The Gut Microbiome Contributes to a Substantial Proportion of the Variation in Blood Lipids

Citation

Published Version
doi:10.1161/CIRCRESAHA.115.306807

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:23473982

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current/terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
The Gut Microbiome Contributes to a Substantial Proportion of the Variation in Blood Lipids

Jingyuan Fu, Marc Jan Bonder, María Carmen Cenit, Ettje F. Tigchelaar, Astrid Maatman, Jackie A.M. Dekens, Eelke Brandsma, Joanna Marczynska, Floris Imhann, Rinse K. Weersma, Lude Franke, Tiffany W. Poon, Ramnik J. Xavier, Dirk Gevers, Marten H. Hofker*, Cisca Wijmenga®, Alexandra Zhernakova*

Rationale: Evidence suggests that the gut microbiome is involved in the development of cardiovascular disease, with the host–microbe interaction regulating immune and metabolic pathways. However, there was no firm evidence for associations between microbiota and metabolic risk factors for cardiovascular disease from large-scale studies in humans. In particular, there was no strong evidence for association between cardiovascular disease and aberrant blood lipid levels.

Objectives: To identify intestinal bacteria taxa, whose proportions correlate with body mass index and lipid levels, and to determine whether lipid variance can be explained by microbiota relative to age, sex, and host genetics.

Methods and Results: We studied 893 subjects from the LifeLines-DEEP population cohort. After correcting for age and sex, we identified 34 bacterial taxa associated with body mass index and blood lipids; most are novel associations. Cross-validation analysis revealed that microbiota explain 4.5% of the variance in body mass index, 6% in triglycerides, and 4% in high-density lipoproteins, independent of age, sex, and genetic risk factors. A novel risk model, including the gut microbiome explained ≤25.9% of high-density lipoprotein variance, significantly outperforming the risk model without microbiome. Strikingly, the microbiome had little effect on low-density lipoproteins or total cholesterol.

Conclusions: Our studies suggest that the gut microbiome may play an important role in the variation in body mass index and blood lipid levels, independent of age, sex, and host genetics. Our findings support the potential of therapies altering the gut microbiome to control body mass, triglycerides, and high-density lipoproteins.

Key Words: body mass index ■ cardiovascular diseases ■ lipids ■ lipoproteins, HDL ■ metabolism

In recent years, the gut microbiome has emerged as an important player in human health.1,2 Gut microbiota comprise thousands of microbial species that are involved in host metabolism by regulating energy extraction, activation of the immune system, drug metabolism, and other processes.3,4 Association of bacterial composition with many diseases has been observed, including immune, inflammatory, and metabolic phenotypes.5–7 Several mechanisms for the downstream effect of microbiota were discovered which also suggest that they play a role in cardiovascular disease (CVD). The microbiota play an important role in choline diet–induced trimethylamine N-oxide production, which has been implicated in CVD.8 A further mouse study has demonstrated that atherosclerosis susceptibility can be transmitted via gut microbiota transplantation.9 Furthermore, dysbiosis in the gut has been shown to induce increased permeability of the intestine, leading to increased systemic levels of bacterial products causing low-grade chronic inflammation.10 This inflammation may directly affect atherogenesis and has also been hypothesized to lead to the development of insulin resistance.

In This Issue, see p 743
Editorial, see p 750

© 2015 The Authors. Circulation Research is published on behalf of the American Heart Association, Inc., by Wolters Kluwer. This is an open access article published under the terms of the Creative Commons Attribution Non-Commercial-NoDerivs License, which permits use, distribution, and reproduction in any medium, provided that the original work is properly cited; the use is noncommercial, and no modifications or adaptations are made.
with concomitant effects on plasma lipids.15 Gut microbiota have also been linked with lipid metabolism through their role in bile acid metabolism. They can also influence the efficiency of energy harvest from ingested food2,15 and play a crucial role in the metabolic processes and development of obesity.

In line with these observations, altering the gut microbiome in humans and mice has shown improvement in metabolic syndrome.13–15 However, the evidence for a causal relationship between the gut microbiome and the development of CVD has not been firmly established for lack of large-scale human studies. Atherosclerosis, a lipid-driven disease, is the main underlying cause of CVD. However, to date, no studies of sufficient size have been done to assess the association between lipids and microbiota. In this study, we performed a systematic analysis of host genome, gut microbiome, body mass index (BMI), and blood lipids in 893 human subjects from the Dutch LifeLines-DEEP cohort.16 We investigated that which gut bacteria were associated with BMI and blood lipids, and how much of the variation in blood lipids could be explained by the gut microbiome, relative to age, sex, BMI, and host genetics.

