.1128/mBio.01492-17>.
- Howard, David M., Mark J. Adams, Toni-Kim Clarke, Jonathan D. Hafferty, Jude Gibson, Masoud Shirali, Jonathan R. I. Coleman, et al. 2019. 'Genome-Wide Meta-Analysis of Depression Identifies 102 Independent Variants and Highlights the Importance of the Prefrontal Brain Regions'. *Nature Neuroscience* 22 (3): 343–52. <https://doi.org/10.1038/s41593-018-0326-7>.
- Hu, Xiao-Fan, Wen-Yong Zhang, Qiang Wen, Wei-Jun Chen, Zhi-Min Wang, Jian Chen, Feng Zhu, et al. 2019. 'Fecal Microbiota Transplantation Alleviates Myocardial Damage in Myocarditis by Restoring the Microbiota Composition'. *Pharmacological Research* 139 (January): 412–21. <https://doi.org/10.1016/j.phrs.2018.11.042>.
- Jani, Bhautesh Dinesh, Ross McQueenie, Barbara I. Nicholl, Ryan Field, Peter Hanlon, Katie I. Gallacher, Frances S. Mair, and Jim Lewsey. 2021. 'Association between Patterns of Alcohol Consumption (Beverage Type, Frequency and Consumption with Food) and Risk of Adverse Health Outcomes: A Prospective Cohort Study'. *BMC Medicine* 19 (1): 8. <https://doi.org/10.1186/s12916-020-01878-2>.
- Jansen, Iris E., Jeanne E. Savage, Kyoko Watanabe, Julien Bryois, Dylan M. Williams, Stacy Steinberg, Julia Sealock, et al. 2019. 'Genome-Wide Meta-Analysis Identifies New Loci and Functional Pathways Influencing Alzheimer's Disease Risk'. *Nature Genetics* 51 (3): 404–13. <https://doi.org/10.1038/s41588-018-0311-9>.
- Jia, Jinzhu, Pan Dou, Meng Gao, Xuejun Kong, Changwei Li, Zhonghua Liu, and Tao Huang. 2019. 'Assessment of Causal Direction Between Gut Microbiota-Dependent Metabolites and Cardiometabolic Health: A Bidirectional Mendelian Randomization Analysis'. *Diabetes* 68 (9): 1747–55. <https://doi.org/10.2337/db19-0153>.
- Johnson, Jethro S., Daniel J. Spakowicz, Bo-Young Hong, Lauren M. Petersen, Patrick Demkowicz, Lei Chen, Shana R. Leopold, et al. 2019. 'Evaluation of 16S rRNA Gene Sequencing for Species and Strain-Level Microbiome Analysis'. *Nature Communications* 10 (1): 5029. <https://doi.org/10.1038/s41467-019-13036-1>.
- Kazemian, Negin, Morteza Mahmoudi, Frank Halperin, Joseph C. Wu, and Sepideh Pakpour. 2020. 'Gut Microbiota and Cardiovascular Disease: Opportunities and Challenges'. *Microbiome* 8 (1): 36. <https://doi.org/10.1186/s40168-020-00821-0>.
- Kettunen, Johannes, Ayşe Demirkan, Peter Würtz, Harmen H. M. Draisma, Toomas Haller, Rajesh Rawal, Anika Vaarhorst, et al. 2016. 'Genome-Wide Study for Circulating Metabolites Identifies 62 Loci and Reveals Novel Systemic Effects of LPA'. *Nature Communications* 7 (March): 11122. <https://doi.org/10.1038/ncomms11122>.
- Kim, Dorothy, Casey E. Hofstaedter, Chunyu Zhao, Lisa Mattei, Ceylan Tanes, Erik Clarke, Abigail Lauder, et al. 2017. 'Optimizing Methods and Dodging Pitfalls in Microbiome Research'. *Microbiome* 5 (May). <https://doi.org/10.1186/s40168-017-0267-5>.
- Koeth, Robert A., Zeneng Wang, Bruce S. Levison, Jennifer A. Buffa, Elin Org, Brendan T. Sheehy, Earl B. Britt, et al. 2013. 'Intestinal Microbiota Metabolism of L-Carnitine, a Nutrient in Red Meat, Promotes Atherosclerosis'. *Nature Medicine* 19 (5): 576–85. <https://doi.org/10.1038/nm.3145>.

