Gut Microbial Associations to Plasma Metabolites Linked to Cardiovascular Phenotypes and Risk
A Cross-Sectional Study


Rationale: Altered gut microbial composition has been linked to cardiovascular diseases (CVDs), but its functional links to host metabolism and immunity in relation to CVD development remain unclear.

Objectives: To systematically assess functional links between the microbiome and the plasma metabolome, cardiometabolic phenotypes, and CVD risk and to identify diet-microbe-metabolism-immune interactions in well-documented cohorts.

Methods and Results: We assessed metagenomics-based microbial associations between 231 plasma metabolites and microbial species and pathways in the population-based LLDe (Lifelines DEEP) cohort (n=978) and a clinical obesity cohort (n=297). After correcting for age, sex, and body mass index, the gut microbiome could explain ≤11.1% and 16.4% of the variation in plasma metabolites in the population-based and obesity cohorts, respectively. Obese-specific microbial associations were found for lipid compositions in the VLDL, IDL, and LDL lipoprotein subclasses. Bacterial L-methionine biosynthesis and a Ruminococcus species were associated to cardiovascular phenotypes in obese individuals, namely atherosclerosis and liver fat content, respectively. Integration of microbiome-diet-inflammation analysis in relation to metabolic risk score of CVD in the population cohort revealed 48 microbial pathways associated to CVD risk that were largely independent of diet and inflammation. Our data also showed that plasma levels rather than fecal levels of short-chain fatty acids were relevant to inflammation and CVD risk.

Conclusions: This study presents the largest metagenome-based association study on plasma metabolism and microbiome relevance to diet, inflammation, CVD risk, and cardiometabolic phenotypes in both population-based and clinical obesity cohorts. Our findings identified novel bacterial species and pathways that associated to specific lipoprotein subclasses and revealed functional links between the gut microbiome and host health that provide a basis for developing microbiome-targeted therapy for disease prevention and treatment. (Circ Res. 2019;124:1808-1820. DOI: 10.1161/CIRCRESAHA.118.314642.)

Key Words: atherosclerosis • inflammation • metabolomics • metagenomics • microbiota
Novelty and Significance

What Is Known?
- The human gut microbiome consists of over 100 trillion bacteria.
- Gut microbial composition and function varies between individuals.
- Diet-host-microbe interactions play an important role in various complex traits and diseases, including blood lipid levels, obesity, and cardiovascular diseases.

What New Information Does This Article Contribute?
- The human gut microbiome is associated with plasma levels of different lipoprotein subclasses of different sizes and composition, fatty acids, and amino acids.
- A total of 48 bacterial metabolic pathways are associated to metabolic risk of cardiovascular disease, with most involved in metabolism of amino acids, carbohydrates, and poly saccharides.
- The bacterial pathway of L-methionine biosynthesis is linked to lower cIMT.

The human gut microbiome contains trillions of bacteria and is associated with body mass index, blood lipids, and cardiovascular disease risk. However, previous studies, mostly based on 16S ribosome RNA sequencing technology, have provided limited information on microbial function. This study presents a large-scale metagenomics-based association study between gut microbiome and plasma metabolism in 2 deeply phenotyped cohorts: a human population cohort and a clinical obesity cohort. We identified various bacterial species and pathways, for example, metabolism of L-methionine, associated to metabolic risk of cardiovascular disease in human population and atherosclerotic plaques in obese individuals. Our findings may provide a basis for understanding the functional links between diet and bacterial and human metabolism in the development of cardiovascular disease.

Nonstandard Abbreviations and Acronyms

- **300-OB**: 300-Obese
- **AA**: amino acid
- **AAT**: alpha-1 antitrypsin
- **BMI**: body mass index
- **cIMT**: carotid intima-media thickness
- **CVD**: cardiovascular disease
- **FDR**: false discovery rate
- **IL**: interleukin
- **IMT**: intima-media thickness
- **LLD**: Lifelines DEEP
- **MRS**: metabolic risk score
- **NBS**: Nijmegen Biomedical Study
- **NIMA**: Nijmegen Biomedical Study—Non-Invasive Measurements of Atherosclerosis
- **NMR**: nuclear magnetic resonance
- **SCFA**: short-chain fatty acid
- **TMAO**: trimethylamine-N-oxide

in digestion and degradation of nutrients, maintenance of digestive tract integrity, stimulation of the immune system, and modulation of the host metabolism. Recent studies indicate a strong link between the gut microbiota and the development of various human diseases, including obesity, insulin resistance, and type 2 diabetes mellitus, as well as gastrointestinal, autoimmune, and cardiovascular diseases (CVDs). Various studies have provided evidence that host-microbe interactions contribute to the pathogenesis of many of these diseases through their impact on metabolism and immunity. Lower bacterial richness (a reduction in the number of different species or bacterial genes) has been associated with an overall increase in adiposity, insulin resistance, dyslipidemia, and inflammatory phenotypes. The distinct gut microbiome profile found in overweight individuals has been shown to have an increased capacity to harvest nutrients from food. Moreover, the gut microbiome is also associated with an individual’s cytokine production capacity in response to different pathogens. Despite a large body of evidence from cross-sectional association studies, the underlying mechanisms are largely unknown, and several putative mechanisms and functional links have been proposed. For instance, the impact of the gut microbiome on insulin sensitivity and glucose homeostasis may be mediated via microbial biosynthesis of branched-chain amino acids (AA), short-chain fatty acids (SCFAs), and N-acyl amides. The fermentation products of dietary fibers, in particular, SCFAs, also have potential roles in the host’s innate and adaptive immunity through modulation of cell proliferation and differentiation, hormone secretion, G-protein–coupled receptor activation, and regulation of colonic Treg cell homeostasis. Finally, inhibition of gut microbiome–induced trimethylamine-N-oxide (TMAO) production can attenuate atherosclerosis development in mice. However, our understanding of the functions of gut microbes of diet-microbe-metabolism-immune interactions in CVD remains limited, leaving a knowledge gap that greatly delays clinical translation. Evaluating the complex interactions between the gut microbiome, host metabolism, and immune system—as affected by intrinsic host and external factors (diet and medication)—requires multiomics, systems biology–based approaches.