Methods

Population Cohort

The LifeLines-DEEP cohort is a subcohort of the LifeLines cohort (167 729 subjects),17 which uses a broad range of investigative procedures in assessing the biomedical, sociodemographic, behavioral, physical, and psychological factors that contribute to the health and disease of the general population. A subset of approximately 1500 participants also took part in LifeLines-DEEP: for these participants, additional biological materials were collected, including genome-wide genotyping and analysis of the gut microbiome composition. A full description of the LifeLines-DEEP data set is given in the article describing the study design.16

Lipid Measurements

We had lipid measurements available for all 1500 LifeLines-DEEP samples. Total cholesterol (TC) was measured with an enzymatic colorimetric method, high-density lipoprotein (HDL) cholesterol with a colorimetric UV method (Modular P analyzer, Roche Diagnostics, Burgdorf, Switzerland). The low-density lipoprotein (LDL) cholesterol concentration was calculated using the Friedewald equation. More details were reported previously.15 The triglyceride level was further log2 transformed.

Genotype Information

All LifeLines-DEEP samples were genotyped using the HumanCytoSNP-12 BeadChip and ImmunoChip, a customized Illumina Infinium array. The data were harmonized,19 merged, and subsequently imputed using the Genome of the Netherlands (GoNL) data set.20,21 Further details and information on the quality control are described by Tiggelaar et al.15 We removed ethnic outliers and genetically related participants from our study.

Microbiome Data Generation

Sequencing

Microbiome data were generated for 1180 LifeLines-DEEP samples. Fecal samples were collected at home within 2 weeks after collection of blood samples and stored immediately at −20°C. After transport on dry ice, all samples were stored at −80°C. Aliquots were made, and DNA was isolated with the AllPrep DNA/RNA Mini Kit (Qiagen; cat. #80204). Isolated DNA was sequenced at the Broad Institute, Boston, using Illumina MiSeq paired-ends. Hypervariable region V4 was selected using forward primer 515F [GTGCACGCMGCCGCGGTAA] and reverse primer 806R [GGACTACHVGGGTWTCTAAT]. We used custom scripts to remove the primer sequences and align the paired-end reads. Details are given in the study by Gevers et al.22

OTU Picking

Selection of unique bacterial sequences, so-called operational taxonomic unit (OTU) picking, was performed using the QIIME (the toolbox for Quantitative Insights Into Microbial Ecology) reference optimal picking, which uses UCLUST (an algorithm to cluster sequence reads based on similarity)23 (version 1.2.2q) to perform the clustering. Matching OTUs to bacteria was done using a primer-specific version of the GreenGenes 13.5 reference database.24 Using TaxMan,25 we created the primer-specific reference database containing only reference entries that matched the selected primers. During this process, we restricted probe-reference mismatches to a maximum of 25%. The 16S regions that were captured by our primers, including the primer sequences, were extracted from the full 16S sequences. For each of the reference sequences, we determined the overlapping part of the taxonomy of each of the reference reads in the clusters and used this overlap as the taxonomic label for the cluster. This process is based on, and similar to, the work described by Bonder et al.,25 Brandt et al.,26 May et al.,27 and Ding et al.28 We used QIIME29 for exploratory analysis and for gathering basic statistics on the microbiome data set.

Quality Control

Overall, for 1021 samples, we had lipid measurements, genotype, and microbiome information. We excluded 99 samples from participants who were taking antibiotic or other potential microbiome-modifying drugs or who were on lipid-lowering medication. The library size of microbial sequencing varied greatly among samples, ranging from 3969 to 336 900 reads. The sequence depth can significantly bias the measures of microbial composition, and rarefaction was widely used to make the library sizes equal by randomly selecting the same number of reads per sample.30,31 We compared the number of samples at different sequence depths and determined the rarefaction depth based on criteria to obtain both the number of reads and the number of samples as high as possible. We rarefied the library size to 15 000 read-depth using the rarefy function in R package vegan (version 2.3-0). At this depth, we only excluded 29 subjects. After these exclusion steps, we had 893 samples (380 men and 513 women) for final analysis. Their characteristics are summarized in Table.

Furthermore, we filtered on the OTU abundance and confined our analysis to 645 OTUs, each of which comprised ≥0.05% of reads and was present in at least 1% of the population. These OTUs accounted for an average of 99% of total reads per sample. The OTUs were assigned to 173 taxonomies that were further truncated to 136 taxonomies after removing identical or highly similar information between different clade levels.

Statistical Analysis

Analysis of Microbial Diversity

The microbial Shannon diversity index was calculated using the diversity function in R package vegan (version 2.3-0).