- Konturek, Peter Christopher, Igor Alexander Harsch, Kathrin Konturek, Monic Schink, Thomas Konturek, Markus F. Neurath, and Yurdaguel Zopf. 2018. 'Gut–Liver Axis: How Do Gut Bacteria Influence the Liver?' *Medical Sciences* 6 (3). <https://doi.org/10.3390/medsci6030079>.
- Kostic, Aleksandar D., Rannik J. Xavier, and Dirk Gevers. 2014. 'The Microbiome in Inflammatory Bowel Diseases: Current Status and the Future Ahead'. *Gastroenterology* 146 (6): 1489–99. <https://doi.org/10.1053/j.gastro.2014.02.009>.
- Kurilshikov, Alexander, Carolina Medina-Gomez, Rodrigo Bacigalupe, Djawad Radjabzadeh, Jun Wang, Ayse Demirkan, Caroline I. Le Roy, et al. 2021. 'Large-Scale Association Analyses Identify Host Factors Influencing Human Gut Microbiome Composition'. *Nature Genetics*, January, 1–10. <https://doi.org/10.1038/s41588-020-00763-1>.
- Lai, Zi-Lun, Ching-Hung Tseng, Hsiu J. Ho, Cynthia K. Y. Cheung, Jian-Yong Lin, Yi-Ju Chen, Fu-Chou Cheng, et al. 2018. 'Fecal Microbiota Transplantation Confers Beneficial Metabolic Effects of Diet and Exercise on Diet-Induced Obese Mice'. *Scientific Reports* 8 (1): 1–11. <https://doi.org/10.1038/s41598-018-33893-y>.
- Langgartner, Dominik, Carolyn A. Vaihinger, Melanie Haffner-Luntzer, Julia F. Kunze, Anna-Lena J. Weiss, Sandra Foertsch, Stephanie Bergdolt, Anita Ignatius, and Stefan O. Reber. 2018. 'The Role of the Intestinal Microbiome in Chronic Psychosocial Stress-Induced Pathologies in Male Mice'. *Frontiers in Behavioral Neuroscience* 12. <https://doi.org/10.3389/fnbeh.2018.00252>.
- Larsson, Susanna C., Magnus Bäck, Jessica M. B. Rees, Amy M. Mason, and Stephen Burgess. 2020. 'Body Mass Index and Body Composition in Relation to 14 Cardiovascular Conditions in UK Biobank: A Mendelian Randomization Study'. *European Heart Journal* 41 (2): 221–26. <https://doi.org/10.1093/eurheartj/ehz388>.
- Lawlor, Deborah A, Kaitlin Wade, Maria Carolina Borges, Tom Palmer, Fernando Pires Hartwig, Gibran Hemani, and Jack Bowden. 2019. 'A Mendelian Randomization Dictionary: Useful Definitions and Descriptions for Undertaking, Understanding and Interpreting Mendelian Randomization Studies'. Preprint. Open Science Framework. <https://doi.org/10.31219/osf.io/6yys7>.
- Liu, Jimmy Z., Suzanne van Sommeren, Hailiang Huang, Siew C. Ng, Rudi Alberts, Atsushi Takahashi, Stephan Ripke, et al. 2015a. 'Association Analyses Identify 38 Susceptibility Loci for Inflammatory Bowel Disease and Highlight Shared Genetic Risk across Populations'. *Nature Genetics* 47 (9): 979–86. <https://doi.org/10.1038/ng.3359>.
- . 2015b. 'Association Analyses Identify 38 Susceptibility Loci for Inflammatory Bowel Disease and Highlight Shared Genetic Risk across Populations'. *Nature Genetics* 47 (9): 979–86. <https://doi.org/10.1038/ng.3359>.
- Liu, Yongmei, Linda Fan, Zhuo Cheng, Lei Yu, Shuo Cong, Yaxin Hu, Lili Zhu, et al. 2021. 'Fecal Transplantation Alleviates Acute Liver Injury in Mice through Regulating Treg/Th17 Cytokines Balance'. *Scientific Reports* 11 (1): 1611. <https://doi.org/10.1038/s41598-021-81263-y>.
- Lopera-Maya, E. A., A. Kurilshikov, A. van der Graaf, S. Hu, S. Andreu-Sánchez, L. Chen, A. Vich Vila, et al. 2020. 'Effect of Host Genetics on the Gut Microbiome in 7,738 Participants of the Dutch Microbiome Project'. *BioRxiv*, December, 2020.12.09.417642. <https://doi.org/10.1101/2020.12.09.417642>.

- Lynch, Susan V., and Oluf Pedersen. 2016. 'The Human Intestinal Microbiome in Health and Disease'. *The New England Journal of Medicine* 375 (24): 2369–79. <https://doi.org/10.1056/NEJMra1600266>.
- Machiela, Mitchell J., and Stephen J. Chanock. 2015. 'LDlink: A Web-Based Application for Exploring Population-Specific Haplotype Structure and Linking Correlated Alleles of Possible Functional Variants'. *Bioinformatics* 31 (21): 3555–57. <https://doi.org/10.1093/bioinformatics/btv402>.
- Mahajan, Anubha, Daniel Taliun, Matthias Thurner, Neil R. Robertson, Jason M. Torres, N. William Rayner, Anthony J. Payne, et al. 2018. 'Fine-Mapping Type 2 Diabetes Loci to Single-Variant Resolution Using High-Density Imputation and Islet-Specific Epigenome Maps'. *Nature Genetics* 50 (11): 1505–13. <https://doi.org/10.1038/s41588-018-0241-6>.
- Malik, Rainer, Ganesh Chauhan, Matthew Traylor, Muralidharan Sargurupremraj, Yukinori Okada, Aniket Mishra, Loes Rutten-Jacobs, et al. 2018. 'Multiancestry Genome-Wide Association Study of 520,000 Subjects Identifies 32 Loci Associated with Stroke and Stroke Subtypes'. *Nature Genetics* 50 (4): 524–37. <https://doi.org/10.1038/s41588-018-0058-3>.
- Marsland, Benjamin J., Aurélien Trompette, and Eva S. Gollwitzer. 2015. 'The Gut-Lung Axis in Respiratory Disease'. *Annals of the American Thoracic Society* 12 Suppl 2 (November): S150-156. <https://doi.org/10.1513/AnnalsATS.201503-133AW>.
- Martinez, Kristina B., Vanessa Leone, and Eugene B. Chang. 2017. 'Microbial Metabolites in Health and Disease: Navigating the Unknown in Search of Function'. *The Journal of Biological Chemistry* 292 (21): 8553–59. <https://doi.org/10.1074/jbc.R116.752899>.
- McCarthy, Shane, Sayantan Das, Warren Kretzschmar, Olivier Delaneau, Andrew R. Wood, Alexander Teumer, Hyun Min Kang, et al. 2016. 'A Reference Panel of 64,976 Haplotypes for Genotype Imputation'. *Nature Genetics* 48 (10): 1279–83. <https://doi.org/10.1038/ng.3643>.
- Mocanu, Valentin, Zhengxiao Zhang, Edward C. Deehan, Dina H. Kao, Naomi Hotte, Shahzeer Karmali, Daniel W. Birch, Kalutota K. Samarasinghe, Jens Walter, and Karen L. Madsen. 2021. 'Fecal Microbial Transplantation and Fiber Supplementation in Patients with Severe Obesity and Metabolic Syndrome: A Randomized Double-Blind, Placebo-Controlled Phase 2 Trial'. *Nature Medicine* 27 (7): 1272–79. <https://doi.org/10.1038/s41591-021-01399-2>.
- Myers, Timothy A., Stephen J. Chanock, and Mitchell J. Machiela. 2020. 'LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations'. *Frontiers in Genetics* 11. <https://doi.org/10.3389/fgene.2020.00157>.
- Ni, Josephine, Gary D. Wu, Lindsey Albenberg, and Vesselin T. Tomov. 2017. 'Gut Microbiota and IBD: Causation or Correlation?' *Nature Reviews. Gastroenterology & Hepatology* 14 (10): 573–84. <https://doi.org/10.1038/nrgastro.2017.88>.
- Nishino, Kyohei, Atsushi Nishida, Ryo Inoue, Yuki Kawada, Masashi Ohno, Shigeki Sakai, Osamu Inatomi, et al. 2018. 'Analysis of Endoscopic Brush Samples Identified Mucosa-Associated Dysbiosis in Inflammatory Bowel Disease'. *Journal of Gastroenterology* 53 (1): 95–106. <https://doi.org/10.1007/s00535-017-1384-4>.
- Pérez-Matute, Patricia, María Íñiguez, María de Toro, Emma Recio-Fernández, and José A. Oteo. 2020. 'Autologous Fecal Transplantation from a Lean State Potentiates Caloric Restriction Effects on Body Weight and Adiposity in Obese Mice'. *Scientific Reports* 10 (1): 9388. <https://doi.org/10.1038/s41598-020-64961-x>.