To do so, we examined both a population-representative and an overweight patient cohort, collectively comprising 1275 individuals. The obese cohort was deeply phenotyped for cardiometabolic traits, fat distribution, and plasma level of TMAO, whereas the population-based cohort was deeply phenotyped for inflammation, diet, and SCFAs (Figure 1A). We first aimed to identify microbial species and metabolic pathways associated with plasma metabolite profiles (Figure 1B). We then identified the relevance of these metabolism-related microbial factors to cardiometabolic phenotypes in the obese patient cohort. Finally, we assessed individual metabolic risk of CVD in the population-based cohort and identified bacterial pathways associated to the CVD risk score and assessed the diet-microbe-metabolism-immune interactions in CVD risk (Figure 1B).
Abstract

All metagenomics and metabolomics data have been made publicly available at the European Genomics-Phenome Archive at accession numbers EGAS00001003508 for the LLD (LifeLines DEEP) cohort and EGAS00001001704 for the 300-OB cohort. In LLD and 300-OB, 231 metabolic traits and metagenomics-based microbial features were identified, respectively, and an MWAS explored the relevance of metabolome-associated microbial features to cardiovascular phenotypes. All metagenomics and metabolism data have been made publicly available at the European Genomics-Phenome Archive at accession numbers EGAS00001003508 for the LLD (LifeLines DEEP) cohort and EGAS00001001704 for the 300-OB cohort. (2) An MWAS assessed the relevance of metabolome-associated microbial features to cardiovascular phenotypes. (3) An MWAS identified microbial features associated with the metabolic risk score (MRS) of cardiovascular disease (CVD) in LLD, which was constructed using 33 metabolic biomarkers, and an MWAS explored the relevance of microbial factors associated with the metabolic risk score (MRS) of cardiovascular disease (CVD) in 300-OB. An MWAS assessed the relevance of metabolome-associated microbial features to cardiovascular phenotypes. (4) An integration analysis assessed the relevance of 78 dietary factors, inflammatory cytokines, and SCFAs in microbial association of MRS in LLD.

**Methods**

All metagenomics and metabolomics data have been made publicly available at the European Genomics-Phenome Archive at accession numbers EGAS00001003508 for the LLD (LifeLines DEEP) cohort and EGAS00001001704 for the 300-OB cohort. In LLD and 300-OB, 231 metabolic traits and metagenomics-based microbial features were identified, respectively, and an MWAS explored the relevance of metabolome-associated microbial features to cardiovascular phenotypes. All metagenomics and metabolism data have been made publicly available at the European Genomics-Phenome Archive at accession numbers EGAS00001003508 for the LLD (LifeLines DEEP) cohort and EGAS00001001704 for the 300-OB cohort. In LLD and 300-OB, 231 metabolic traits and metagenomics-based microbial features were identified, respectively, and an MWAS explored the relevance of metabolome-associated microbial features to cardiovascular phenotypes. 

**Cohorts**

**LLD Cohort**

LLD is a subcohort of the large, prospective, population-based LifeLines cohort (167,729 subjects) from the north of the Netherlands. A subset of 1539 participants with deep omics profiling makes up LLD. Participants volunteered to participate in LLD from April to August 2013, with blood and fecal samples collected in the same period (within 2 weeks). High-quality metagenomic data and detailed dietary data (78 dietary factors) information are available for 1135 LLD participants, and 1046 LLD individuals were profiled for plasma metabolites and inflammation data (see sections below). We further excluded 57 participants who were taking antibiotics or were using thrombocyte aggregation inhibitors other than acetylsalicylic acid and carbasalate calcium, or a contraindication for magnetic resonance imaging were excluded from the study. Participants who used lipid-lowering therapy temporarily discontinued this medication 4 weeks before the measurements. All women were postmenopausal and did not use hormonal replacement therapy. All subjects completed questionnaires about lifestyle, medication use, and previous diagnosis of hypertension and diabetes mellitus. For all participants, blood samples for metabolomics analysis were collected in the morning after an overnight fast. All underwent comprehensive assessment of cardiovascular profile and fat distribution, as detailed below. Five samples of metagenomic data failed to pass the quality control, leaving 297 individuals for further analyses.