Two-Part Model for Association Analysis

We observed that the distribution of the abundance of OTUs or taxonomies departed significantly from a normal distribution because of the fact that bacteria were not presented in many samples. Only 50 of 645 OTUs (7.7%) were presented in >90% of samples, whereas
### Table. Summary of Physical Characteristics of 893 LifeLines-DEEP Subjects

<table>
<thead>
<tr>
<th>Age, y</th>
<th>BMI</th>
<th>HDL-C, mmol/L</th>
<th>LDL-C, mmol/L</th>
<th>TC, mmol/L</th>
<th>TG, mmol/L</th>
<th>Log2-TG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>Women</td>
<td>Total</td>
<td>Men</td>
<td>Women</td>
<td>Total</td>
<td>Men</td>
</tr>
<tr>
<td>44.7±12.9</td>
<td>44.6±12.9</td>
<td>44.6±12.9</td>
<td>18±78</td>
<td>18±80</td>
<td>18±80</td>
<td>1.39±1.23</td>
</tr>
<tr>
<td>25.4±3.3</td>
<td>25.1±4.6</td>
<td>25.2±4.1</td>
<td>16.9±39.3</td>
<td>16.9±44.9</td>
<td>16.9±44.9</td>
<td>3.36±0.89</td>
</tr>
<tr>
<td>15.3±1.00</td>
<td>5.04±1.01</td>
<td>5.08±1.01</td>
<td>2.5±8.7</td>
<td>2.4±9.7</td>
<td>2.4±9.7</td>
<td>1.39±1.23</td>
</tr>
<tr>
<td>0.20±0.81</td>
<td>−0.21±0.67</td>
<td>−0.04±0.76</td>
<td>−2.2±3.8</td>
<td>−1.79±2.02</td>
<td>−2.2±3.81</td>
<td>−2.2±3.8</td>
</tr>
</tbody>
</table>

**BMI** indicates body mass index; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; and TG, triglycerides.

448 OTUs (69.5%) were detected in <50% of samples. At the taxonomic level, 32 of 136 (9.5%) taxonomies were detected in >90% of samples, whereas 60 taxonomies (44.1%) were detected in <50% of samples. There are different explanations for the detection rate: (1) the bacteria are really absent in the samples; (2) the abundance levels of bacteria are lower and not to be detected at the current sequencing and rarefication depth; (3) the abundance levels are similar and it is a random effect because of the sequencing or rarefication procedure. We therefore adopted a novel, 2-part model that was developed to account for both binary (detected/undetected) and quantitative features. This approach overcomes the problem of a non-normal distribution, which is a feature of the majority of gut bacteria OTUs or taxa.

The 2-part model is illustrated in Online Figure I. The first part describes a binomial analysis that tests for association of detecting a microbe (represented by an OTU or taxonomy) with a trait. The binary feature (b) describes a binomial analysis that tests for association of detecting a microbe is a feature of the majority of gut bacteria OTUs or taxa. The abundance level (q) of a microbe was the log10 transformed read count per individual. The abundance level (q) is a binary feature, β is the estimated effect for the binary effect, and e represents the residuals.

The second part of the quantitative analysis tests for association between the lipid level and the abundance of bacteria, but only for the subjects where that microbe is present. The abundance level (g) of a microbe was the log$_{10}$ transformed read count per individual. The quantitative model is written as: $y = \beta g + e$, where $g$ refers to the trait level (BMI or lipid level) per individual after adjusting for age and sex, $b$ is a binary feature, $\beta$ is the estimated effect for the binary effect, and $e$ represents the residuals.

To further combine the effect of both binary and quantitative analysis, a meta $P$ value was derived using an unweighted $Z$ method. Then, a final association $P$ value per microbe-trait pair was assigned from the minimum of $P$ values from the binary analysis, quantitative analysis, and meta-analysis. The association $Z$ score was calculated based on the $Z$ distribution. If the association direction was negative, the $Z$ score was assigned a positive value. If the association direction is positive, the $Z$ score was assigned a positive value.

The association $P$ value was set as the minimum value of 3 $P$ values and the distribution of the association $P$ values could be skewed, so we therefore performed 1000x permutation tests to control the false discovery rate (FDR). For each permutation, we randomized the gut microbial composition across individuals and performed the 2-part analysis on permuted data. At a certain $P$ cutoff, the average number of the detected significance ($N_j$) in 1000x permutations was defined as the false positive, and its ratio to the detected positive ($N_j$) in the real analysis was the FDR. We controlled the FDR at 0.05.