- Pierce, Brandon L, Habibul Ahsan, and Tyler J VanderWeele. 2011. 'Power and Instrument Strength Requirements for Mendelian Randomization Studies Using Multiple Genetic Variants'. *International Journal of Epidemiology* 40 (3): 740–52. <https://doi.org/10.1093/ije/dyq151>.
- Qin, Youwen, Aki S. Havulinna, Yang Liu, Pekka Jousilahti, Scott C. Ritchie, Alex Tokolyi, Jon G. Sanders, et al. 2020. 'Combined Effects of Host Genetics and Diet on Human Gut Microbiota and Incident Disease in a Single Population Cohort'. *MedRxiv*, September, 2020.09.12.20193045. <https://doi.org/10.1101/2020.09.12.20193045>.
- Rhee, Eugene P., Jennifer E. Ho, Ming-Huei Chen, Dongxiao Shen, Susan Cheng, Martin G. Larson, Anahita Ghorbani, et al. 2013. 'A Genome-Wide Association Study of the Human Metabolome in a Community-Based Cohort'. *Cell Metabolism* 18 (1): 130–43. <https://doi.org/10.1016/j.cmet.2013.06.013>.
- Rühlemann, Malte Christoph, Britt Marie Hermes, Corinna Bang, Shauni Doms, Lucas Moitinho-Silva, Louise Bruun Thingholm, Fabian Frost, et al. 2021. 'Genome-Wide Association Study in 8,956 German Individuals Identifies Influence of ABO Histo-Blood Groups on Gut Microbiome'. *Nature Genetics*, January, 1–9. <https://doi.org/10.1038/s41588-020-00747-1>.
- Sanna, Serena, Natalie R. van Zuydam, Anubha Mahajan, Alexander Kurilshikov, Arnau Vich Vila, Urmo Vösa, Zlatan Mujagic, et al. 2019. 'Causal Relationships between Gut Microbiome, Short-Chain Fatty Acids and Metabolic Diseases'. *Nature Genetics* 51 (4): 600–605. <https://doi.org/10.1038/s41588-019-0350-x>.
- Sartor, R. Balfour, and Gary D. Wu. 2017. 'Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of Inflammatory Bowel Diseases and Therapeutic Approaches'. *Gastroenterology* 152 (2): 327–339.e4. <https://doi.org/10.1053/j.gastro.2016.10.012>.
- Schippa, Serena, and Maria Pia Conte. 2014. 'Dysbiotic Events in Gut Microbiota: Impact on Human Health'. *Nutrients* 6 (12): 5786–5805. <https://doi.org/10.3390/nu6125786>.
- Schloss, Patrick D. 2018. 'Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research'. *MBio* 9 (3). <https://doi.org/10.1128/mBio.00525-18>.
- Shan, Zhilei, Taoping Sun, Hao Huang, Sijing Chen, Liangkai Chen, Cheng Luo, Wei Yang, et al. 2017. 'Association between Microbiota-Dependent Metabolite Trimethylamine-N-Oxide and Type 2 Diabetes'. *The American Journal of Clinical Nutrition* 106 (3): 888–94. <https://doi.org/10.3945/ajcn.117.157107>.
- Shreiner, Andrew B., John Y. Kao, and Vincent B. Young. 2015. 'The Gut Microbiome in Health and in Disease'. *Current Opinion in Gastroenterology* 31 (1): 69–75. <https://doi.org/10.1097/MOG.0000000000000139>.
- Slob, Eric A. W., and Stephen Burgess. 2020. 'A Comparison of Robust Mendelian Randomization Methods Using Summary Data'. *Genetic Epidemiology* 44 (4): 313–29. <https://doi.org/10.1002/gepi.22295>.
- Sun, Hong-Hao, and Feng Tian. 2018. 'Inflammatory Bowel Disease and Cardiovascular Disease Incidence and Mortality: A Meta-Analysis'. *European Journal of Preventive Cardiology* 25 (15): 1623–31. <https://doi.org/10.1177/2047487318792952>.
- Sze, Marc A., and Patrick D. Schloss. 2016. 'Looking for a Signal in the Noise: Revisiting Obesity and the Microbiome'. *MBio* 7 (4). <https://doi.org/10.1128/mBio.01018-16>.
- Takahashi, Kenichiro, Atsushi Nishida, Takehide Fujimoto, Makoto Fujii, Makoto Shioya, Hirotugu Imaeda, Osamu Inatomi, Shigeki Bamba, Mitsushige Sugimoto, and Akira