**300-OB Cohort**

Between 2014 and 2016, 302 individuals aged 55 to 80 years were enrolled in the 300-OB study at the Radboud University Medical Center, Nijmegen, the Netherlands. All had a body mass index (BMI) >27 kg/m², and the majority (n=227) had participated in the NBS-NIMA1 study (Nijmegen Biomedical Study—Non-Invasive Measurements of Atherosclerosis I)—a population-based survey of Nijmegen residents. We recruited another 75 participants, acquaintances of previously included subjects, who fulfilled the inclusion criteria of age >55 years and BMI >27 kg/m². Most of these new participants were unrelated subjects, with only 9 being family members of previously included subjects; we, therefore, did not separately evaluate or incorporate the potential clustering of subjects. Subjects with a recent cardiovascular event (myocardial infarction, transient ischemic attack, and stroke <6 months), a history of bariatric surgery or bowel resection, inflammatory bowel disease, renal dysfunction, increased bleeding tendency, use of oral or subcutaneous anticoagulant therapy, use of thrombocyte aggregation inhibitors other than acetylsalicylic acid and carbasalate calcium, or a contraindication for magnetic resonance imaging were excluded from the study. Participants who used lipid-lowering therapy temporarily discontinued this medication 4 weeks before the measurements. All women were postmenopausal and did not use hormonal replacement therapy. All subjects completed questionnaires about lifestyle, medication use, and previous diagnosis of hypertension and diabetes mellitus. For all participants, blood samples for metabolomics analysis were collected in the morning after an overnight fast. All underwent comprehensive assessment of cardiovascular profile and fat distribution, as detailed below. Five samples of metagenomic data failed to pass the quality control, leaving 297 individuals for further analyses.

The summary characteristics of both cohorts are described in the Table. We noted that there are differences in age, sex, and BMI between cohorts that can represent potential confounders. Therefore, these variables were corrected for in subsequent analyses.

**Cardiovascular Phenotyping in the 300-OB Cohort**

Cardiovascular assessment was performed at Radboud University Medical Center (Online Table 1). Vascular studies included the measurement of carotid intima-media thickness (IMT), carotid plaque presence, and maximum plaque thickness. Measurements were performed after an overnight fast or in the afternoon 6 hours after a standardized
breakfast. Participants were asked to abstain from caffeinated products for at least 12 hours and to not smoke for 12 hours before the visit. Testing was performed in a quiet temperature-controlled room with patients in supine position. After a resting period of at least 30 minutes, baseline resting diameter, distensibility, and wall thickness of the carotid artery were assessed by a well-trained sonographer using a 7.5-MHz transducer of a Mylab class C ultrasound device (Esaote Biomedica, Genoa, Italy) connected to a computer with a data acquisition board (Artlab, Esaote Europe BV, Maastricht, the Netherlands). Ultrasound parameters were set to optimize longitudinal B-mode images of the lumen/arterial wall interface. The carotid IMT and diameter measurements were performed in the proximal 1-cm straight portion of the carotid artery in 3 different angles (90°, 120°, and 180°) for 6 heartbeats. Measurements were recorded during the diastolic phase. Measurement of carotid IMT was performed using an automatic boundary detection system based on radio frequency processing–based measurement (Artlab).29 The primary outcome variable was defined as the mean carotid IMT of the different angles.30

Subsequently the presence of plaque and the maximum thickness of plaques in the common carotid, internal carotid, external carotid artery, and at the carotid bulbus were measured. The presence of plaque was defined as focal thickening of the wall of at least 1.5× the mean IMT or an IMT >1.5 mm, according to the Mannheim IMT consensus.31 Furthermore, fat distribution was assessed using magnetic resonance imaging, including volumes of visceral adipose tissue and subcutaneous adipose tissue, divided into deep and superficial subcutaneous adipose tissue, respectively. Hepatic fat content was quantified using localized proton magnetic resonance spectroscopy (for detailed Methods, see the Online Data Supplement).

### Microbiome Data Profiling

#### Metagenomic Shotgun Sequencing

High-quality metagenomics data were already available for the 1135 LLD participants. For this study, we performed metagenomic sequencing of the 300-OB cohort using a similar protocol and analysis pipeline. In brief, fecal and blood samples were collected within 2 weeks for LLD participants and within 1 to 2 days for 300-OB participants to reduce potential bias introduced by sampling. Further processing of all sample sets was identical to LLD. In brief, DNA was isolated with the AllPrep DNA/RNA Mini Kit (Qiagen, Hilden, Germany; catalog No. 80206) with the addition of mechanical lysis. Metagenomic shotgun sequencing was performed using the Illumina HiSeq platform (illumina, San Diego, CA), with an average of 3.0 Gb data (around 32.3 million reads) obtained per sample. Reads were quality filtered using our in-house pipeline. Sequencing adapters were removed using Trimomatic (v.0.32).32 Human reads were removed by mapping the data to the human reference genome (version NCBI37) with Bowtie2 (v.2.1.0).

#### Identifying Microbial Taxa Abundances

The profile of microbial composition was determined using MetaPhlan 2.2.33 and it reported 1772 microbial taxonomies in our data. We normalized the taxonomy data using log- and inverse rank-sum transformation and further corrected for age and sex with linear regression in R (v.3.4.3). We confined the analysis to the 188 common microbial species (Online Table II) present in >10% of samples.

#### Identifying Abundances of Bacterial Metabolic Pathways

The abundance of metabolic pathways was determined using HUMAnN2 (http://huttenhower.sph.harvard.edu/humann2), which maps reads to a customized database of functionally annotated pan-genomes. This analysis mapped reads to UniProt Reference Clusters (UniRef50, http://www.uniprot.org) and then further grouped them to 773 pathways from the MetaCyc metabolic pathway database (www.metacyc.org). Only pathways present in >25% of samples (562 pathways; Online Table III) were used in our downstream analysis. For the nonzero gene counts per MetaCyc pathway, we performed log- and inverse rank-sum transformations, followed by correction for the effects of age and sex using linear regression.

