The gut microbiome across the cardiovascular risk spectrum

Femke M. Prins [1,2], Valerie Collij [1†], Hilde E. Groot [2†], Johannes R. Björk [1],
J. Casper Swarte [1], Sergio Andreu-Sánchez [3,4], Bernadien H. Jansen[1],
Jingyuan Fu [3,4], Hermie J.M. Harmsen [5], Alexandra Zhernakova [3], Erik Lipsic[2],
Pim van der Harst [6], Rinse K. Weersma [1,3*‡], and Ranko Gacesa [1,3‡]

[1]University of Groningen, University Medical Center Groningen, Department of Gastroenterology and Hepatology, Groningen, The Netherlands; [2]University of Groningen, University Medical Center Groningen, Department of Cardiology, Groningen, The Netherlands; [3]University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands; [4]University of Groningen, University Medical Center Groningen, Department of Pediatrics, Groningen, The Netherlands; [5]University of Groningen, University Medical Center Groningen, Department of Medical Microbiology and Infection prevention, Groningen, The Netherlands; and [6]Department of Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands

|	
Aims	Despite treatment advancements, cardiovascular disease remains a leading cause of death worldwide. Identifying new targets is crucial for enhancing preventive and therapeutic strategies. The gut microbiome has been associated with coronary artery disease (CAD), however our understanding of specific changes during CAD development remains limited. We aimed to investigate microbiome changes in participants without clinically manifest CAD with different cardiovascular risk levels and in patients with ST-elevation myocardial infarction (STEMI).
Methods and results	In this cross-sectional study, we characterized the gut microbiome using metagenomics of 411 faecal samples from individuals with low ($n = 130$), intermediate ($n = 130$), and high ($n = 125$) cardiovascular risk based on the Framingham score, and STEMI patients ($n = 26$). We analysed diversity, and differential abundance of species and functional pathways while accounting for confounders including medication and technical covariates. *Collinsella stercoris*, *Flavonifractor plautii*, and *Ruthenibacterium lactatiformans* showed increased abundances with cardiovascular risk, while *Streptococcus thermophilus* was negatively associated. Differential abundance analysis revealed eight species and 49 predicted metabolic pathways that were differently abundant among the groups. In the gut microbiome of STEMI patients, there was a depletion of pathways linked to vitamin, lipid, and amino acid biosynthesis.
Conclusion	We identified four microbial species showing a gradual trend in abundance from low-risk individuals to those with STEMI, and observed differential abundant species and pathways in STEMI patients compared to those without clinically manifest CAD. Further investigation is warranted to gain deeper understanding of their precise role in CAD progression and potential implications, with the ultimate goal of identifying novel therapeutic targets.
Lay summary	Despite previous studies demonstrating dysbiosis in STEMI patients, our understanding of the precise microbiome changes across the cardiovascular risk spectrum remains limited. This study addresses this knowledge gap by providing insights into the gut microbiome composition of individuals across varying cardiovascular risk levels and STEMI patients. By examining the gut microbiome of carefully selected participants from the general population with three different risk levels and a unique group of STEMI patients, we identified microbial species and pathways with differential abundance across the groups. Several of these species and pathways are associated with inflammation and lipid metabolism, which are key factors in CAD development.

* Corresponding author. Tel: +31 6 41132824, Fax: +31 50 361 9306, Email: r.k.weersma@umcg.nl
† Shared second author.
‡ Shared last author.

- *Collinsella stercoris, Flavonifractor plautii*, and *Ruthenibacterium lactatiformans* are increasingly abundant, while *Streptococcus thermophilus* is decreasingly abundant across the cardiovascular risk spectrum.
- The gut microbiome of STEMI patients showed eight differentially abundant species compared to groups at risk. Notably, four of these species, characterized by an elevated abundance in STEMI patients, have not been previously reported.

Graphical Abstract
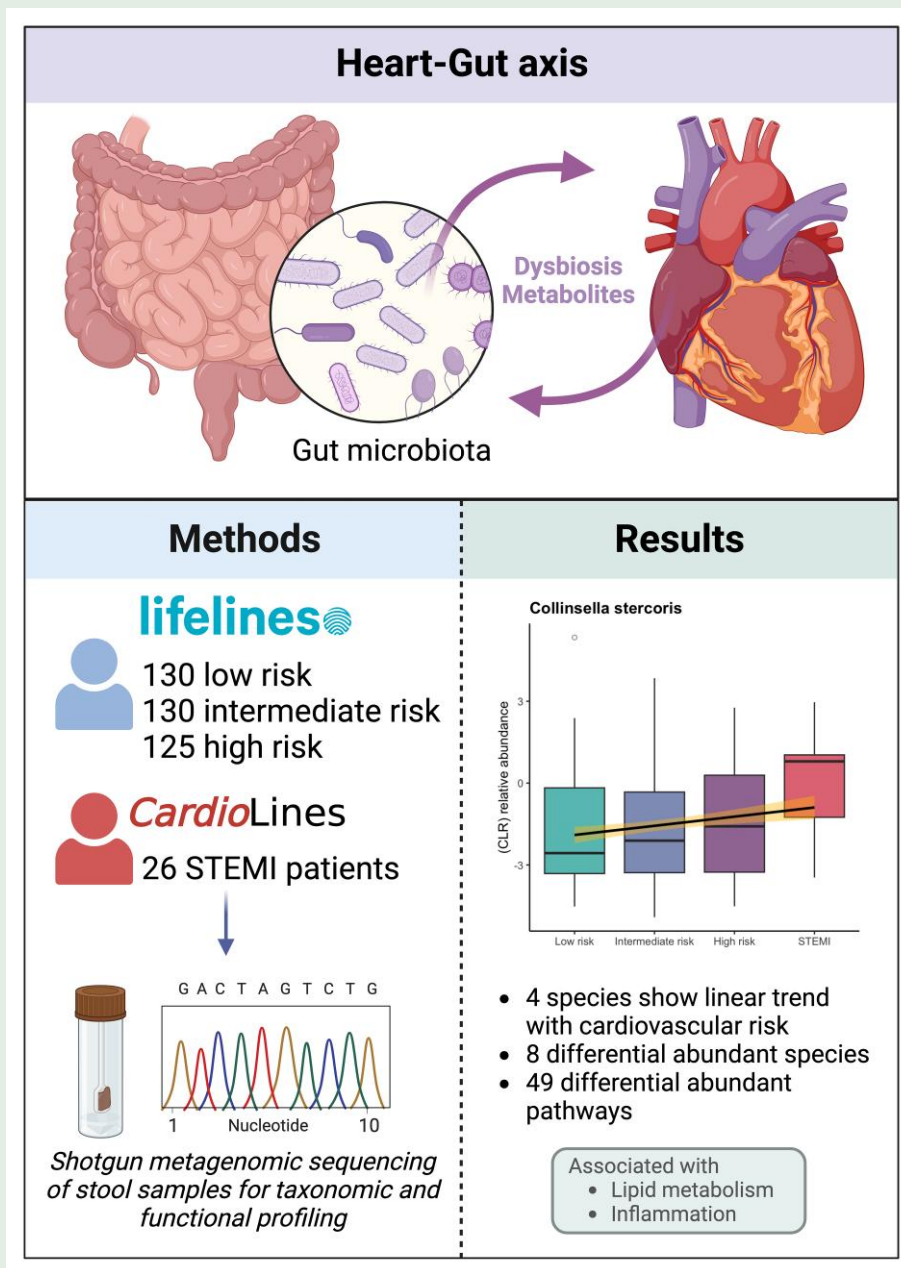
Created with BioRender.com