- Andoh. 2016. ‘Reduced Abundance of Butyrate-Producing Bacteria Species in the Fecal Microbial Community in Crohn’s Disease’. *Digestion* 93 (1): 59–65. <https://doi.org/10.1159/000441768>.
- Timmers, Paul Rhj, Ninon Mounier, Kristi Lall, Krista Fischer, Zheng Ning, Xiao Feng, Andrew D. Bretherick, et al. 2019. ‘Genomics of 1 Million Parent Lifespans Implicates Novel Pathways and Common Diseases and Distinguishes Survival Chances’. *ELife* 8 (January). <https://doi.org/10.7554/eLife.39856>.
- Turnbaugh, Peter J., Micah Hamady, Tanya Yatsunenko, Brandi L. Cantarel, Alexis Duncan, Ruth E. Ley, Mitchell L. Sogin, et al. 2009. ‘A Core Gut Microbiome in Obese and Lean Twins’. *Nature* 457 (7228): 480–84. <https://doi.org/10.1038/nature07540>.
- Vatanen, Tommi, Aleksandar D. Kostic, Eva d’Hennezel, Heli Siljander, Eric A. Franzosa, Moran Yassour, Raivo Kolde, et al. 2016. ‘Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans’. *Cell* 165 (4): 842–53. <https://doi.org/10.1016/j.cell.2016.04.007>.
- Verbanck, Marie, Chia-Yen Chen, Benjamin Neale, and Ron Do. 2018. ‘Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred from Mendelian Randomization between Complex Traits and Diseases’. *Nature Genetics* 50 (5): 693–98. <https://doi.org/10.1038/s41588-018-0099-7>.
- Villa, Christopher R., Wendy E. Ward, and Elena M. Comelli. 2017. ‘Gut Microbiota-Bone Axis’. *Critical Reviews in Food Science and Nutrition* 57 (8): 1664–72. <https://doi.org/10.1080/10408398.2015.1010034>.
- Vujkovic, Marijana, Shweta Ramdas, Kimberly M. Lorenz, Carolin V. Schneider, Joseph Park, Kyung M. Lee, Marina Serper, et al. 2021. ‘A Genome-Wide Association Study for Nonalcoholic Fatty Liver Disease Identifies Novel Genetic Loci and Trait-Relevant Candidate Genes in the Million Veteran Program’. Preprint. Genetic and Genomic Medicine. <https://doi.org/10.1101/2020.12.26.20248491>.
- Vujkovic-Cvijin, Ivan, Jack Sklar, Lingjing Jiang, Loki Natarajan, Rob Knight, and Yasmine Belkaid. 2020. ‘Host Variables Confound Gut Microbiota Studies of Human Disease’. *Nature* 587 (7834): 448–54. <https://doi.org/10.1038/s41586-020-2881-9>.
- Walker, Venexia M, Neil M Davies, Gibran Hemani, Jie Zheng, Philip C Haycock, Tom R Gaunt, George Davey Smith, and Richard M Martin. 2019. ‘Using the MR-Base Platform to Investigate Risk Factors and Drug Targets for Thousands of Phenotypes’. *Wellcome Open Research* 4 (November). <https://doi.org/10.12688/wellcomeopenres.15334.2>.
- Walter, Jens, Anissa M. Armet, B. Brett Finlay, and Fergus Shanahan. 2020. ‘Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents’. *Cell* 180 (2): 221–32. <https://doi.org/10.1016/j.cell.2019.12.025>.
- Walters, William A., Zech Xu, and Rob Knight. 2014. ‘Meta-Analyses of Human Gut Microbes Associated with Obesity and IBD’. *FEBS Letters* 588 (22): 4223–33. <https://doi.org/10.1016/j.febslet.2014.09.039>.
- Wang, Hui, Yuan Lu, Yan Yan, Shanshan Tian, Dongjie Zheng, Dongjing Leng, Cao Wang, Jingfeng Jiao, Zhiguo Wang, and Yunlong Bai. 2020. ‘Promising Treatment for Type 2 Diabetes: Fecal Microbiota Transplantation Reverses Insulin Resistance and Impaired Islets’. *Frontiers in Cellular and Infection Microbiology* 9. <https://doi.org/10.3389/fcimb.2019.00455>.

- Wang, Thomas J., Martin G. Larson, Ramachandran S. Vasam, Susan Cheng, Eugene P. Rhee, Elizabeth McCabe, Gregory D. Lewis, et al. 2011. 'Metabolite Profiles and the Risk of Developing Diabetes'. *Nature Medicine* 17 (4): 448–53. <https://doi.org/10.1038/nm.2307>.
- Wang, Zeneng, Elizabeth Klipfell, Brian J. Bennett, Robert Koeth, Bruce S. Levison, Brandon Dugar, Ariel E. Feldstein, et al. 2011. 'Gut Flora Metabolism of Phosphatidylcholine Promotes Cardiovascular Disease'. *Nature* 472 (7341): 57–63. <https://doi.org/10.1038/nature09922>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. 'Welcome to the Tidyverse'. *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wilmanski, Tomasz, Christian Diener, Noa Rappaport, Sushmita Patwardhan, Jack Wiedrick, Jodi Lapidus, John C. Earls, et al. 2021. 'Gut Microbiome Pattern Reflects Healthy Ageing and Predicts Survival in Humans'. *Nature Metabolism* 3 (2): 274–86. <https://doi.org/10.1038/s42255-021-00348-0>.
- Wortelboer, Koen, Max Nieuwdorp, and Hilde Herrema. 2019. 'Fecal Microbiota Transplantation beyond Clostridioides Difficile Infections'. *EBioMedicine* 44 (June): 716–29. <https://doi.org/10.1016/j.ebiom.2019.05.066>.
- Wuttke, Matthias, Yong Li, Man Li, Karsten B. Sieber, Mary F. Feitosa, Mathias Gorski, Adrienne Tin, et al. 2019. 'A Catalog of Genetic Loci Associated with Kidney Function from Analyses of a Million Individuals'. *Nature Genetics* 51 (6): 957–72. <https://doi.org/10.1038/s41588-019-0407-x>.
- Xu, Fang. 2018. 'Health-Risk Behaviors and Chronic Conditions Among Adults with Inflammatory Bowel Disease — United States, 2015 and 2016'. *MMWR. Morbidity and Mortality Weekly Report* 67. <https://doi.org/10.15585/mmwr.mm6706a4>.
- Yavorska, Olena O., and Stephen Burgess. 2017. 'MendelianRandomization: An R Package for Performing Mendelian Randomization Analyses Using Summarized Data'. *International Journal of Epidemiology* 46 (6): 1734–39. <https://doi.org/10.1093/ije/dyx034>.
- Zhang, Zhengxiao, Valentin Mocanu, Chenxi Cai, Jerry Dang, Linda Slater, Edward C. Deehan, Jens Walter, and Karen L. Madsen. 2019. 'Impact of Fecal Microbiota Transplantation on Obesity and Metabolic Syndrome—A Systematic Review'. *Nutrients* 11 (10). <https://doi.org/10.3390/nu11102291>.
- Zhao, Qingyuan, Jingshu Wang, Gibran Hemani, Jack Bowden, and Dylan S. Small. 2018. 'Statistical Inference in Two-Sample Summary-Data Mendelian Randomization Using Robust Adjusted Profile Score', January. <https://arxiv.org/abs/1801.09652v3>.
- Zheng, Jie, Denis Baird, Maria-Carolina Borges, Jack Bowden, Gibran Hemani, Philip Haycock, David M. Evans, and George Davey Smith. 2017. 'Recent Developments in Mendelian Randomization Studies'. *Current Epidemiology Reports* 4 (4): 330–45. <https://doi.org/10.1007/s40471-017-0128-6>.
- Zheng, Jie, A Mesut Erzurumluoglu, Benjamin L Elsworth, John P Kemp, Laurence Howe, Philip C Haycock, Gibran Hemani, et al. 2017. 'LD Hub: A Centralized Database and Web Interface to Perform LD Score Regression That Maximizes the Potential of Summary Level GWAS Data for SNP Heritability and Genetic Correlation Analysis'. *Bioinformatics* 33 (2): 272–79. <https://doi.org/10.1093/bioinformatics/btw613>.
- Zhou, Wei, Jonas B. Nielsen, Lars G. Fritsche, Rounak Dey, Maiken E. Gabrielsen, Brooke N. Wolford, Jonathon LeFaive, et al. 2018. 'Efficiently Controlling for Case-Control