## Plasma Metabolome Profiling

For both cohorts, we profiled a wide range of plasma metabolites using nuclear magnetic resonance (NMR) and Nightingale Biomarker Analysis Platform.34 This platform provides measures of 231 plasma metabolite traits (Online Table IV), including total lipid concentrations and relative compositions of 14 lipidprotein subclasses, lipoprotein particle sizes, apolipoproteins, cholesterol, glycerides and phospholipid concentrations, various glycolysis components, fatty acid composition, inflammation, fluid balance, ketone bodies, and AAs. The NMR metabolomics platform has recently been used in several epidemiological, genome-wide association, and functional genetic studies.35–37 To further validate platform precision, we compared NMR measures of several traits with corresponding routine lipid measurements, including concentrations of HDL, LDL, triglycerides, and total cholesterol (Online Figure I). For these traits, we observed high correlation rates (R>0.89), in agreement with earlier platform validation results.37 All the metabolomics measurements in LLD samples were performed in 1 batch, whereas the 300-OB samples were randomized and measured in 2 batches. The 300-OB data were, therefore, corrected for batch effects using linear regression before performing further analyses. There were 2.2% missing values across all data. Given the high correlation structure of metabolites, missing values could be imputed using the Principal Components Imputation method implemented in the missMDA package (v.1.1.2) for R, using the first 10 principal components. Before microbiome-metabolome association analysis, because of nonlinear dependency of metabolomic traits to covariates, we used locally weighted scatterplot smoothing to correct for sex-dependent age and BMI effects.

### Adipokine and Cytokine Profiling in LLD

The panel of cytokines for the LLD cohorts (IL [interleukin]-1β, IL-6, IL-8, IL-10, IL-12p70, and TNF-α [tumor necrosis factor-α]) was measured by ProcartaPlex multiplex immunoassay (ebioscience, San Diego, CA) according to protocols described before.38 Other inflammation markers were measured using commercially available sandwich ELISA kits (R&D Systems, Minneapolis, MN), including leptin, adiponectin, IL-18, IL-18BP, resistin, and AAT (alpha-1 antitrypsin).

### SCFA Profiling in LLD

For LLD, we measured fecal levels of acetate, propionate, butyrate, valerate, and caproate by gas chromatography–mass spectrometry following the method of Garcíía-Villalba et al.40 The abundance of acetate in plasma was obtained through plasma metabolome profiling using NMR. All SCFA measurements were corrected for age and sex using locally weighted scatterplot smoothing.
TMAO, Choline, Betaine, and Citrulline Profiling in 300-OB

TMAO, choline, L-carnitine, betaine, and γ-butyrobetaine in plasma were analyzed by ultra-high performance liquid chromatography in combination with isotope dilution tandem mass spectrometry. In short, 25 μL plasma was pipetted into 96-well plates, 25 μL internal standard solution was added (containing TMAO-D9, choline-D9, L-carnitine-D3, and betaine-D11), followed by 300 μL 80% acetonitrile and 1% formic acid in Millipore water. Mass spectrometric detection was performed on a XEVO TQ-s system (Waters). Analytes were detected in positive mode and selected reaction monitoring mode. The respective quantifier ion transitions were as follows: m/z 76.15>58.3 for TMAO, m/z 104.2>60.3 for choline, m/z 162.2>103.25 for L-carnitine, m/z 118.2>59.3 for betaine, and m/z 146.25>60.3 for γ-butyrobetaine. All analytes were baseline separated from each other.

Statistical Analysis

All statistical analyses were performed using the R statistical language (v.3.4.3).

Microbiome-Metabolome Association and Explained Variance

(1) To identify the associations between metabolic and microbial factors, association analyses were performed in both LLD and 300-OB. Microbiome features included both microbial pathways and species. After adjusting for age, sex, BMI, and smoking, Spearman correlation of metabolic traits and microbiome features was used to identify associations. The analysis was confined to pairwise nonzero values. False discovery rate (FDR) was estimated using 100× permutations. We also calculated the permutation-based family-wise error rate and report the top association at family-wise error rate <0.05 level. (2) To estimate the proportion of variation in plasma metabolism explained by microbial factors, we used a least absolute shrinkage and selection operator (LASSO) from the R package glmnet (v.2.0.16) that included all identified microbial pathways and species as predictors. The independent and most dominant microbial features were selected automatically, and the variation of each metabolite explained by the selected microbial factors was then estimated by least absolute shrinkage and selection operator. Moreover, we used the same least absolute shrinkage and selection operator algorithm to estimate the proportion of metabolic variation in 300-OB explained by microbial factors identified in LLD.

Association to Cardiometabolic Phenotypes

The bacterial pathways and species identified were further tested for association with cardiometabolic phenotypes in 300-OB, including fat distribution and atherosclerosis phenotypes. The analysis was performed using linear regression, with cardiometabolic phenotype as outcome and microbiome feature as predictor, and treating age, sex, and BMI as covariates. Quantitative outcomes were adjusted using inverse rank-sum transformation; for binary outcomes, logistic regression was used instead of linear regression. Furthermore, to assess to what extent bacteria-derived TMAO and its relevant metabolites could underlie microbiome associations to cardiometabolic pathways, we performed extra analysis with adjustment for the plasma levels of TMAO, choline, betaine, and L-carnitine. Significant associations for each phenotypic trait are reported at empirical FDR <0.05 level based on 100× permutations.