Keywords   Gut microbiome • STEMI • Cardiovascular risk • Coronary artery disease

Introduction

The human gut microbiome, the trillions of microbes that reside in our gastrointestinal tract, plays a vital role in health and disease. This complex system provides us with defence mechanisms against harmful pathogens and aids in nutrient absorption and digestion.[1] In recent years, studies have shed light on the interplay between the gut microbiome and multiple diseases including coronary artery disease (CAD).[2–4]

An imbalance in the microbial communities of the gut microbiome, known as dysbiosis, can trigger systemic inflammation via various mechanisms, and this is believed to play a role in the development of atherosclerosis and pathogenesis of CAD.[2,5,6] In fact, several studies have confirmed the occurrence of dysbiosis in patients with ST-elevation myocardial infarction (STEMI). Specifically, increased abundances of *Proteobacteria, Aerococcaceae,* and *Enterobacteriaceae* have been reported.[7,8] *Enterobacteriaceae* are known to be pro-inflammatory as the interaction of the endotoxin lipopolysaccharide and immune cells triggers an inflammatory response.[9] On the other hand, decreased abundances of *Firmicutes* and *Lactobacillales* have been reported in STEMI patients. These bacteria are recognized for their anti-inflammatory properties because of their involvement in short-chain fatty acids (SCFA) metabolism.[7] Furthermore, it has been observed that STEMI patients show a decreased abundance of the *Lachnospiraceae* family. This beneficial microbe family has been linked to lower serum levels of trimethylamine *n*-oxide, a gut metabolite recognized to have pro-atherosclerotic effects.[8]

Understanding the complex mechanisms that contribute to disease development and progression is pivotal for developing new preventive and therapeutic strategies for CAD. Therefore, it is important to broaden the scope of investigation beyond conventional cardiovascular risk factors. The potential role of the gut microbiome as a preventive or therapeutic modality for CAD, and as a contributor to cardiovascular risk assessment, is intriguing. However, although dysbiosis signatures have been observed in STEMI patients, there is still a lack of knowledge about the microbial composition in individuals prior to an acute STEMI event. To identify microbial targets for future interventions, it is crucial to determine whether specific changes in microbiome composition occur in the early stages of CAD development.

With this in mind, we studied whether dysbiosis signatures of CAD are already present in individuals at risk for cardiovascular events. Our study aimed to characterize the gut microbiome's composition across a spectrum of cardiovascular risk profiles, ranging from healthy individuals to those with intermediate/high-cardiovascular risk and patients with acute STEMI. We hypothesized that the abundance of microbes reported to be decreased or increased in the STEMI group would demonstrate a gradual trend from low-risk individuals to those with STEMI. We analysed shotgun metagenomes derived from faecal samples of 385 individuals with varying cardiovascular risk profiles and 26 patients with acute STEMI and we compared the metagenomes of different groups on taxonomic and predicted microbial functional levels.

Methods

Cohort descriptions

For this study, we used two Dutch cohorts: (i) myocardial infarction patients participating in the CardioLines Biobank and (ii) general population participants of the Dutch Microbiome Project (DMP).[10] The CardioLines Biobank was established in 2013 and collected body materials from patients with heart disease treated in the University Medical Center Groningen (UMCG, METc number 2012/296). For the present study, inclusion criteria were presence of STEMI, an age older than 18 years, and <6 h since the start of STEMI complaints. Major exclusion criteria were previous myocardial infarction or percutaneous coronary intervention, presence of inflammatory diseases or other heart diseases, usage of anti-inflammatory medication or usage of tavegil/desloratadine. Verbal consent was obtained from all patients at admission, followed by written informed consent. Participants from the DMP were included as groups at risk and a detailed description of the DMP has been published previously.[10] In short, the cohort is part of the Lifelines study (METc number 2017/152), which consists of 167 000 participants from the general population of the northern provinces of the Netherlands for whom stool, blood, and metadata information such as diseases and medication are collected. Written consent was signed by all DMP participants.

Faecal sample collection

All faecal samples for both cohorts were collected according to the same standardized protocol. The CardioLines samples are samples of the first stool after admission to the hospital for STEMI. Hospital nurses were asked to freeze stool samples directly after production on dry ice and thereafter in the freezer. For the DMP, stool sample collection was performed by participants at home using a standardized stool collection kit provided by the UMCG. The frozen samples were then collected and stored at −80°C until further processing.

Microbial extraction and metagenomic sequencing

Microbial DNA was isolated with the QIAmp Fast DNA Stool Mini Kit (Qiagen, Germany) according to the manufacturer's instructions, using the QIAcube robot system for automated sample preparation. All CardioLines samples were handled by the same research technicians who handled the DMP samples. Whole-genome shotgun metagenomic sequencing and library preparation was performed at NovoGene (Hong Kong) using the Illumina HiSeq 2000 sequencing platform.

Metagenomic data processing

Metagenomic sequencing reads mapping to Illumina adapters and low-quality reads were trimmed and filtered out using KneadData tools (v0.5.1). The Bowtie2 tool (v2.3.4.1)[11] was used to remove reads aligning to the human genome (hg19). For the resulting reads, the MetaPhlAn (v3) tool was used to generate taxonomic abundance profiles and the HUMAnN3 pipeline (v0.11.1) for profiling microbial pathway abundances.[12] Taxonomic profiles were subsequently filtered on bacterial and archaeal features. The dataset for the CardioLines cohort contained 526 taxa (2 kingdoms, 10 phyla, 22 classes, 30 orders, 48 families, 107 genera, and 256 species). That for the DMP cohort consisted of 942 taxa (2 kingdoms, 12 phyla, 23 classes, 36 orders, 60 families, 137 genera, and 406 species).

Group stratification and description

Participants from the DMP were stratified into three groups based on cardiovascular risk as determined by their Framingham score and medical history. Framingham score was calculated using the R package *CVrisk* (v1.1.0).[13] This score computes the 10-year risk for atherosclerotic cardiovascular disease events using the following parameters: sex, age, total cholesterol, HDL-cholesterol, systolic blood pressure, blood pressure medication, current smoking, and diabetes. Risk groups were matched (5:1) with STEMI patients ($n = 26$) based on age, sex, and body mass index (BMI) (factors known to affect the gut microbiome) using the R package *MatchIt* (v4.5.1, method = nearest),[14] resulting in the following groups, predominantly composed of males:

(1) Low risk ($n = 130$): DMP participants with no pre-existing disease, no medication use and a Framingham score <10%.
(2) Intermediate risk ($n = 130$): DMP participants without myocardial infarction and with a Framingham score of 10–19%.
(3) High risk ($n = 125$): DMP participants without myocardial infarction and with a Framingham score of 20% or higher.