- Imbalance and Sample Relatedness in Large-Scale Genetic Association Studies'. *Nature Genetics* 50 (9): 1335–41. <https://doi.org/10.1038/s41588-018-0184-y>.
- Zhuang, Zhenhuang, Ruotong Yang, Wenxiu Wang, Lu Qi, and Tao Huang. 2020. 'Associations between Gut Microbiota and Alzheimer's Disease, Major Depressive Disorder, and Schizophrenia'. *Journal of Neuroinflammation* 17 (1): 288. <https://doi.org/10.1186/s12974-020-01961-8>.
- Zoll, Jereon, Mark N. Read, Sarah E. Heywood, Emma Estevez, Jessica P. S. Marshall, Helene L. Kammoun, Tamara L. Allen, Andrew J. Holmes, Mark A. Febbraio, and Darren C. Henstridge. 2020. 'Fecal Microbiota Transplantation from High Caloric-Fed Donors Alters Glucose Metabolism in Recipient Mice, Independently of Adiposity or Exercise Status'. *American Journal of Physiology-Endocrinology and Metabolism* 319 (1): E203–16. <https://doi.org/10.1152/ajpendo.00037.2020>.

### Acknowledgements

We would like to thank all study participants as well as all investigators of the studies that were used throughout the course of this investigation. EG holds a masters research award from the *Fonds de recherche du Québec: Santé*. (FRQS). BJA and ST hold junior scholar awards from the FRQS. Part of this study was supported by the European Union through the European Regional Development fund. The work of Estonian Genome Center, Univ. of Tartu has been supported by the European Regional Development Fund and grants No. GENTRANSMED (2014-2020.4.01.15-0012), MOBERA5 (Norface Network project no 462.16.107) and 2014-2020.4.01.16-0125. This study was also funded by the European Union through Horizon 2020 research and innovation programme under grant no 810645 and through the European Regional Development Fund project no. MOBEC008 and Estonian Research Council Grant PUT1660.

## Figure legends

**Figure 1. Overview of the multidimensional Mendelian randomization framework used to investigate the association between gut dysbiosis (inflammatory bowel disease, microbial metabolites and microbial taxa abundance) and chronic diseases and longevity.**

**Figure 2. Forest plot of the association of inflammatory bowel disease with all 10 health outcomes.** A) Genetic correlation B) MR causal estimates with inverse variance weighted multiplicative random effect.

**Figure 3. Balloon plot of the association of microbial fecal metabolites, microbial pathway and plasma metabolites with all 10 health outcomes.** Associations with LDL cholesterol are added as positive controls. Non-available (NA) associations stem from a lack of overlapping SNPs or proxies between exposure and outcome resulting in fewer than three genetic instruments in the harmonized data set. Non-significant associations ( $p$ -value  $> 0.05$ ) are depicted with crosses.

**Figure 4. Balloon plot of the association of gut microbial taxa abundance with all 10 health outcomes.** Associations with LDL cholesterol are added as positive controls. Non-available (NA) associations stem from a lack of overlapping SNPs or proxies between exposure and outcome resulting in fewer than three genetic instruments in the harmonized data set. Non-significant associations ( $p$ -value  $> 0.05$ ) are depicted with crosses.

## Tables

Table 1. Description of cohorts used in this study

	Trait	Cohort of study	Sample size	Populations	(Author, year)	pubmed id
<b>Exposures</b>						
Diseases	Inflammatory Bowel Disease	only european individuals from BEL1, BEL2, CEDARS, CHOP, GERMAN, NIDDK, WTCCC, NORWEGIAN and SWEDISH.	38,155 cases - 48,485 controls	European	(Liu et al., 2015)	26192919
functional microbial pathways	PWY-5022	LifeLines-DEEP	952	European	(Sanna et al., 2019)	30778224
fecal metabolites	Propionate	LifeLines-DEEP	898	European	(Sanna et al., 2019)	30778224
fecal metabolites	TMAO, indole-3-propionate (IPA), carnitine, betaine and choline	Framingham Heart Study offspring cohort	2076	European	(Rhee et al., 2013)	23823483
	leucine, isoleucine and valine and acetate	EGCUT, ERF, FTC, FR97, COROGENE, GenMets, HBCS, KORA, LLS, NTR, NFBC, PredictCVD, PROTE, YFS	24,925	European	(Kettunen et al., 2016)	27005778
Microbial Abundance	Microbial relative abundance	PopGen, FoCus, KORA FF4, SHIP	8,956	European	(Rühlemann et al., 2021)	33462482
	Microbial relative abundance	BSPSPC, CARDIAb, CARDIAw, COPSAC, DanFunD, FGFP, FOCUS, GEM, GenR, KSCS, LLD, METSIM, MIBS, NGRC, NTR, PNP, POPCOL, RS3, SHIP, HCHS/SOL, TwinsUK	18,34	Mixed (>72%European)	(Kurilshikov et al., 2021)	33462485
<b>Outcomes</b>						
Chronic diseases	Coronary artery disease	CARDIoGRAMplusC4D and UK Biobank	122,733 cases - 424,528 controls	European	(Van der Harst, 2018)	29212778
	Ischemic stroke	MEGASTROKE consortium	67,162 cases - 454,450 controls	European	(Malik, 2018)	29531354