Microbiome Association to Metabolic Risk of CVD

To estimate individual metabolic risk for CVD development in the population cohort, we used 33 established metabolic biomarkers for CVD measured using the same NMR platform and associated with future CVD incidents in 3 perspective cohorts. We first constructed each individual’s CVD metabolic risk score (MRS) using a weighted risk model:

$$\text{MRS} = \sum_{i=1}^{n} b_i M_i,$$

where $M_i$ is the scaled level of the $i$-th metabolic marker in serum, not adjusted for phenotype, and $b_i$ is a hazard ratio for the corresponding effect of each metabolic marker on the CVD risk as reported by Würzt et al. The MRS showed a normal distribution.

We then tested MRS association to microbial pathways and species, adjusting for age, sex, BMI, and smoking. The significance was controlled at FDR <0.05 based on 100× permutations. For microbiome pathways, we also examined to what extent these pathways are driven by specific taxa by calculating Spearman correlations between pathway abundance and taxa abundance.

Integration Association With Dietary Factors, Inflammatory Markers, and Stool Levels of SCFAs

To better understand the functional properties of the MRS-associated bacterial pathways in relation to metabolism, inflammation, diet, and SCFAs, we conducted an integration analysis with 12 inflammatory markers, 78 dietary factors, and stool levels of 5 SCFAs.

First, we computed pairwise associations between MRS-associated pathways and each of these factors, using linear regression adjusted for age, sex, BMI, and smoking and controlling for FDR of 0.05 using 100× permutations per dataset (cytokines, diet, SCFA) separately. All traits were transformed using inverse rank-sum transformation before analysis.

Second, to elaborate whether pathway-cytokine associations were dependent on MRS, we recalculated these associations using linear regression, additionally adjusting for MRS as a covariate.

Finally, to explore whether the microbial pathways were associated to variance of MRS independently of inflammatory markers and diet, we performed stepwise model selection for each pathway, with MRS as outcome and pathway, age, sex, BMI, smoking, all pathway-associated cytokines, and all pathway-associated diet categories as predictors. This was done using the stepAIC function from R package MASS (v.7.3.50). At each selection step, predictors were selected by both forward and backward direction using Akaike information criterion value as an indicator of goodness-of-fit. The model with the highest Akaike information criterion was selected as the best model. Pathways that survived in the best-fit model were considered independent predictors.

Results

Gut Microbiome Associated With Plasma Metabolomics in Healthy and Obese Individuals

This study included 1275 individuals from 2 independent Dutch cohorts: 978 subjects from the population-based cohort LLD in the Northern part of the Netherlands (the provinces of Groningen, Drenthe, and Friesland), with an average age of 44.5 years (18–81 years), an average BMI of 25 (16–45), and 42% male, and 297 individuals were from the 300-OB cohort on the territory of Nijmegen province, with an average age of 67 (54–81 years), an average BMI of 30.7 (26.3–45.5), and 55% male (Table). Both the LLD and 300-OB cohorts followed cohort-, disease-, and drug-specific exclusion criteria (see Methods). We measured both serum metabolomics and gut metagenomics profiles in both cohorts (Figure 1A). After quality check and imputation of ≥22% missing values, a total of 188 microbial species (Online Table II), 562 bacterial metabolic pathways (Online Table III), and 231 metabolic traits (Online Table IV) were subjected to association analysis. After correcting for age, sex, and BMI, microbial associations were detected for 210 metabolites in LLD at FDR 0.05 level, with 64 associations to 12 species and 4135 associations to 308 pathways. To evaluate whether imputation of missing values in the metabolite data (see Methods) introduced any systematic bias, we re conducted association analyses by removing missing values following 2 approaches and found concordant results (Online Figure II). For instance, after removing missing values, 3952 significant associations were detected at FDR 0.05 level, and 3597 of these overlapped with 4135 associations revealed by missing

Downloaded from http://ahajournals.org by on August 15, 2019
values imputation. In the 300-OB cohort, microbial associations were detected for 42 metabolites, with 1 association to species and 105 associations to 19 pathways. All significant associations at FDR <0.05 level can be downloaded from https://github.com/alexa-kur/NMR_microbiome. Most of the microbial factors identified showed modest effects and could jointly explain, on average, 3.7% of the variation in LLD and 7.7% of the variation in 300-OB (Figure 2). The highest levels of variation explained were 11.1% for glycoprotein N-acetyls in LLD and 16.4% for XS_VLDL_C_percent in 300-OB. The smaller number of associations in 300-OB was likely because of its smaller sample size, but our data also show that the microbial factors identified in LLD generally had lower predictive value for metabolic variation in 300-OB (Figure 2). This indicates there are some genuine differences in microbial associations between the population and the obese cohort. We further compared association strength and directions for the top associations from both cohorts (Online Figures III and IV), including 16 associations to microbial species (Online Table V) and 304 associations to pathways (Online Table VI) at family-wise error rate 0.05 level. We observed some obese-specific associations, particularly for the relative composition of lipoprotein subclasses. For instance, we found significant associations of Ruminococcus species sp_5_1_39BFAA to XXL-VLDL_PL_percent (r=0.30, P=4.6×10^{-7}) in 300-OB that were completely absent in LLD (P>0.05), even with its much larger sample size (Online Table V). We also observed several opposite associations between LLD and 300-OB. For instance, the pathway of pyruvate fermentation to acetate and lactate (PWY-5100) was positively associated to hepatic fat content (liver fat; r=0.21, P<0.001) and maximum steatosis (Ruminococcus sp_5_1_39BFAA, P=1.35×10^{-6}). Interestingly, a large number of associated pathways were related to end point association of hepatic fat content and MRS association of L-methionine biosynthesis and glutamate family AAs (L-proline, L-arginine, L-histidine, and L-histidine), branched-chain AAs (L-valine), hydrophobic AAs (L-threonine), aromatic AAs (L-phenylalanine and L-lysine), and sulfur-containing AAs (L-methionine). Most AA pathways were associated to lower MRS, except for positive associations detected for L-methionine and L-threonine. Other associated pathways were mostly involved in, among others, fermentation, carbohydrates, and sugar derivative metabolism (Online Table X).