Description of the microbiome

We started our investigation with computation of diversity measures. To achieve this, we merged taxonomy data from both cohorts and the metabolic pathways, keeping the features that were present in both cohorts. Species that did not overlap between cohorts were examined for their mean relative abundance and prevalence, as documented in Supplementary material online, *Tables S3 and S4*. Of these species, we decided to keep features with a prevalence >5% in the respective cohort (one species) for the calculation of alpha diversity.

For alpha diversity, defined as the diversity of species within a sample (considering its richness, number of different taxonomic groups and evenness and the distribution of abundances of the taxonomic groups), we computed the first three Hill numbers: richness, Shannon diversity index, and the inverse of the Simpson's index. As one moves from richness to the Simpson's index, rare species are increasingly penalized. We computed these indices with the *diversity* function in the R *vegan* (v2.5–7)[15] package.

Statistical differences in mean diversities between groups were tested using the non-parametric Wilcoxon–Mann–Whitney test.

Metagenomic sequencing generates compositional data, which means that information can only be acquired as relative abundances, independent of the absolute microbial load in a sample. That means that if one microbial feature increases, other features decrease accordingly.[16] Not taking the nature of this data into account may lead to false positive results. To facilitate statistical testing for the following analyses, we applied the centred log-ratio (CLR) transformation, which put simply, compares the relative abundance of each microbiome in a sample to the 'average' microbe in the same sample. To deal with zero values, we implemented a pseudocount by calculating it as half the minimum abundance. Furthermore, we applied a prevalence filter of ≥10%, resulting in 136 species that were kept for subsequent analyses.

Beta diversity, a term describing the extent of difference in microbial composition between samples, was assessed using the Aitchison distance, which corresponds to the Euclidean distance of CLR-transformed relative abundances. For this, we used the *vegdist* function in the R *vegan* package. To visualize our data and to examine any potential clustering of the samples, we used principal coordinate analysis (PCoA). Principal coordinate analysis is a dimension–reduction technique used for analysis of high-dimensional data. Principal coordinates were constructed and plotted with the *cmdscale* function. The means of the first five principal components were tested using a Wilcoxon test for each pairwise group comparison (see Supplementary material online, Table S5). To test the community configurations observed in the PCoA plots, we performed Permutational Multivariate Analysis of Variance (PERMANOVA), using the *adonis2* function from the R *vegan* package with 1000 permutations. We tested the proportion of variance in beta diversity explained by phenotypes in a univariable and a multivariable analysis. In the multivariable analysis, we included covariates previously shown to influence the composition of the microbiome, specifically the use of proton pump inhibitors (PPI) and sequence read depth.

Post hoc analysis of risk score

Following the above-mentioned analyses, we aimed to investigate the specific effects of the individual components within the Framingham risk score. Here, we analysed whether certain components of the risk score had a substantial impact on the significant results we observed. To address this, we performed PERMANOVA analyses once more, this time including all eight components of the Framingham score as the variables of interest.

Differential abundance analysis

Differential abundance analysis is a method that can detect variations in the relative abundances of taxonomic groups or pathways between different groups, such as between healthy and diseased individuals. We used linear models to investigate the differential abundance while taking into account the covariates PPI use and sequence read depth (CLR abundance (taxon/pathway) ∼ risk_group + proton_pump_inhibitor + reads). To compare groups, we used the *emmeans* (v1.8.6)[17] package in R, which calculates the estimated marginal means of each group from the fitted model. Any comparison can then be defined, comparing the marginal means between two groups. In total, we performed seven comparisons, including six pairwise comparisons and a 'merged-group' comparison that consisted of the low-, intermediate-, and high-risk groups grouped together vs. the STEMI group. For each contrast, we used the Benjamini–Hochberg false discovery rate (FDR) to correct for multiple testing, with significance set at an FDR < 0.05.

Results

Clinical characteristics

Table 1 presents the clinical characteristics of the study population. In this study, we included four distinct groups: 130 individuals from the general population with a low cardiovascular risk (mean age 59.9 years, 77% males), 130 individuals with an intermediate cardiovascular risk (mean age 62.3 years, 100% males), 125 individuals with a high-cardiovascular risk (mean age 65.4 years, 100% males), and 26 STEMI patients (mean age 63.5 years, 85% males). More than half of the study participants (56.2%) reported a history of current or former smoking. In terms of medication usage, the intermediate and high-risk groups showed some use of medication, especially beta blockers and statins (respectively 23.2% and 19.2% in the high-risk group).

Differences in diversity measures observed in the gut microbiome across different cardiovascular risk profiles

We investigated the diversity of the microbiome on species level to determine whether the microbial community composition and complexity differed among the study groups. We first analysed the alpha diversity. The measure of richness, which reflects the total number of unique bacterial taxa present in a sample, demonstrated a significant difference between the low-risk group and the STEMI group (mean 66.8, SD = 11.1 vs. mean 72.2, SD = 10.5, $P = 0.034$) and the high-risk group and the STEMI group (mean 66.2, SD = 13.9 and mean 72.2, SD = 10.5, $P = 0.047$), also when correcting for sequencing read depth. For the intermediate risk group vs. other groups, no significant differences were observed. Our analysis using the Shannon and Simpson's diversity measures, which not only accounted for taxa richness but also for evenness, did not reveal any differences in alpha diversity between groups, suggesting no distinct microbial compositions. Next, using the inter-sample Aitchison distance (beta diversity), we visualized differences in taxonomic composition through PCoA. No evident clustering of groups was observed in the first five principal components (Figure 1). However, when comparing the centroids of the groups in PCs 2–4, we did observe significant differences, suggesting a relation between inter-individual distance and risk groups (Figure 1).

Cardiovascular risk profiles, based on the Framingham score, explain part of the variance in the gut microbiome

We subsequently conducted PERMANOVA analyses to test the proportion of variance of the gut microbiome that could be explained by the cardiovascular risk profile. In a multivariable analysis, in which we included the covariates PPI use and sequence read depth, risk profile explained a significant proportion of the microbiome inter-sample distance variance ($R^2 = 0.011$, $P \leq 0.001$). This indicates that there are significant differences in microbial composition between groups. As a *post hoc* analysis, we evaluated the individual variables of the Framingham risk score to determine if the results we observed were driven by a specific variable. Here, we found that four out of eight risk-score variables significantly explained variance in taxonomic composition; diabetes ($R^2 = 0.007$, $P \leq 0.001$), age ($R^2 = 0.006$, $P \leq 0.001$), smoking ($R^2 = 0.004$, $P = 0.022$) and HDL ($R^2 = 0.003$, $P = 0.024$).