This method accounts for the complicated features of the microbial data and maximizes the power. If the association $P$ value comes from the binary model, indicating the effect is only because of the presence/absence of the microbe, the abundance of the microbe in the samples does not matter. If the association $P$ value comes from the quantitative model, this indicates the abundance level of the microbe associates with the trait, and the absence of the microbe has no influence. The explanation would be another microbe takes its place for both binary (detected/undetected) and quantitative features. This approach overcomes the problem of a non-normal distribution, which is a feature of the majority of gut bacteria OTUs or taxa.

### Estimating the Variance Explained by the Gut Microbiome

To estimate the proportion of variation in BMI and lipids that could be explained by the gut microbiome, we performed a 1000x cross validation. Each time we split the data randomly into an 80% discovery set and a 20% validation set. In the discovery set, a total of $n$ number of significantly associated OTUs was identified at a certain $P$ value, and the effect sizes of binary and quantitative features of each OTU ($\beta_1$ and $\beta_2$) were estimated. Then, the risk of the gut microbiome on BMI or lipids ($r_{g}$) for each individual in the validation set was calculated using an additive model:

$$r_g = \sum (\beta_1 + b + \beta_2 g).$$

The variation in BMI and blood lipids explained by the gut microbiome was represented as the squared correlation coefficient ($R^2$) between the traits and $r_g$ after correcting for age and sex. To ensure the robustness of our estimation, we repeated the cross-validation 100x and calculated the average value of the explained variation. We hypothesized that many microbes may contribute a small effect but may not be confidently detected at an FDR of 0.05. Therefore, we did this analysis at different significant $P$ levels ranging from 1×10⁻⁵ to 0.1.

### Genetic Risk Score Calculation

A total of 157 lipid-associated single nucleotide polymorphisms (SNPs) and 97 BMI-associated SNPs were extracted from the literature. To ensure completeness and effect size estimates were made for each SNP and each lipid type. We excluded 3 SNPs for which genotypes could not be successfully imputed in the LifeLines cohort: rs9411489 at the ABO locus, rs3177928 at the HLA locus, and rs12016871 at the MTTF3 locus. Thus, our final study included genetic information for 96 BMI-associated SNPs and 155 lipid-associated SNPs, including 71 for HDL, 56 for LDL, 40 for triglycerides, and 72 for TC. We then computed weighted genetic risk scores ($r_g$) for BMI and lipids, as described previously.

The association analysis between individual SNPs and the gut microbiome was performed using the analysis pipeline developed in the LifeLines cohort: rs9411489 at the ABO locus, rs3177928 at the HLA locus, and rs12016871 at the MTTF3 locus. Thus, our final study included genetic information for 96 BMI-associated SNPs and 155 lipid-associated SNPs, including 71 for HDL, 56 for LDL, 40 for triglycerides, and 72 for TC. We then computed weighted genetic risk scores ($r_g$) for BMI and lipids, as described previously.

The association analysis between individual SNPs and the gut microbiome was performed using the analysis pipeline developed in the LifeLines cohort: rs9411489 at the ABO locus, rs3177928 at the HLA locus, and rs12016871 at the MTTF3 locus. Thus, our final study included genetic information for 96 BMI-associated SNPs and 155 lipid-associated SNPs, including 71 for HDL, 56 for LDL, 40 for triglycerides, and 72 for TC. We then computed weighted genetic risk scores ($r_g$) for BMI and lipids, as described previously.

The association analysis between individual SNPs and the gut microbiome was performed using the analysis pipeline developed in the LifeLines cohort: rs9411489 at the ABO locus, rs3177928 at the HLA locus, and rs12016871 at the MTTF3 locus. Thus, our final study included genetic information for 96 BMI-associated SNPs and 155 lipid-associated SNPs, including 71 for HDL, 56 for LDL, 40 for triglycerides, and 72 for TC. We then computed weighted genetic risk scores ($r_g$) for BMI and lipids, as described previously.

The association analysis between individual SNPs and the gut microbiome was performed using the analysis pipeline developed in the LifeLines cohort: rs9411489 at the ABO locus, rs3177928 at the HLA locus, and rs12016871 at the MTTF3 locus. Thus, our final study included genetic information for 96 BMI-associated SNPs and 155 lipid-associated SNPs, including 71 for HDL, 56 for LDL, 40 for triglycerides, and 72 for TC. We then computed weighted genetic risk scores ($r_g$) for BMI and lipids, as described previously.