We also assessed to what extent specific bacterial taxa can drive MRS-related microbial pathways. For this purpose, we identified the top taxon for each pathway that showed the strongest association between the abundances of the taxon and the pathway (Online Table X). What we found is that the relative contribution of the top taxa varied greatly: the correlation coefficients between top taxa and pathways ranged from 0.26 to 0.89, with an average value of 0.60. This suggests that some pathways are driven by one dominant bacterial player, whereas others may be attributed to many different players. For instance, phylum Bacteroidetes—including class Bacteroidia and families Bacteroidaceae and Rikenellaceae—is the major player in 17 of 31 lower MRS–associated pathways, in particular GDP-mannose biosynthesis and glutamate family AAs (r=0.8). In contrast, the top players in L-methionine metabolism, the Ruminococcus genus and Actinobacteria phylum, only contributed a modest effect (Online Table X).

### Linkage of MRS-Related Pathways to Inflammation and Diet

To gain deeper insight into the contribution of host-microbe-diet interactions to metabolism and inflammation, which both underlie susceptibility to CVD, we conducted a systematic integration analysis between the 48 MRS-associated bacterial pathways and the plasma level of 12 cytokines (as a readout of low-grade inflammation; Online Table XI) and 78 dietary factors (see Methods). After adjustment for age, sex, and BMI, 14 associations between 12 pathways and 5 cytokines were significant at FDR 0.05 level (Figure 4; Online Table XI). The associations detected were also largely independent of MRS and remained significant after adjusting for MRS (Online Table XII). Most associations were found to IL members, namely 8 associations to IL-10, 3 associations to IL-6, 1 to IL-12p70, and 1 to IL-18bp. Elevated levels of these IL members have previously been linked to increased risk of CVD. The pathways associated with these ILs were related to bacterial AA
Figure 2. Variation of plasma metabolites explained by the identified gut microbiome. Bar plot shows the variation of 231 metabolites explained by identified microbial factors in LLD (Lifelines DEEP; left) and in 300-OB (300-Obese; right). Each bar represents 1 metabolite. Bar color indicates type of metabolite according to the color key at right. We also assessed to what extent the microbial factors identified in the LLD cohort could explain the metabolic variation in the 300-OB cohort. These data are shown by the gray bars on the right. Lipoprotein subclasses are classified by their size, density, and content, respectively. Size: L indicates large; M, middle; S, small; XL, very large; XS, very small; density: HDL, high-density lipoprotein; IDL, intermediate-density lipoprotein; (Continued)
biosynthesis (proline, ornithine, threonine, citrulline, tyrosine, and arginine), although IL-10 and IL-18 were also associated to GDP-mannose metabolism glycosylation and homolactic fermentation. Moreover, the bacterial glycolysis III pathway from glucose (GLYCOYSIS) was positively associated to plasma level of adiponectin, which is known to be involved in glucose metabolism regulation. This finding shows a possible interaction between the host and the gut microbiome in glucose metabolism.

Diet is known to be an important factor that affects metabolism, CVD risk, and the gut microbiome. Among 78 dietary factors, 34 were associated to MRS-associated microbial pathways at FDR 0.05 level, after adjustment for age, sex and BMI, and smoking (Online Table XIII). The dietary factors linked with lower CVD-associated pathways included higher intake of fruits, vegetables, nuts, fish, tea and red wine, and a protein-rich or gluten-free diet. Higher intakes of carbohydrates, fat, total calories, sweetened drinks, bread, and dairy products were associated with microbial pathways linked to higher CVD risk (Figure 4).

To further estimate to what extent the microbiome-MRS associations were dependent on effects of diet and inflammation, we included diet and cytokines in the stepwise regression model and selected the best model with the highest Akaike information criterions (see Methods). This analysis showed that 43 of 48 pathways survived feature selection and were included as predictors, which indicates that they were significantly associated to MRS variation independent of diet and inflammation (Online Table XIV).

**Plasma Levels Rather Than Stool Levels of SCFAs Are More Relevant to CVD**

Because fruit and vegetable intake was mostly associated to MRS-associated pathways, we hypothesized that bacterial production of SCFAs by fiber fermentation may underlie the mechanism. We measured stool levels of 5 different SCFAs (Online Table XI) and found they were associated to most of the MRS-associated microbial pathways (Figure 4). To our surprise, the association directions we observed were in contrast to the previously suggested beneficial effect of SCFAs. Furthermore, no significant associations were found between stool levels of SCFAs and MRS (Online Table XV). Because 95% of the SCFAs produced in the gut are rapidly absorbed by colonocytes, and only 5% are actually secreted into the feces, we hypothesized that blood levels of SCFAs might be more relevant to CVD risk. We, therefore, extracted the plasma level of acetate—the most abundant SCFA—from the NMR-based metabolic profiling. Indeed, the plasma level of acetate was associated to lower MRS ($P = −0.09$; $P = 7.2 \times 10^{-4}$) and to 29 MRS-associated pathways at $P < 0.05$, all with the expected effect directions (Online Table XVI).