Decreased abundance of anti-inflammatory species and increased abundance of lipid-metabolism-associated species in the STEMI group compared to groups with lower cardiovascular risk

We used linear models to investigate trends in the relative abundances of species across cardiovascular risk profiles with increasing risk. Here we found three species that showed a positive association with increasing cardiovascular risk (*Collinsella stercoris,* coeff = 0.27, FDR = 0.049; *Flavonifractor plautii,* coeff = 0.53, FDR = 0.039; and *Ruthenibacterium*

### Table 1 Clinical characteristics of the different groups

Phenotype	Low-risk group (130)	Intermediate-risk group (130)	High-risk group (125)	STEMI group (26)	Overall	FDR (BH)—Kruskal Wallis
Age (years)	59.9 (7.57)	62.3 (7.62)	65.4 (6.29)	63.5 (11.5)	62.6 (7.82)	<0.001
Sex (male)	100 (76.9%)	130 (100%)	125 (100%)	22 (84.6%)	377 (91.7%)	<0.001
Body mass index (kg/m^2)	26.8 (3.94)	26.0 (3.97)	27.5 (3.42)	26.7 (3.71)	26.8 (3.82)	0.001
Systolic blood pressure (mmHg)	131 (14.6)	142 (16.1)	153 (17.8)	127 (19.5)	141 (18.7)	<0.001
Diastolic blood pressure (mmHg)	75.9 (8.31)	80.2 (10.2)	82.7 (9.90)	81.8 (12.7)	79.7 (10.1)	<0.001
Heart rate (beats/min)	66.3 (9.70)	68.4 (12.2)	70.8 (13.7)	78.0 (16.4)	69.1 (12.6)	0.002
Smoking	50 (38.5%)	73 (56.2%)	91 (72.8%)	17 (65.4%)	231 (56.2%)	<0.001
Diabetes	0 (0%)	8 (6.2%)	34 (27.2%)	0 (0%)	42 (10.2%)	<0.001
Hypercholesterolaemia	0 (0%)	35 (26.9%)	51 (40.8%)	4 (15.4%)	90 (21.9%)	<0.001
Hypertension	0 (0%)	39 (30.0%)	59 (47.2%)	4 (15.4%)	102 (24.8%)	<0.001
Artery disease	0 (0%)	1 (0.8%)	4 (3.2%)	2 (7.7%)	7 (1.7%)	0.023
CVA	0 (0%)	4 (3.1%)	3 (2.4%)	0 (0%)	7 (1.7%)	0.240
Leucocytes (10^9/L)	5.71 (1.47)	6.37 (1.71)	6.98 (1.61)	11.5 (3.50)	6.67 (2.23)	<0.001
Creatinine (µmol/L)	87.0 (13.7)	91.2 (12.6)	95.1 (17.2)	77.7 (11.2)	90.2 (15.1)	<0.001
Total cholesterol (mmol/L)	5.32 (0.839)	5.47 (0.975)	5.31 (1.13)	5.18 (0.851)	5.36 (0.980)	0.325
HDL (mmol/L)	1.53 (0.376)	1.24 (0.303)	1.09 (0.325)	1.12 (0.320)	1.28 (0.380)	<0.001
LDL (mmol/L)	3.53 (0.768)	3.77 (0.864)	3.55 (1.05)	3.93 (0.821)	3.64 (0.902)	0.033
Triglyceride (mmol/L)	1.11 (0.562)	1.69 (1.05)	2.25 (2.20)	1.19 (0.912)	1.65 (1.47)	<0.001
Proton pump inhibitor	0 (0%)	17 (13.1%)	17 (13.6%)	2 (7.7%)	36 (8.8%)	<0.001
Beta-blocker	0 (0%)	11 (8.5%)	29 (23.2%)	1 (3.8%)	41 (10.0%)	<0.001
Angiotensin antagonist	0 (0%)	8 (6.2%)	14 (11.2%)	0 (0%)	22 (5.4%)	0.001
Calcium antagonist	0 (0%)	6 (4.6%)	10 (8.0%)	0 (0%)	16 (3.9%)	0.103
ACE inhibitor	0 (0%)	11 (8.5%)	17 (13.6%)	2 (7.7%)	30 (7.3%)	0.001
Antithrombotic	0 (0%)	13 (10.0%)	15 (12.0%)	0 (0%)	28 (6.8%)	0.001
P2Y12 inhibitors	0 (0%)	1 (0.8%)	0 (0%)	0 (0%)	1 (0.2%)	0.556
Diuretics (low ceiling)	0 (0%)	6 (4.6%)	20 (16.0%)	0 (0%)	26 (6.3%)	<0.001
Antiarrhythmics	0 (0%)	2 (1.5%)	0 (0%)	0 (0%)	2 (0.5%)	0.250
Lipid-lowering meds	0 (0%)	0 (0%)	4 (3.2%)	0 (0%)	4 (1.0%)	0.033
Statin	0 (0%)	10 (7.7%)	24 (19.2%)	0 (0%)	34 (8.3%)	<0.001
Oral antidiabetic	0 (0%)	6 (4.6%)	18 (14.4%)	0 (0%)	24 (5.8%)	<0.001
Insulin	0 (0%)	1 (0.8%)	5 (4.0%)	0 (0%)	6 (1.5%)	0.047
Inhalation medication	0 (0%)	1 (0.8%)	1 (0.8%)	0 (0%)	2 (0.5%)	0.746
Reads (×10^6)	26.4 (5.6)	26.8 (3.73)	26.4 (3.3)	19.6 (4.05)	26.1 (4.62)	<0.001

Clinical characteristics of all groups. Values for categorical variables = counts (proportion, %) for the presence of that phenotype and for numerical variables = mean (SD).
CVA, cerebral vascular accident; HDL, high density lipoprotein; LDL, low density lipoprotein; ACE, angiotensin-converting enzyme.

lactatiformans, coeff = 0.39, FDR = 0.003), while *Streptococcus thermophilius* showed a negative association (coeff = −0.33, FDR = 0.0499).

We then performed a differential abundance analysis using linear models in which we tested six pairwise comparisons between all groups and one comparison of the low, intermediate and high-risk groups grouped together vs. the STEMI group. While our previous analysis tested for taxa that linearly increased or decreased their abundance between groups, this analysis allows for non-linear changes between groups. In total, we found eight unique species showing differential abundance between cardiovascular risk groups (FDR < 0.05, *Figure 2*). Among these species, certain ones showed differential abundance in multiple comparisons, and two species were overlapping with those detected in the linear analysis. We found two species to be differently abundant in the low risk vs. STEMI comparison (*R. lactatiformans* FDR = 0.029 and *C. stercoris* FDR = 0.008), six species when comparing intermediate risk vs. STEMI (*Rothia mucilaginosa* FDR = 0.049, *Odoribacter splanchnicus* FDR = 0.049,