The association analysis between individual SNPs and the gut microbiome was performed using the analysis pipeline developed in the LifeLines cohort: rs9411489 at the ABO locus, rs3177928 at the HLA locus, and rs12016871 at the MTTF3 locus. Thus, our final study included genetic information for 96 BMI-associated SNPs and 155 lipid-associated SNPs, including 71 for HDL, 56 for LDL, 40 for triglycerides, and 72 for TC. We then computed weighted genetic risk scores ($r_g$) for BMI and lipids, as described previously.

The association analysis between individual SNPs and the gut microbiome was performed using the analysis pipeline developed in the LifeLines cohort: rs9411489 at the ABO locus, rs3177928 at the HLA locus, and rs12016871 at the MTTF3 locus. Thus, our final study included genetic information for 96 BMI-associated SNPs and 155 lipid-associated SNPs, including 71 for HDL, 56 for LDL, 40 for triglycerides, and 72 for TC. We then computed weighted genetic risk scores ($r_g$) for BMI and lipids, as described previously.
where \( r_g \) is the calculated genetic risk and \( r_m \) is the highest microbial risk we determined. The variation explained by each risk model was calculated in 100× cross validation, as described above. To evaluate the significance of microbial contribution, the ANOVA test was used to compare the performance of the risk models \( r_g \) and \( r_m \): the average of \( F \) values of the ANOVA test from 100× cross validation was calculated, and the \( P \) value was determined based on the \( F \)-distribution. As BMI and lipids are highly correlated, we also investigated whether the gut microbiome can contribute to lipid levels independent of BMI. To do so, we tested 4 risk models of lipids, including BMI as a risk factor:

\[
\begin{align*}
    r_1 &= \text{age} + \text{gender}; \\
    r_2 &= \text{age} + \text{gender} + \text{BMI}; \\
    r_3 &= \text{age} + \text{gender} + \text{BMI} + r_g; \quad \text{and} \\
    r_4 &= \text{age} + \text{gender} + \text{BMI} + r_g + r_m.
\end{align*}
\]

**Results**

**Microbial Diversity in the LifeLines-DEEP Cohort**

After quality control, our study included 893 human subjects. The study cohort had a wide range of age, BMI, and blood lipids levels (Table). We assessed how variable the gut microbial composition was in the cohort in terms of microbial richness and diversity. The microbial richness reflects the number of OTUs per individual. The cohort had on average 238 OTUs per individual, ranging from 44 to 355. When individuals were grouped into different bins based on their richness, we observed that age and the proportion of women were higher in the richer OTU groups (Figure 1). The Spearman correlation showed that the richness was significantly higher in women (\( P=0.0055 \)) and increased with age (\( P=5.87 \times 10^{-12} \); Online Table I). Given the abundance of OTUs, we computed the microbial diversity (Shannon’s diversity index) and observed similar significant correlations for age and sex (Online Table I). We then investigated whether bacterial richness and diversity were correlated with BMI and lipid levels. After correcting for age and sex, OTU richness was negatively correlated with BMI (\( P=3.8 \times 10^{-10} \)) and triglycerides (\( P=1.37 \times 10^{-4} \)), but positively correlated with HDL (\( P=8.3 \times 10^{-5} \)). We did not observe significant correlations between microbial richness and LDL or TC levels (Online Table I).

**Association of Bacteria With Lipid Metabolites**

We next tested for association between the individual bacterial OTU, BMI, and blood lipid levels. After adjusting for age and sex, we identified 148 associated OTUs at FDR=0.05: 66 OTUs were associated with BMI, 114 with triglycerides, and 34 with HDL (Online Tables II–IV). We did not detect any significant association at OTU level for LDL or TC. Of the 148 associated OTUs, 12 were shared by all 3 traits (BMI, triglyceride, and HDL), 29 OTUs were shared by BMI and triglycerides, and 4 by BMI and HDL, whereas 21, 64, and 9 OTUs were specifically associated with BMI, triglyceride, and HDL, respectively (Online Figure II). At the taxonomic level, we identified 50 significant associations for 34 unique taxonomies at FDR=0.05: 22 were associated with BMI, 23 with triglycerides, 4 with HDL, and 1 with LDL (Figure 2; Online Table V). We found that 18 associations (36%) were detected by binary analysis (presence/absence); 4 associations (8%) were detected by the quantitative model; and 28 associations (56%) were detected by the meta-analysis of binary and quantitative analyses (Online Table V). Although most of the associated taxonomies were shared across lipid metabolites and BMI, several microbes were predominantly linked to lipids rather than BMI. For example, the family *Clostridiaceae/Lachnospiraceae* (\( N^{32} \)) was negatively associated with BMI (\( P=9.8 \times 10^{-7} \) and \( P=0.0002 \), respectively), a nominal significance to other lipids, and no association with BMI (\( P=0.1 \)).