**Discussion**

This study presents a comprehensive exploration of the relationship of the gut microbiome with plasma metabolites, metabolic risk of CVD, and cardiometabolic phenotypes in 1275 individuals from a population cohort and a cohort of overweight individuals with cardiovascular and metabolic complications. We conducted an integrated analysis of diet, metagenome, plasma metabolites, inflammatory cytokines, and stool SCFAs. Previously, relationships of the gut microbiome to plasma metabolites and metabolic traits were investigated using 16S rRNA sequencing. To our knowledge, this study is the largest metagenome-based study to date and provides richer information into the functional link between the gut microbiome and CVD risk. It is also the first study to simultaneously address microbial association with cardiometabolic phenotypes in obesity and with metabolic risk for future development of CVD in a population-based nonpatient cohort. Because of the scope of the NMR-based metabolic platform, this study primarily focused on lipid-related traits, including various lipoprotein subclasses and fatty acids. Microbial associations to a broader spectrum of metabolism, therefore, still need to be identified.
Although our group had previously established the association of the gut microbiome with routine blood lipid level, our current data show that it is not only levels of lipoproteins but also their size and composition that is important and differentially associated with microbiome taxonomic and functional composition. After correcting for age, sex, BMI, and smoking, microbial associations were identified for 210 metabolic traits in the LLD population cohort and for 35 metabolic traits in the 300-OB cohort. The microbial factors identified showed modest effects, jointly explaining ≤11.1% of the variation in plasma level of glycoprotein acetylts in LLD and ≤16.4% of variation in XS_VLDL_C_percent in 300-OB. Our data also show obese-specific microbial associations, in particular for lipid compositions in VLDL, IDL, and LDL lipoprotein subclasses. Furthermore, our data identified that bacterial L-methionine biosynthesis and the species Ruminococcus_5_1_39BFAA were associated to atherosclerosis and liver fat content, respectively, in our obese cohort and that 48 bacterial pathways were associated to MRS of CVD in our population cohort. Finally, we integrated the microbiome and metabolomics with diet, SCFAs, and low-grade inflammation, with most associations being detected to fruit
and vegetable intake, plasma levels of adiponectin, and several IL family members (IL-10, IL-6, IL-2P70, and IL-18P). We also observed that the microbiome-MRS associations identified were largely independent of diet and cytokines. However, the current dietary information was obtained from food questionnaires, and the inaccuracy of self-reported data can attenuate the power in dietary analysis. Our data further show that plasma level of SCFAs is more relevant to CVD risk than stool SCFA levels. We did not confirm the previously observed association of TMAO to CVD in our 300-OB cohort, and similar negative associations were reported in a study of young adults (CARDIA cohort). However, another study in a German cohort did report associations between carotid IMT and TMAO but also observed substantial loss in association signal after adjusting for age. Moreover, our study revealed TMAO association to visceral fat, which supports the previous finding. When comparing microbial associations in different human cohorts, it is important to keep in mind that differences in diet, genetic background, and environmental exposures can affect the gut microbiome, TMAO production, and CVD risk, thereby resulting in conflicting findings. For instance, the 300-OB cohort contained only elderly, obese Dutch individuals whose plasma levels of TMAO, at an average of 5.18 μmol/L, were higher than those reported in young or population-based cohorts.

Among the bacterial associations we identified, the bacterial pathways of L-methionine biosynthesis showed consistent links with plasma metabolites, MRS of CVD, and atherosclerotic plaques, and these pathways were driven by lower fruit intake. These observations are consistent with some previous findings. For instance, supplementation of the glutamine family AAs is predicted to have a beneficial effect toward decreasing CVD risk, while L-methionine and its metabolic products S-adenosyl-L-methionine and L-homocysteine have been associated with CVD incidence and complication. Methionine has an essential role in a number of cellular processes, including the initiation of protein synthesis, the methylation of DNA, and metabolism of xenobiotics. It is also a crucial factor in the biosynthesis of cysteine, phospholipids, and polyamine. It is hypothesized that L-methionine induces atherosclerosis by increasing plasma homocysteine levels, as L-methionine can be converted to homocysteine directly or through S-adenosyl-L-methionine. Hyperhomocysteinemia has also been related to CVD development. A recent meta-analysis addressing the effects of low homocysteine by folate acid supplementation found a 10% reduction in risk for stroke and a 4% reduction in risk for CVD. Individuals with homocysteinuria—a genetic disorder characterized by severe hyperhomocysteinemia—experience severe atherosclerotic disease at a young age. A number of mechanisms have been proposed to explain the induction of atherosclerosis by elevated homocysteine levels, for example, through endothelial dysfunction, an increase in proliferation of vascular smooth muscle cells, oxidative damage with deterioration of arterial wall elastic material, and a reduction of HDL cholesterol levels. Furthermore, homocysteinemia has been shown to promote the attraction of monocytes and production of proinflammatory cells. Homocysteine also induces macrophage maturation in vessel walls with enhanced vascular inflammation. Recently, Wang et al revealed a proinflammatory status via NLRP3 (Nod-like receptor family pyrin domain-containing 3) inflammasome activation in hyperhomocysteinemia induced by a high methionine diet in apoE-deficient mice. Our data thus highlight that bacterial metabolism of L-methionine is also associated to the development of CVD in humans—an observation that was confirmed in both our population-representative cohort and the CVD-enriched cohort of overweight individuals. The association between bacterial metabolism of L-methionine and atherosclerotic plaques was observed to be independent of TMAO metabolism (L-carnitine, choline, betaine, and TMAO) and BMI.