C. stercoris FDR = 0.005, *Bacteroides vulgatus* FDR = 0.008, *Bacteroides uniformis* FDR = 0.008 and *Agathobaculum butyriciproducens* FDR = 0.033), four species when comparing high risk vs. STEMI (*Ruminococcaceae bacterium D5* FDR = 0.044, *C. stercoris* FDR = 0.019, *B. vulgatus* FDR = 0.014, and *B. uniformis* FDR = 0.019) and four species were differently abundant when comparing STEMI vs. all other groups. Focusing on the species that were differently abundant in the STEMI group compared to all other groups, we observed *C. stercoris*, a species that has been associated with cholesterol metabolism (displaying a negative correlation with serum HDL-cholesterol and a positive correlation with triglycerides[18]), to be enriched in the STEMI group (FDR = 0.005). *Agathobaculum butyriciproducens*, a butyric acid-producing bacteria, was also enriched in the STEMI group compared to other groups (FDR = 0.037). A lower relative abundance of *Bacteroides* species (*B. vulgatus* FDR = 0.013, *B. uniformis* FDR = 0.019), known for their ability to produce SCFAs, was observed in the STEMI group compared to all other groups.
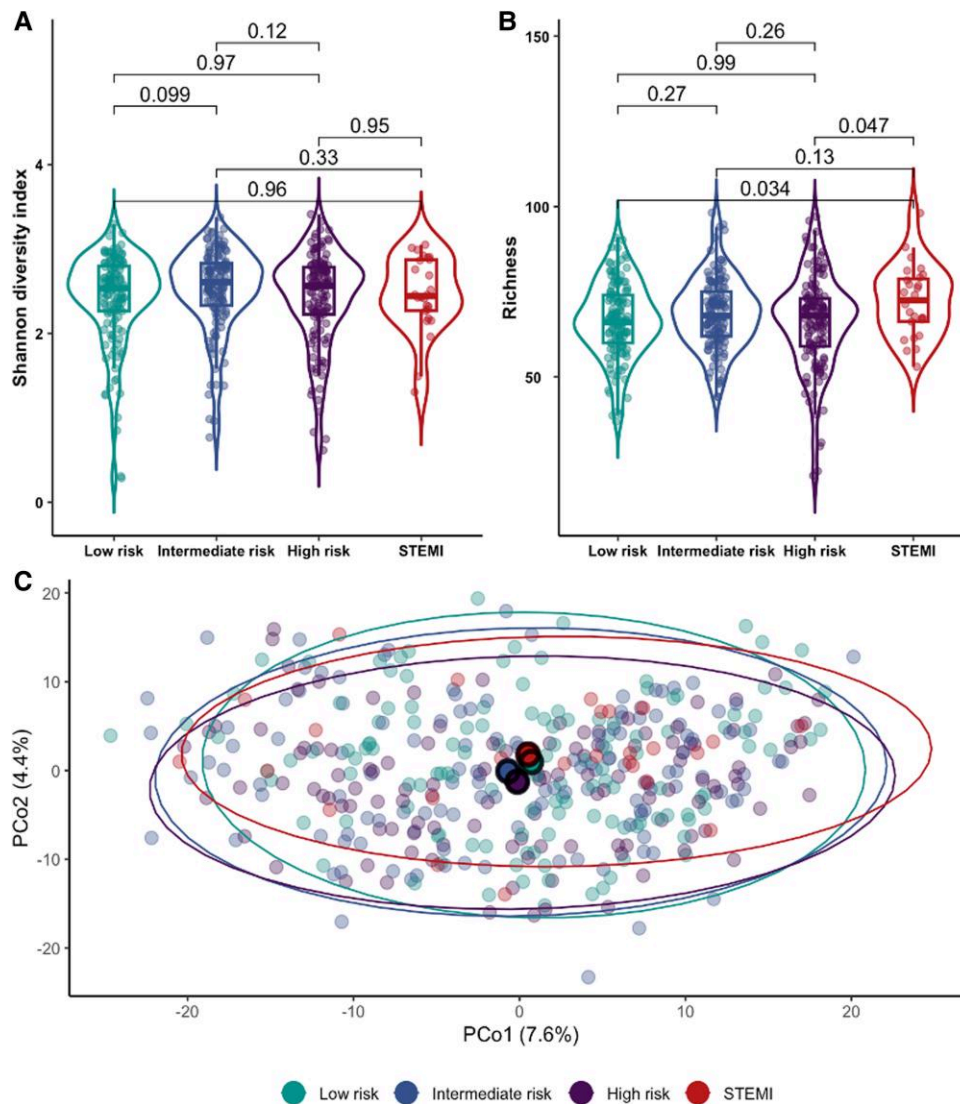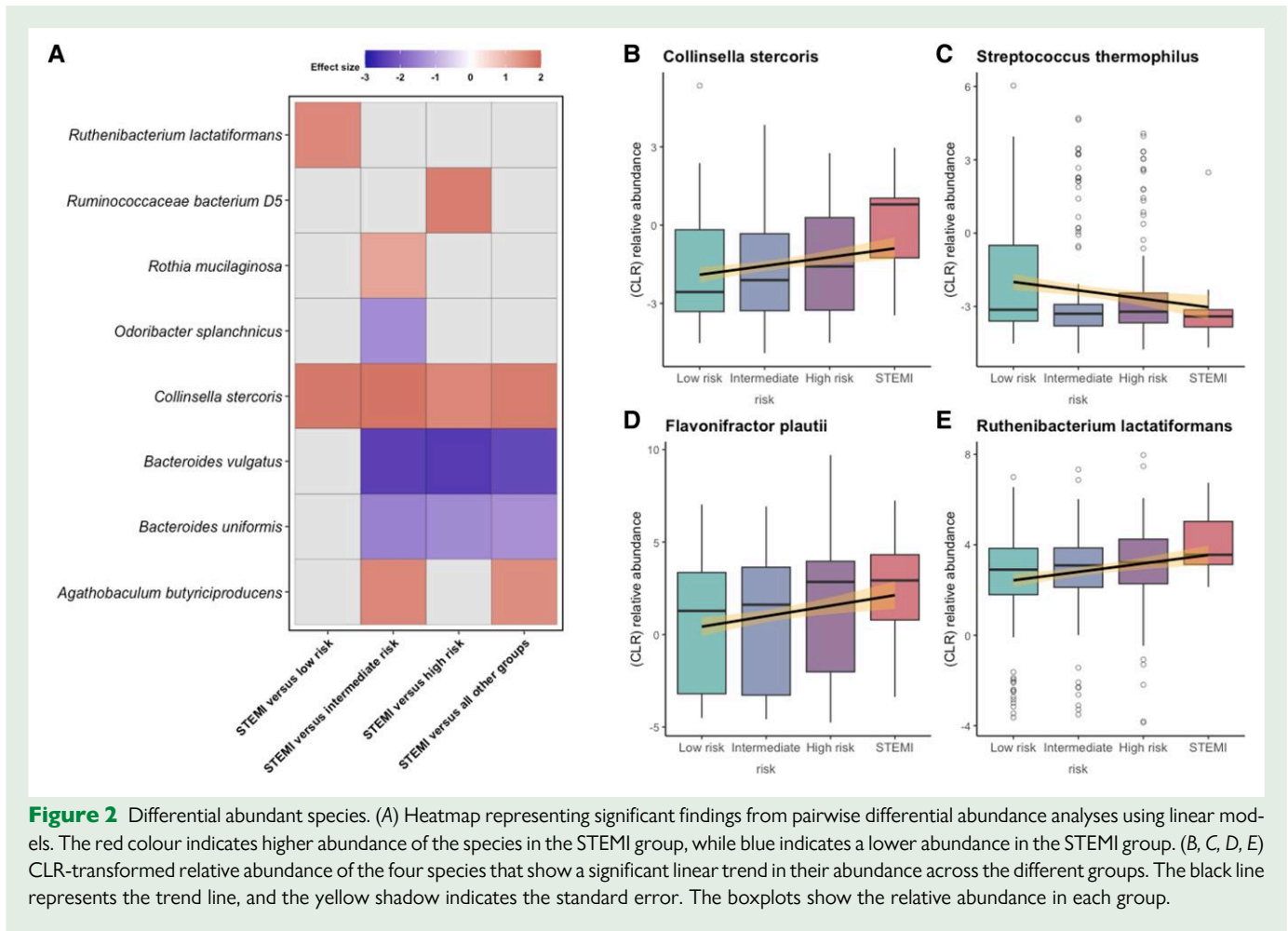
Figure 1 Alpha and beta diversity measures across groups with varying cardiovascular risk. (*A*) Violin plots illustrating the Shannon diversity index for four groups: low risk, intermediate risk, high risk, and STEMI. The *P*-values of the pairwise Wilcoxon tests are shown, which show no significant differences between groups. (*B*) Violin plots that show the richness of each group. The *P*-values resulting from Wilcoxon tests for each comparison are indicated. Low risk vs. STEMI and high risk vs. STEMI are significantly different ($P = 0.034$ and $P = 0.047$). (*C*) Principal coordinates plot depicting the Aitchison distances at species level, using CLR-transformed relative abundance data. Individual data points are represented by dots. The large circle represents the centroid, which represents the mean composition of each group.