We confirmed several previously described bacterial associations with obesity. An increased abundance of genus *Akermannia* (\( N^{34} \)) has been associated with a decrease in BMI (\( P=0.0005 \)).\(^{32}\) We also confirmed the association of both the family *Christensenellaceae* (phylum *Firmicutes*; \( N^{37} \)) and the phylum *Tenericutes* (mainly represented by order *RF-39*; \( N^{13} \)) with low BMI (\( P=9.8 \times 10^{-7} \) and \( P=0.0002 \), respectively), as reported in the TwinsUK cohort.\(^{34}\) In addition, we identified a novel and strong association of these particular bacteria with lower levels of triglyceride (\( P=2.1 \times 10^{-5} \) and \( P=2.7 \times 10^{-7} \), respectively) and higher levels of HDL (\( P=0.0047 \) and \( P=0.0006 \), respectively). We also observed several new associations with BMI and levels of triglyceride and HDL, such as genus *Eggerthella* (\( N^{8} \)) with increased triglyceride (\( P=4.1 \times 10^{-5} \)) and decreased HDL (\( P=6.3 \times 10^{-5} \)), and family *Pasteurellaceae* (\( N^{3} \)) with decreased triglyceride (\( P=6.2 \times 10^{-5} \)). The genus *Butyrivibrios* (\( N^{8} \)) was previously linked to a lean phenotype in mice after fecal transplantation from twins discordant for obesity.\(^{14}\) Our study shows that this genus is strongly associated with decreased triglyceride (\( P=4.7 \times 10^{-5} \)) and nominally associated with BMI and HDL in humans.

![Figure 1. The richness of the gut microbiome.](image)
To anticipate how much BMI and blood lipids can be modulated by the gut microbiome, it is important to estimate what proportion of variation in these metabolic traits can be explained by the microbiome. To do so, we performed a 100× cross-validation analysis by splitting the data set randomly into an 80% discovery set and a 20% validation set. The OTUs identified at $P=1\times 10^{-5}$ level in the discovery set explained 2.74% variation in BMI in the validation set, 3.83% in triglycerides, 2.46% in HDL, 0.01% variation in LDL, and 0.01% in TC. When the association significance decreased and the risk model included more (but less-significant) OTUs, the explained variation increased to 4.57% in BMI, 6.0% in triglycerides, 4.0% in HDL, but was only 1.5% in LDL and 0.7% in TC (Figure 3A). To test the robustness of our estimation, we rerarefied the OTU library 100× and repeated the

**Variance of Blood Lipid Explained by Microbiota Composition**

To anticipate how much BMI and blood lipids can be modulated by the gut microbiome, it is important to estimate what proportion of variation in these metabolic traits can be explained by the microbiome. To do so, we performed a 100× cross-validation analysis by splitting the data set randomly into an 80% discovery set and a 20% validation set. The OTUs identified at $P=1\times 10^{-5}$ level in the discovery set explained 2.74% variation in BMI in the validation set, 3.83% in triglycerides, 2.46% in HDL, 0.01% variation in LDL, and 0.01% in TC. When the association significance decreased and the risk model included more (but less-significant) OTUs, the explained variation increased to 4.57% in BMI, 6.0% in triglycerides, 4.0% in HDL, but was only 1.5% in LDL and 0.7% in TC (Figure 3A). To test the robustness of our estimation, we rerarefied the OTU library 100× and repeated the
Microbiota Contribute Significantly to Lipid Variation, Independently of Age, Sex, and Genetics

Evidence has already shown that the gut microbiome can be shaped by host genetics.38 We further tested whether the explained variation in the gut microbiome was independent of genetic factors by testing the association between the gut microbiome and genetic risk scores.46 To date, 157 genetic loci have been reported to be associated with blood lipid levels,31 and 97 loci have been associated with BMI.34 In our cohort, these SNPs collectively explained 2.1% variation in BMI (P = 1.66×10^{−5}), 3.4% in triglycerides (P = 3.22×10^{−8}), 7.5% in HDL (P < 2.2×10^{−16}), 4.6% in LDL (P = 8.0×10^{−11}), and 5.6% in TC (P = 7.7×10^{−11}), after correcting for age and sex. However, we did not observe any significant association between the microbiome and the genetic risk at FDR = 0.05. We did not find a significant association for either single SNPs (Online Table VI) or for the combined lipid and BMI genetic risk scores (Online Table VII). Our results indicated that the proportion of variation in BMI and lipid levels explained by the gut microbiome was different from that explained by genetic variation. Therefore, we further assessed whether the microbiome could make a significant contribution to the explained variation beyond age, sex, and genetic factors. Our analysis unambiguously showed that age, sex, genetics, and the gut microbiome collectively explained 11.3% of the variation in BMI, 17.1% in triglycerides, and 25.9% in HDL cholesterol, with the microbiome making a significant contribution to the explained variation in BMI (P = 4.1×10^{−7}), triglycerides (P = 4.5×10^{−5}), and HDL (P = 2.7×10^{−4}; Figure 3B). When we included BMI as a risk factor, the total explained variation in lipids increased to 25% in triglycerides, 37.4% in HDL, 22.3% in LDL, and 22.3% in TC (Online Figure IV). The microbiome made a lesser, but still significant, contribution to triglycerides (P = 4×10^{−3}) and HDL (P = 0.026). Our study therefore indicates that the gut microbiome can explain a substantial proportion of the variation, independent of age, sex, BMI, and genetics.