Our study also identified several functional links between the gut microbiome and metabolic profile that may predict future CVD events, in particular, the association of bacterial pathways related to metabolism of AAs, carbohydrates, and polysaccharides (specifically, GDP-mannose) with MRS. This highlights the potential of microbiome-targeting therapy for CVD prevention and treatment. Some of these pathways are dominantly driven by a specific taxon such as class Bacteroidia. However, there are several pathways, including bacterial pathways of L-methionine metabolism, which are driven by many different taxa, each with modest or small effect. This suggests potential applications of different microbiome-targeting approaches in controlling a certain taxon or bacterial pathway, for example, through personalized dietary control combined with prebiotic and probiotic treatment. Interestingly, the top players in MRS-associated bacterial metabolism we identified were in line with previous findings. For instance, genus Faecalibacterium, Subdoligranulum, species Methanobrevibacter smithii, Eubacterium eligens, and others were top players in lower MRS-associated pathways. Previous studies have suggested their health-promoting properties: Faecalibacterium members have been associated to lower intestinal and adipose tissue inflammation, lower levels of members of Rikenellaceae family have been associated to liver disease and obesity, and levels of different members of Bacteroidaceae family have been associated with numerous host properties, acting as both mutualistic and pathogenic cohabitants.

We acknowledge several limitations in our study. First, both the LLD and 300-OB cohorts comprised of participants of Dutch ethnicity. The reported results are thus likely biased toward a region-specific genetic background and diet, and both are known to affect both metabolism and the gut microbiome. Second, this was an association analysis based on a cross-sectional design, which means that the underlying causality and mechanism remain unclear. The Mendelian randomization approach is considered to be a powerful method to assess causality. However, several recent studies have shown that genetics and microbiome likely exert independent additive effects on CVD-related phenotypes that include, among others, blood lipid levels, CVD-related proteins, and BMI. This may limit the power of our Mendelian randomization analysis to illustrate underlying causality. Longitudinal studies and further
functional studies are thus essential to reveal the underlying mechanism and causality.

Conclusions

Our study presents an integrated analysis of the relationship between the gut microbiome and host metabolomics, cardiovascular metabolic phenotypes, and metabolic risk of CVD in humans. Importantly, we investigated microbial association in both a population-based and an obese cohort.

We identified numerous associations of functional properties and microbial species in the gut microbiome with plasma metabolic traits, including lipoprotein particle composition, fatty acid saturation, and glycoprotein N-acetyl. Some of the microbial factors identified were also linked with clinical outcomes in obese subjects, including hepatic fat content and atherosclerosis. In our population-representative cohort, the combined metabolic risk that represents the probability of having a CVD event in future is associated with numerous microbiome functional parameters such as biosynthesis and degradation of AAs, fermentation, and carbohydrate and sugar derivative metabolism.

Altogether, our study highlights microbial associations to current and future clinical outcomes related to CVD. The microbial factors identified and their interactions with diet and inflammation, such as association of bacterial L-methionine biosynthesis with CVD risk and current atherosclerosis, provide a rationale for the future studies, including intervention and prospective experiments. These may contribute to the development of preventive or therapeutic strategies aimed at modulating the microbiome to reduce the burden of cardiovascular events.

Acknowledgments

We thank the participants and staff of LifeLines DEEP for their collaboration. The study was approved by the University Medical Center Groningen review board, ref. M12.113965. We thank Jackie Dekens, Mathieu Platteel, Maria Carmen Centi, Astrid Maatman, and Jody Arens for management and technical support and Kate Mc Intrye for editing the manuscript. We gratefully acknowledge Vincenzo Postino of the National Research Council Institute of Clinical Physiology, Pisa, Italy, for providing the HIPPO FAT software. We thank Floris Imhann, Arnau Vich Vila, and Rinse K. Weersma for discussions. C. Wijmenga, M.G. Netea, A. Zhernakova, and J. Fu codirected the study and share senior authorship. Additional information: we dedicate this article to the memory of Marten Hofker (1956–2016).

Sources of Funding

This project was funded by Dutch Heart Foundation (CVON2012-03, IN-CONTROL) to M.H. Hofker, M.G. Netea, D.P.Y. Koonen, F. Kuipers, J.H.W. Rutten, N.P. Riksien, A. Zhernakova, L.A.B. Joosten and J. Fu; from the Top Institute Food and Nutrition, Wageningen, the Netherlands, to C. Wijmenga (TiFN GH001); and by the Netherlands Organization for Scientific Research (NWO) to J. Fu (NWO-VIDI 864.13.013), A. Zhernakova (NWO-VIDI 016.Vidi.178.056), and M. Oosting (NWO-VENI 016.176.006). A. Zhernakova also holds a European Research Council (ERC) Starting Grant (ERC No. 715772) and a Rosalind Franklin Fellowship from the University of Groningen. This research also received funding under the European Union Seventh Framework Program. C. Wijmenga is supported by an FP7/2007–2013/ERC Advanced Grant (agreement 2012–322698) and by a Spinoza Prize from NWO (SPI 94–226). M.G. Netea holds an ERC Consolidator Grant (ERC No. 310372) and a Spinoza Prize from NWO (SPI 94–212). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosures

None.

References