Multiple pathways related to vitamin biosynthesis and amino acid biosynthesis are decreased in the STEMI group

To examine the functional potential of the microbial community in individuals with varying levels of cardiovascular risk, we analysed the relative abundance of metabolic pathways. We selected pathways that were present in ≥10% of the samples and used a linear model for pairwise comparisons between groups. Of the 408 pathways identified, 49 unique pathways had distinct differences in abundance between the different cardiovascular risk groups (*Figure 3*). Of these, only four pathways were increased in the STEMI group in comparison to groups with lower risk, while the majority of the pathways were decreased

in the STEMI group in contrast to other groups. These decreased pathways were associated with many processes such as nucleotide biosynthesis, fatty acid, and lipid biosynthesis and carbohydrates biosynthesis. We observed a decrease in four amino acid biosynthesis pathways in the STEMI group. These pathways are involved in: (i) L-citrulline biosynthesis, linked to nitric oxide synthesis and suggested to reduce blood pressure through supplementation,[19] (ii) aspartate/asparagine biosynthesis, shown to promote production of inflammatory cytokines,[20] (iii) L-orthinine biosynthesis, linked to a higher incidence of cardiovascular death when the L-arginine/L-orthinine ratio is low,[21] and (iv) L-lysine biosynthesis, where a previous study found that individuals with pre-existing type 2 diabetes and elevated lysine levels are at the highest risk of developing cardiovascular disease.[22] In addition,

Figure 2 Differential abundant species. (A) Heatmap representing significant findings from pairwise differential abundance analyses using linear models. The red colour indicates higher abundance of the species in the STEMI group, while blue indicates a lower abundance in the STEMI group. (B, C, D, E) CLR-transformed relative abundance of the four species that show a significant linear trend in their abundance across the different groups. The black line represents the trend line, and the yellow shadow indicates the standard error. The boxplots show the relative abundance in each group.

we found six vitamin biosynthesis pathways decreased in the STEMI group including vitamin B7 (biotin), vitamin B1 derivatives, and vitamin B6.

Discussion

In this study, we explored the microbial composition and functional potential of the gut microbiome in groups with varying degrees of cardiovascular risk and STEMI patients. Our analysis revealed four species that showed a significant linear trend in their abundance from low-risk individuals to STEMI patients. *Collinsella stercoris*, *F. plautii*, and *R. lactatiformans* showed an increasing abundance with higher cardiovascular risk, whereas *S. thermophilius* showed a decreasing abundance. Furthermore, we found eight species and 49 pathways to be differently abundant among the groups. Regarding alpha diversity measures, we observed differences in richness between the STEMI group and both the low- and high-risk groups. Variations in alpha diversity and richness have been linked to health conditions before, suggesting that the pre-existing health conditions of the STEMI patients might affect this finding. Nevertheless, it is important to note that another study within the DMP cohort revealed no significant correlations between the pre-existing health conditions considered in our study and alpha diversity.[10] Furthermore, this finding of increased richness was not consistent in the other alpha diversity measures, and the biological relevance remains uncertain. Additionally, the risk groups accounted for a significant proportion of the microbiome inter-sample distance variance.

An intriguing finding is the gradual increase in the abundance of the species *C. stercoris* in groups with increasing cardiovascular risk.

As previously reported, patients with high-cardiovascular risk show an increased abundance of the genus *Collinsella*.[23] These microbes are believed to modulate host metabolism by affecting intestinal cholesterol absorption, decreasing glycogenesis in the liver and increasing triglyceride synthesis.[5] The role of *Collinsella* in cholesterol metabolism was confirmed by a study that demonstrated a positive correlation between serum cholesterol and *Collinsella*.[24] Given that hypercholesterolaemia is a well-established risk factor for CAD, an increased abundance of *Collinsella* could contribute to the risk of developing CAD by promoting hypercholesterolaemia. Nevertheless, not all reports on *Collinsella* in CAD present consistent findings; a reduction in the abundance of *Collinsella* in coronary heart disease patients has also been reported, but this was suggested to be due to treatment with statins.[25] In our study, some participants received statin treatment; however, a previous study found little effect of statins on the overall composition of the microbiome, and these effects were not specifically associated with *Collinsella*.[26]

Bacterial species that were significantly decreased in the STEMI group included *B. vulgatus*, *B. uniformis*, and *O. splanchnicus*. These findings are consistent with previous studies that reported decreased *Bacteroides* abundance in patients with CAD.[25] *Bacteroides* are common microbes in the gut and are considered commensal bacteria. It has been suggested that *B. vulgatus* has an anti-inflammatory effect by affecting microbial lipopolysaccharide synthesis.[27] *Bacteroides uniformis* has been studied in a mouse model, where its administration was found to reduce serum cholesterol and triglyceride levels, improve glucose metabolism and insulin sensitivity, reduce leptin levels and enhance immune function.[28] *Odoribacter splanchnicus* is a common SCFA producer,

Figure 3 Differential abundant microbial predicted pathways across groups. Linear models analysis identified 49 predicted microbial pathways that showed significant alterations in relative abundance between groups (FDR < 0.05). The heatmap displays the effect size, where red indicates an increased abundance in the STEMI group, while blue indicates a decreased abundance. Additionally, the left-side colour-coded columns show the functional classes and categories to which each pathway belongs.

and a reduced species abundance has been associated with diseases such as metabolic syndrome.[29] Higher abundance of the genus *Odoribacter* has been linked to lower blood pressure in overweight and obese pregnant women.[30]

In addition, in STEMI patients, we observed an enrichment of species belonging to the *Oscillospiraceae* family (previously names *Ruminococcacaea*), which includes bacteria known for their butyrate production. To the best of our knowledge, the three species identified in this study that belong to this family (*R. lactatiformans, A. butyriciproducens*, and *Ruminococcaceae bacterium D5*) have not previously been described in relation to cardiovascular disease,

although other members of this family have been reported in that context. Previous studies have reported contradictory findings regarding the abundance of the *Oscillospiraceae* family. Some studies found increased abundance of species of this family in patients with severe (cardiac) disease, while others noted decreased abundance in CAD patients.[25,31,32] This discrepancy could reflect individual species within the same family having distinct functions. Additionally, a novel finding of our study is the increased abundance in STEMI patients of *Rothia mucilaginosa*, a pathogenic species typically found in the oral cavity and has been associated with many infections such as meningitis and endocarditis.[33] To increase our understanding of the differentially abundant

species identified in this study, it is crucial that future studies are conducted to replicate these findings.