Discussion

Obesity and aberrant levels of blood lipids are associated with a high risk of CVD. Studying the effect of the gut microbiome on BMI and blood lipid levels yields insight into the role of the microbiome in the development of CVD. Although animal studies have shown that microbiota can influence lipid metabolism,39 no large-scale studies have been performed in humans thus far. Here, we investigated the impact of the gut microbiome on BMI and blood lipid levels in 893 human subjects from the LifeLines-Deep cohort. The power of our study is reflected by 3 factors. First, to our knowledge, it is the largest association study linking the gut microbiome to blood lipids in humans to date. Second, our cohort represented a wide range of ages, BMI, and blood lipids, as well as microbial composition. We also had detailed medication information per individual and could exclude those taking lipid-lowering or antibiotic medication. Moreover, we adopted a novel and powerful 2-part model to account for both the binary and quantitative features of microbial data. We established associations for 34 taxonomies with BMI and blood lipid levels, and we estimated that gut microbiota composition can explain ≤6% of the variation in lipid levels and that this effect is independent of age, sex, and host genetics.

Our results for the microbiota associated with BMI are in line with a recent study of 416 twin-pairs from the TwinsUK population38; in particular, we confirmed that lower abundances of families Christensenellaceae, Rikenellaceae, class Mollicutes, genus Dehalobacterium, and kingdom Archaea were associated with a high BMI. Of 22 independent taxa associated with BMI by our study, 16 were also accessed in the TwinsUK study: 11 (68.8%) showed significant association with BMI (P < 0.05) with the same direction of effect as we found (Online Table V). We also identified a correlation of decreased bacterial diversity with increased BMI, which is in line with previous observations.40

However, many of the taxonomies we identified are novel findings. Several of the identified bacteria are known to be involved in the bile acid metabolic pathway. In particular, order Bacteroidales (phylum Bacteroidetes) and family Clostridiaceae (phylum Firmicutes) are both negatively correlated with BMI and triglycerides and known to be involved in bile acid metabolism.41 Bile acid activity of commensal bacteria is involved in a complex interplay with host hepatic enzymes, and together they promote digestion and absorption of dietary lipids.42 Interestingly, several small-scale studies reported lowered cholesterol on using probiotics with bile salt hydrolytic activity.43,44 Our study found support for the role of bacterial bile acids in lipid metabolism. Another pathway enriched in several associated bacteria is short chain fatty acid metabolism. Both orders Bacteroidales and Clostridiaceae, identified in our study; are involved in short chain fatty acid metabolism.41 Short chain fatty acids are produced by microbiota from dietary fibers, effect host body energy homeostasis, and are protective against metabolic syndrome, type 2 diabetes mellitus, and atherosclerosis.15,45–47

To firmly establish the gut microbiome as a risk factor for obesity and aberrant levels of blood lipids, we have been able to estimate that the microbiome could explain 4.57% to 6% of the variation in BMI, triglyceride and HDL, respectively. We did not detect any significant association between the gut microbiome and genetic predisposition to obesity and aberrant levels of blood lipids, suggesting the variation explained by the microbiome is independent of that explained by genetic variants. It should be noted, however, that the genetic risk score was limited to our established 157 lipid-associated SNPs31 and 97 BMI-associated SNPs,34 which together only explain a small proportion of the heritability of lipid levels. We might have missed the effect of other, not yet discovered SNPs. Our risk model included age, sex, genetic variation, and gut microbiome and explained 11.3% of the variation in BMI, 17.1% in triglycerides, and 25.9% in HDL, significantly outperforming the risk model without the microbiome. Because blood lipids and BMI are highly correlated with each other and many associated bacteria were shared, we investigated whether the observed effect of the gut microbiome on lipids might just be the confounded effect of BMI. We showed that by including BMI in the risk model, the gut microbiome made a smaller, but significant, contribution to the variation in triglycerides and HDL, suggesting that the microbiome affects blood lipids partly independently of BMI. Our results therefore indicate that the gut microbiome...
is a potentially important player in blood lipid metabolism. In contrast to genetics, sex, and age (all fixed characteristics), an individual’s microbiota composition can be modified by diet, prebiotics and probiotics, and fecal transplantation. Studies have shown that diet can alter the gut microbiome.54 Our study has not addressed how much of the association we observed between gut microbiome and blood lipids might be explained by diet. A better understanding of this could provide more insights into the role of diet in microbiome and lipid metabolism.