	Type 2 diabetes	DIAGRAM consortium	74,124 cases - 824,006 controls	European	(Mahajan, 2018)	30297969
	Non-alcoholic fatty liver disease	eMERGE, UK Biobank, Estonian Biobank and FinnGen	8434 cases - 770,180 controls	European	(Ghodsian, 2021)	unpublished
	Chronic kidney disease	CKDGen consortium	41,395 cases - 439,303 controls	European	(Wuttke, 2019)	31152163
	Osteoporosis	UK Biobank	7547 cases - 455,386 controls	European	UKB	PheCode_743.1 in the UKB
	Depression	PGC and UK Biobank	170,756 cases - 329,443 controls	European	(Howard, 2019)	30718901
	Alzheimer's disease	PGC-ALZ, IGAP, ADSP and UK Biobank	246,363 cases - 561,190 controls	European	(Jansen, 2019)	30617256
Lifespan	Parental lifespan	LifeGen and UK Biobank	n = 1012,240	European	(Timmers, 2019)	30642433
	Longevity	100-plus/LASA/ADC, AGES, CEPH, CHS, DKLS, FHS, GEHA, HRS, LLFS LLS, Longevity, MrOS, Newcastle 85 +, RS, SOF, Vitality 90 +,	11,262 cases - 25,483 controls	European	(Deelen, 2019)	31413261

# A multidimensional Mendelian randomization study on the impact of gut dysbiosis on human health and longevity

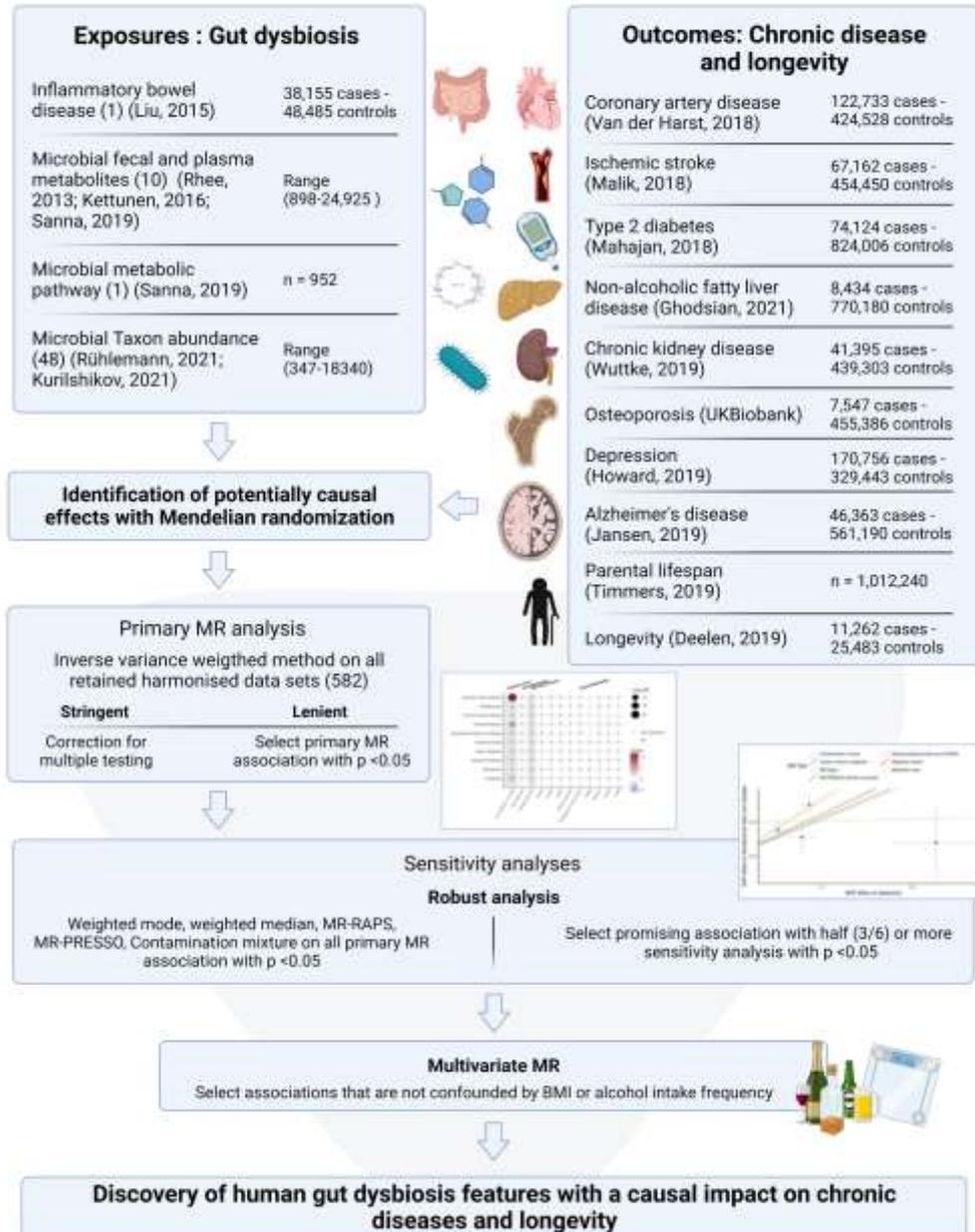
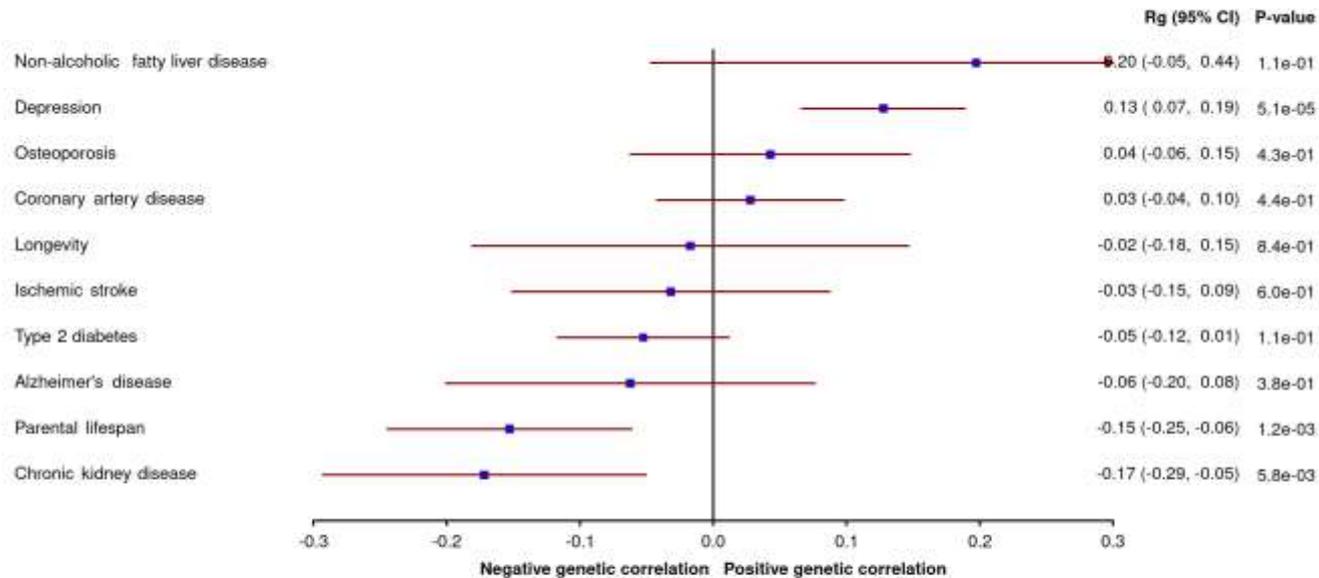