We used metagenomic data to examine the functional potential of the microbiome in each group. We found that in the STEMI group, compared to the other risk groups, microbial pathways related to carbohydrate biosynthesis, fermentation, and energy metabolism were enriched. Of the pathways decreased in the STEMI group, many were related to nucleotides, vitamin, amino acids and carbohydrate biosynthesis. One pathway of interest is the *pyridoxal 5'-phosphate biosynthesis I* pathway, which is involved in the biosynthesis of vitamin B6. Vitamin B6 is an essential micronutrient known for its anti-inflammatory properties and is involved in more than 160 different biochemical processes. Previous studies have reported significant associations between low vitamin B6 levels and increased risk of coronary heart disease. More specifically, a significant inverse linear relation of plasma pyridoxal 5'-phosphate (the active coenzyme form of vitamin B6) levels and risk of coronary heart disease in women has been reported.[34] However, it should be emphasized that the decreased abundance of this pathway in the STEMI group is for the microbial metabolic pathway not the human pathway. We did not assess vitamin B6 or pyridoxal 5'-phosphate levels in the current study. As a result, it remains unknown if the observed microbial pathway abundance correlates with reduced serum levels in the host.

This study has several strengths. Firstly, we used shotgun sequencing to obtain metagenomic data, which provides improved taxonomic resolution compared to the more commonly used 16S rRNA gene sequencing techniques. Furthermore, this approach allowed us to investigate functional profiles of metabolic pathways of the microbial communities. Secondly, the inclusion of participants from a general population cohort for the risk groups ensures that selection bias is minimized. In addition, the availability of comprehensive metadata from this cohort made it possible to calculate risk scores. Thirdly, a notable strength is that the STEMI group had minimal medication usage, as cardiovascular medication was only started after the acute event and sample collection. This represents a common scenario in which patients presenting with STEMI are often not previously known by a cardiologist, resulting in the absence of medication regimens before the event.

However, our study also has some limitations that should be addressed. Firstly, our STEMI group had a relatively small sample size of 26 patients. Future studies would benefit from a larger sample size for the groups to increase statistical power. Furthermore, it may be interesting to include other acute coronary syndrome conditions, such as non-STEMI and unstable angina pectoris to the spectrum of cardiovascular risk groups. Secondly, the intermediate- and high-risk groups included in the study consisted only of male participants. While it is recognized that cardiovascular disease disproportionally affects men in the general population, with a ratio of approximately 2 to 1,[35] the context of age is essential to consider. In our study, all included women were aged 50 or older. Given that oestrogen levels decline after menopause, impacting cardiovascular risk, future studies should explore alternative cohorts that included women across a broader range of age groups with cardiovascular risk to make the findings more generalizable. Thirdly, the complex and multifactorial nature of the crosstalk between the gut microbiome and host in CAD makes it difficult to determine causality when using *in silico* analyses alone. Therefore, future studies should include experimental designs, such as randomized trials with interventions and culturing of faecal bacteria to provide further insights into the directional and causal relationships of the gut microbiome and CAD. Lastly, we are aware that diet and lifestyle can significantly influence the gut microbiome, but the present study lacked data on these factors and we could not adjust for them. However, we did account for BMI, which also reflects important aspects of lifestyle and diet. Additionally, the participants from both cohorts were selected from the same population in the northern Netherlands, where dietary habits have been shown to be largely stable.[10] This minimizes potential biases associated with variations in diet and lifestyle.

Overall, the aetiology of myocardial infarction is multifaceted, involving stochastic, genetic, and environmental factors. Understanding these mechanisms is crucial for developing new targets for prevention and therapy. We broadened the scope of investigation beyond conventional cardiovascular risk factors, hypothesizing the gut microbiome's involvement in STEMI. Therefore, we explored if dysbiosis signatures of the gut microbiome seen in CAD/STEMI patients could be detected in individuals at risk. We identified four species to show a gradual trend in abundance with increasing cardiovascular risk and eight species to be differently abundant in the STEMI group compared to individuals at risk. However, with this study we are unable to disentangle yet one of the most intriguing questions within this field about whether dysbiosis precedes disease or is a result of disease. To gain a deeper understanding of this, future studies should aim to clarify the causal direction. Possible approaches would include conducting longitudinal follow-up within the DMP, conducting animal studies, and exploring organ-on-a-chip technology. Our results contribute to the existing body of research on the relationship between CAD and the gut microbiome and offer insight into potential therapeutic directions for future mechanistic and interventional studies.

Supplementary material

Supplementary material is available at *European Journal of Preventive Cardiology*.

Authors' contribution

F.P., H.E., and P.v.H., contributed to the conception and design of the work. F.P., H.E., and B.H. contributed to the acquisition of data for the work. F.P., S.A.S., R.G., J.S., and J.B. contributed to the analysis. F.P., H.E., V.C., and R.G. contributed to the interpretation of the data. F.P., H.E., and V.C. drafted the manuscript. J.F., H.H., A.Z., E.L., P.v.H., R.K., and R.G. critically reviewed the manuscript.

Data availability

Codes used for this study can be found at: https://github.com/GRONINGEN-MICROBIOME-CENTRE/STEMI_microbiome.git. The data from the CardioLines cohort underlying this article can be shared upon reasonable request to the corresponding author. The raw microbiome sequencing data, processed microbiome data, and basic phenotypes of the DMP cohort are available at the European Genome-Phenome Archive under accession EGAS00001005027 (https://ega-archive.org/studies/EGAS00001005027). Access to this dataset requires a minimal access procedure. Access to other phenotypic data of the DMP cohort is subject to approval by the LifeLines Data Access Committee. If researchers wish to request access to the DMP data, they are required to fill out an application form, which can be found at: https://www.lifelines.nl/researcher/how-to-apply.

References

1. Rinninella E, Raoul P, Cintoni M, Franceschi F, Miggiano G, Gasbarrini A, *et al.* What is the healthy gut Microbiota composition? A changing ecosystem across age, environment, diet, and diseases. *Microorganisms* 2019;**7**:14.

2. Trøseid M, Andersen GØ, Broch K, Hov JR. The gut microbiome in coronary artery disease and heart failure: current knowledge and future directions. *eBioMedicine* 2020;**52**:102649.

3. Mansuri NM, Mann NK, Rizwan S, Mohamed AE, Elshafey AE, Khadka A, *et al.* Role of gut microbiome in cardiovascular events: a systematic review. *Cureus* 2022;**14**:e32465.

4. Witkowski M, Weeks TL, Hazen SL. Gut microbiota and cardiovascular disease. *Circ Res* 2020;**127**:553–570.

5. Choroszy M, Litwinowicz K, Bednarz R, Roleder T, Lerman A, Toya T, *et al.* Human gut microbiota in coronary artery disease: a systematic review and meta-analysis. *Metabolites* 2022;**12**:1165.