Our study supports the potential of microbiota-modifying intervention to correct lipid disbalance and thereby help prevent CVD. From potential to action, the next steps are to validate the associations we report in independent cohorts and to prove there is a causal axis of gut microbiome–lipids–CVD in functional studies. It is essential to gain more mechanistic insight into the functioning of the gut microbiome, although research in humans is still in its infancy.

The gut microbiomes in our study were profiled by 16s ribosomal RNA gene sequencing. This technology can identify microbial taxonomies and composition, but has limitations in identifying genetically specific species and strains. Furthermore, 16s ribosomal RNA gene sequencing provides little information on bacterial genes and their functions. With the decreasing cost of metagenome sequencing and development of techniques for culturing and for functional studies of gut bacteria, we expect to learn more about the levels of bacterial genes, metabolic pathways, and their functions in the future.

In conclusion, we have observed a strong association between the gut microbial composition and the variation in BMI and blood lipid levels, which is independent of age, sex, and host genetics. This observation provides insight into the microbiome’s role in regulating metabolic processes during the development of CVD. We established associations for a total of 34 intestinal bacteria taxonomies with BMI and blood lipids. We observed that the gut microbiota makes a significant contribution, beyond that of clinical risk factors and genetics, to the individual variance seen in BMI and to the blood levels of triglycerides and HDL, but that it has little effect on LDL or TC levels. Our results highlight the potential of therapies that alter the gut microbiome to control body mass, triglycerides, and HDL in CVD prevention.

In moving from potential to action, it will be essential to identify the causal axis of microbiome–lipids–CVD and to gain more mechanistic insight into the gut bacteria functions.

Acknowledgments
We thank the LifeLines-DEEP participants and the LifeLines staff in Groningen for their collaboration. We thank Jackie Senior and Kate Mc Intyre for editing the article and Mathieu Platteel for practical and analytic work.

Sources of Funding
This project was funded by grants from the Top Institute Food and Nutrition, Wageningen, to C. Wijmenga, A. Zhernakova, and E.F. Tigchelaar (GH001), the Netherlands Organization for Scientific Research to J. Fu (NWO-VIDI 864.13.013), CardioVascular Onderzoek Nederland to M.H. Hofker and A. Zhernakova (CVON 2012-03), and NWO (Netherlands Organization for Scientific Research) grants to L. Franke (ZonMW-VIDI 917.14.374) and R.K. Weersma (ZonMW-VIDI 016.136.308). A. Zhernakova holds a Rosalind Franklin Fellowship (University of Groningen) and M.C. Centi holds a postdoctoral fellowship from the Fundación Alfonso Martín Escudero. This research received funding from the European Community’s Health Seventh Framework Programme (FP7/2007–2013, grant agreement 259867).

Disclosures
None.

References
The bacterial community in the human gut (known as the microbiome) has been referred to as an extra organ, or the second human genome, because of its important role in an individual's health. As such, it is referred to as an extra organ, or the second human genome, because of its important role in an individual's health. We now know that the microbiome plays a significant role in various aspects of human health, including digestive function, immune system development, and the prevention and treatment of various diseases. This is why the study of the microbiome is becoming increasingly important in the field of human health research.

To understand the microbiome, it is important to consider its composition and diversity. The human microbiome is composed of trillions of microorganisms, including bacteria, viruses, and fungi, which inhabit different parts of the human body, such as the skin, mouth, gut, and respiratory tract. The diversity of the microbiome varies between individuals and can be influenced by various factors, such as genetics, diet, and lifestyle.

The microbiome is also influenced by environmental factors, such as exposure to antibiotics and other antimicrobial agents, which can alter the composition of the microbiome and affect its function. It is important to consider these factors when studying the microbiome, as they can influence the results of microbiome studies.

In conclusion, the study of the microbiome is a rapidly growing field of research, with significant implications for human health. As we continue to learn more about the microbiome, we can expect to see new and innovative approaches to the prevention and treatment of various diseases.