Figure 1

A

Genetic correlation ( $R_g$ ) between  
inflammatory bowel disease and 10 health outcomes



B

Primary MR analysis for inflammatory bowel disease  
and 10 health outcomes

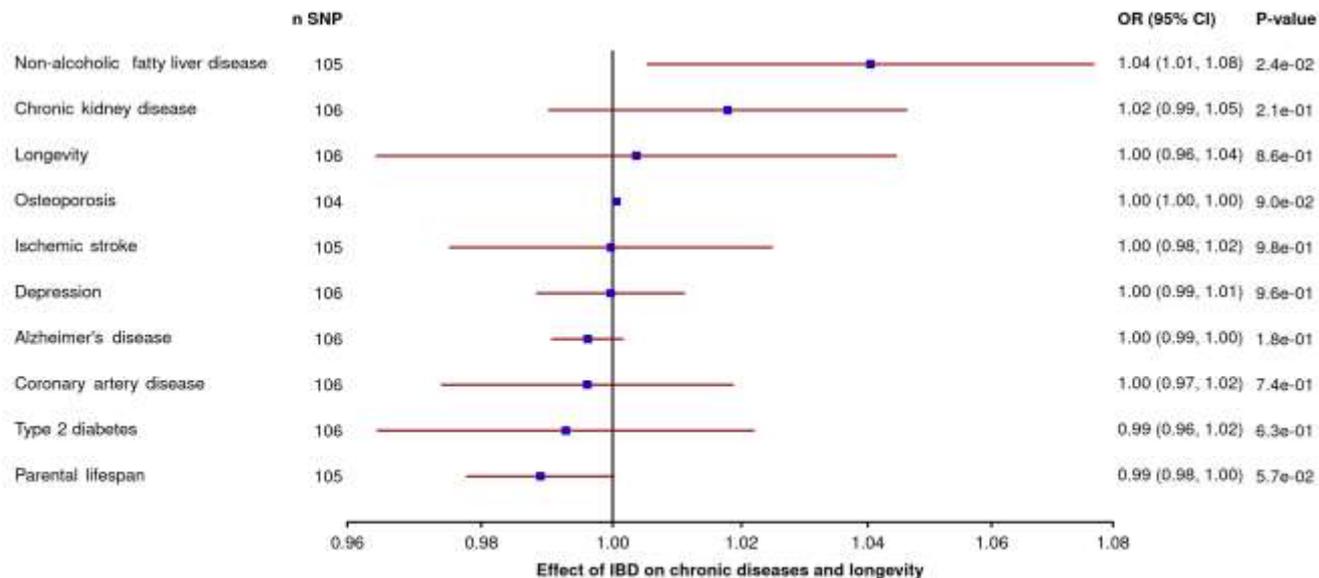


Figure 2

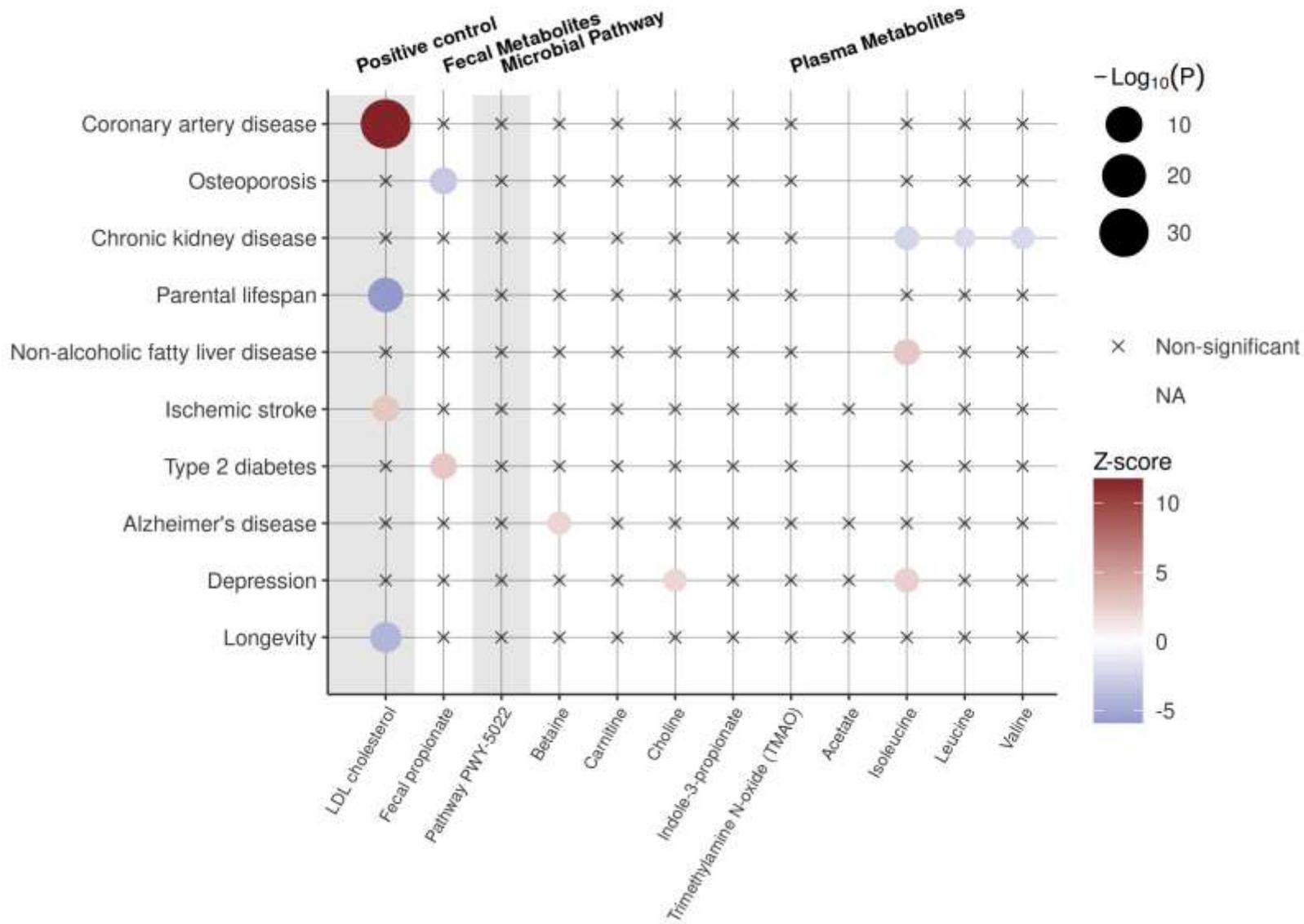


Figure 3

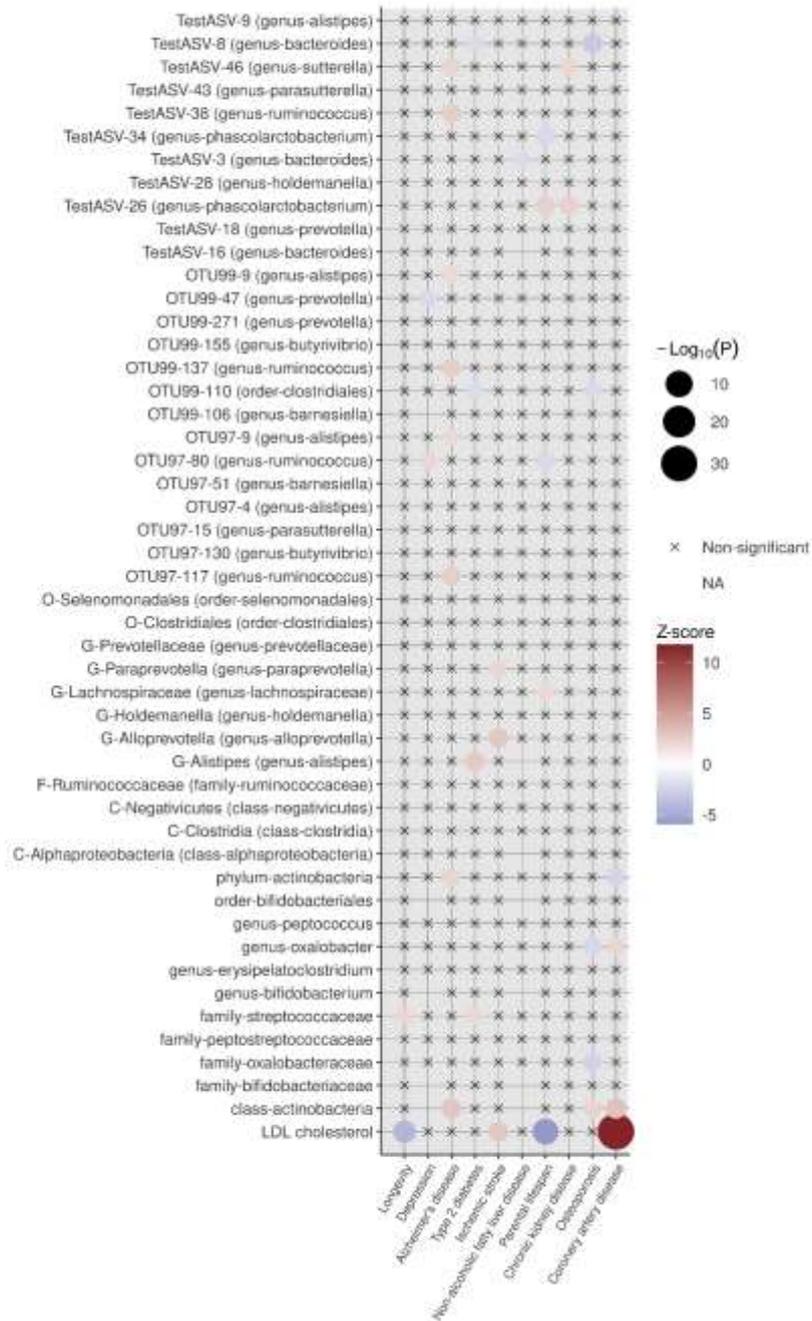


Figure 4