6. Groot HE, van de Vegte YJ, Verweij N, Lipsic E, Karper JC, van der Harst P. Human genetic determinants of the gut microbiome and their associations with health and disease: a phenome-wide association study. *Sci Rep* 2020;**10**:14771.

7. Kwun JS, Kang SH, Lee HJ, Park H-K, Lee W-J, Yoon C-H, *et al.* Comparison of thrombus, gut, and oral microbiomes in Korean patients with ST-elevation myocardial infarction: a case–control study. *Exp Mol Med* 2020;**52**:2069–2079.

8. Gao J, Yan KT, Wang JX, Dou J, Wang J, Ren M, *et al.* Gut microbial taxa as potential predictive biomarkers for acute coronary syndrome and post-STEMI cardiovascular events. *Sci Rep* 2020;**10**:2639.

9. Baldelli V, Scaldaferri F, Putignani L, Del Chierico F. The role of Enterobacteriaceae in gut microbiota dysbiosis in inflammatory bowel diseases. *Microorganisms* 2021;**9**:697.

10. Gacesa R, Kurilshikov A, Vich Vila A, Sinha T, Klaassen MAY, Bolte LA, *et al.* Environmental factors shaping the gut microbiome in a Dutch population. *Nature* 2022;**604**:732–739.

11. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–359.

12. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Turnbaugh P, Franco E, Brown CT, eds. *eLife* 2021;**10**:e65088.

13. Castro V. CVrisk: Compute Risk Scores for Cardiovascular Diseases. Version 1.1.0. 2021. https://github.com/vcastro/CVrisk (15 June 2023).

14. Ho D, Imai K, King G, Stuart EA. Matchit: nonparametric preprocessing for parametric causal inference. *J Stat Softw* 2011;**42**:1–28.

15. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, *et al.* vegan: Community Ecology Package. R package version 2.5-7. 2022. https://CRAN.R-project.org/package=vegan (15 June 2023).

16. Swarte JC, Li Y, Hu S, Björk JR, Gacesa R, Vich Vila A, *et al.* Gut microbiome dysbiosis is associated with increased mortality after solid organ transplantation. *Sci Transl Med* 2022;**14**:eabn7566.

17. Russell V. Lenth. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.6. 2023, https://CRAN.R-project.org/package = emmeans (15 June 2023).

18. Astbury S, Atallah E, Vijay A, Aithal GP, Grove JI, Valdes AM. Lower gut microbiome diversity and higher abundance of proinflammatory genus Collinsella are associated with biopsy-proven nonalcoholic steatohepatitis. *Gut Microbes* 2020;**11**:569–580.

19. Allerton TD, Proctor DN, Stephens JM, Dugas TR, Spielmann G. Irving BA. l-citrulline supplementation: impact on cardiometabolic health. *Nutrients* 2018;**10**:921.

20. Wang H, Zheng X, Liu B, Xia Y, Xin Z, Deng B, *et al.* Aspartate metabolism facilitates IL-1β production in inflammatory macrophages. *Front Immunol* 2021;**12**:753092.

21. Ishinoda Y, Masaki N, Hitomi Y, Taruoka A, Kawai A, Iwashita M, *et al.* A low arginine/ornithine ratio is associated with long-term cardiovascular mortality. *J Atheroscler Thromb.* Published online 11 February 2023;**30**:1364–1375.

22. Razquin C, Ruiz-Canela M, Clish CB, Li J, Toledo E, Dennis C, *et al.* Lysine pathway metabolites and the risk of type 2 diabetes and cardiovascular disease in the PREDIMED study: results from two case-cohort studies. *Cardiovasc Diabetol* 2019;**18**:151.

23. Karlsson FH, Fåk F, Nookaew I, Tremaroli V, Fagerberg B, Petranovic D, *et al.* Symptomatic atherosclerosis is associated with an altered gut metagenome. *Nat Commun* 2012;**3**:1245.

24. Lahti L, Salonen A, Kekkonen RA, Salojärvi J, Jalanka-Tuovinen J, Palva A, *et al.* Associations between the human intestinal microbiota, *Lactobacillus rhamnosus* GG and serum lipids indicated by integrated analysis of high-throughput profiling data. *PeerJ* 2013;**1**:e32.

25. Zhang Y, Xu J, Wang X, Ren X, Liu Y. Changes of intestinal bacterial microbiota in coronary heart disease complicated with nonalcoholic fatty liver disease. *BMC Genomics* 2019;**20**:862.

26. Vila AV, Collij V, Sanna S, Sinha T, Imhann F, Bourgonje AR, *et al.* Impact of commonly used drugs on the composition and metabolic function of the gut microbiota. *Nat Commun* 2020;**11**:362.

27. Yoshida N, Emoto T, Yamashita T, Watanabe H, Hayashi T, Tabata T, *et al. Bacteroides vulgatus* and *Bacteroides dorei* reduce gut microbial lipopolysaccharide production and inhibit atherosclerosis. *Circulation* 2018;**138**:2486–2498.

28. Gauffin Cano P, Santacruz A, Moya Á, Sanz Y. Bacteroides uniformis CECT 7771 ameliorates metabolic and immunological dysfunction in mice with high-fat-diet induced obesity. *PLoS One* 2012;**7**:e41079.

29. Hiippala K, Barreto G, Burrello C, Diaz-Basabe A, Suutarinen M, Kainulainen V, *et al.* Novel *Odoribacter splanchnicus* strain and its outer membrane vesicles exert immunoregulatory effects in vitro. *Front Microbiol* 2020;**11**:575455.

30. Gomez-Arango LF, Barrett HL, McIntyre HD, Callaway LK, Morrison M, Dekker Nitert M, *et al.* Increased systolic and diastolic blood pressure is associated with altered gut microbiota composition and butyrate production in early pregnancy. *Hypertens Dallas Tex* 2016;**68**:974–981.

31. Kim M, Huda MN, Bennett BJ. Sequence meets function—microbiota and cardiovascular disease. *Cardiovasc Res* 2022;**118**:399–412.

32. Liu H, Chen X, Hu X, Niu H, Tian R, Wang H, *et al.* Alterations in the gut microbiome and metabolism with coronary artery disease severity. *Microbiome* 2019;**7**:68.

33. Haddad S, Saade Y, Ramlawi B, Kreidieh B, Gilbert B, Rao S. Native valve endocarditis complicated by abscess formation caused by *Rothia mucilaginosa. IDCases* 2021;**26**:e01348.

34. Hu L, Li Y, Liu Z, Fan F, Xu B, Xu R, *et al.* Association of plasma vitamin B6 with coronary heart disease in patients undergoing diagnostic coronary angiography: new insight on sex differences. *Front Cardiovasc Med* 2021;**8**:789669.

35. Walli-Attaei M, Joseph P, Rosengren A, Chow CK, Rangarajan S, Lear SA, *et al.* Variations between women and men in risk factors, treatments, cardiovascular disease incidence, and death in 27 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. *Lancet (London, England)* 2020;**396**:97–109.