Title

Role of Gut Microbiota in Statin-Associated New-Onset Diabetes-A Cross-Sectional and Prospective Analysis of the FINRISK 2002 Cohort.

Journal

Arteriosclerosis, Thrombosis, and Vascular Biology, 44(2)

Authors

Koponen, Kari

Kambur, Oleg

Joseph, Bijoy

et al.

Publication Date

2024-02-01

DOI

10.1161/ATVBAHA.123.319458

Peer reviewed

Arteriosclerosis, Thrombosis, and Vascular Biology

CLINICAL AND POPULATION STUDIES

Role of Gut Microbiota in Statin-Associated New-Onset Diabetes—A Cross-Sectional and Prospective Analysis of the FINRISK 2002 Cohort

Kari Koponen, Oleg Kambur, Bijoy Joseph, Matti O. Ruuskanen, Pekka Jousilahti, Rodolfo Salido, Caitriona Brennan, Mohit Jain, Guillaume Meric, Michael Inouye, Leo Lahti, Teemu Niiranen, Aki S. Havulinna, Rob Knight, Veikko Salomaa

BACKGROUND: Dyslipidemia is treated effectively with statins, but treatment has the potential to induce new-onset type-2 diabetes. Gut microbiota may contribute to this outcome variability. We assessed the associations of gut microbiota diversity and composition with statins. Bacterial associations with statin-associated new-onset type-2 diabetes (T2D) risk were also prospectively evaluated.

METHODS: We examined shallow-shotgun-sequenced fecal samples from 5755 individuals in the FINRISK-2002 population cohort with a 17+-year-long register-based follow-up. Alpha-diversity was quantified using Shannon index and beta-diversity with Aitchison distance. Species-specific differential abundances were analyzed using general multivariate regression. Prospective associations were assessed with Cox regression. Applicable results were validated using gradient boosting.

RESULTS: Statin use associated with differing taxonomic composition (R^2, 0.02%; q=0.02) and 13 differentially abundant species in fully adjusted models (MaAsLin; q<0.05). The strongest positive association was with *Clostridium sartagoforme* (β=0.37; SE=0.13; q=0.02) and the strongest negative association with *Bacteroides cellulosilyticus* (β=−0.31; SE=0.11; q=0.02). Twenty-five microbial features had significant associations with incident T2D in statin users, of which only *Bacteroides vulgatus* (HR, 1.286 [1.136–1.457]; q=0.03) was consistent regardless of model adjustment. Finally, higher statin-associated T2D risk was seen with *[Ruminococcus] torques* ($\Delta HR_{statins}$, +0.11; q=0.03), *Blautia obeum* ($\Delta HR_{statins}$, +0.06; q=0.01), *Blautia* sp. *KLE 1732* ($\Delta HR_{statins}$, +0.05; q=0.01), and beta-diversity principal component 1 (ΔHR_{statin}, +0.07; q=0.03) but only when adjusting for demographic covariates.

CONCLUSIONS: Statin users have compositionally differing microbiotas from nonusers. The human gut microbiota is associated with incident T2D risk in statin users and possibly has additive effects on statin-associated new-onset T2D risk.

GRAPHIC ABSTRACT: A graphic abstract is available for this article.

Key Words: diabetes mellitus, type 2 ■ metagenome ■ microbiota ■ prospective studies ■ statins

Dyslipidemia and cardiovascular diseases are the leading causes of morbidity and mortality worldwide.[1] Inhibitors of HMG-CoA (3-hydroxy-3-methylglutaryl coenzyme A) reductase, also known as statins, are the most common pharmacotherapeutic tools used in their prevention and treatment.[2,3] Statins' primary function, that is, upregulation of hepatic LDL (low-density lipoprotein)-cholesterol receptors, has proven to be an effective way of lowering plasma LDL-cholesterol and consequently the risk for cardiovascular diseases.[3] However, to some, treatment comes with side effects in the form of an increased risk for type-2 diabetes (T2D), among other complications.

See accompanying editorial on page 488

Correspondence to: Kari Koponen, MSc, PO Box 30, 00271 Helsinki, Finland. Email kari.koponen@alumni.helsinki.fi
Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/ATVBAHA.123.319458.
For Sources of Funding and Disclosures, see page 485–486.

Arterioscler Thromb Vasc Biol is available at www.ahajournals.org/journal/atvb

Nonstandard Abbreviations and Acronyms

HDL	high-density lipoprotein
HMG-CoA	3-hydroxy-3-methylglutaryl coenzyme A
LDL-C	low-density lipoprotein cholesterol
SANOD	statin-associated new-onset diabetes
T2D	type 2 diabetes

The risk for statin users to develop T2D ranges between 9% and 45% according to meta-analyses of observational studies and randomized controlled trials.[4,5] A wide variety of candidate mechanisms for explaining the pathology of statin-associated new-onset diabetes (SANOD) have been presented, including induced hepatic gluconeogenesis,[6] disturbed glycemic control in hepatocytes,[7] insulin resistance via inhibited signal transduction,[8–10] and impaired insulin secretion via multiple intracellular mechanisms.[9,11–13] An epigenetic component via microRNA signaling has also been presented.[14]

The term microbiota entails the community of bacteria, archaea, viruses, fungi, and other single-cell organisms inhabiting our bodies, as well as their vast collective of genes, which complement the metabolic capabilities of the human body.[15] Recent advances in technology and methodology surrounding microbiota have presented novel ways to evaluate health associations and risk factors.[16,17] Most of this research has focused on the fecal gut microbiota, the microbes living in the lumen of the large intestine, which acts as a proxy for the microbial constituents living on the epithelial lining of the intestine. Given its dynamic nature, the gut microbiota holds potential for explaining some of the variability in the development of SANOD.

Increasing evidence indicates that the human gut microbiota is a contributor to interindividual variation in statins' therapeutic effects.[18] It has also been connected with overall development of T2D in the general population.[19] However, knowledge on the associations between the human gut microbiota and T2D risk in statin users specifically, and whether the gut microbiota modulates this risk, is still limited. Statin use has been noted to be associated with higher richness and lower dysbiosis of the gut microbiota via lower occurrence of a proinflammatory enterotype.[20] This, in turn, is reflected in lower levels of systemic inflammation, the role of which in T2D pathophysiology is of increasing interest.[20,21] Also, microbially derived metabolites, mainly short-chain fatty acids, have been shown to reduce hepatic fat content[22] and improve peripheral insulin sensitivity.[23,24] Statins have been previously associated with an increased potential to produce short-chain fatty acids.[25] Adverse associations have also been presented, where statin use in the presence of microbial endotoxins has been shown in animal models to induce insulin resistance in adipose tissue.[26]

Highlights

- Statin use associates with compositional differences in the gut microbiota at the population level.
- Incidence of new-onset type-2 diabetes in statin users is linked with gut microbiota composition.
- Higher *[Ruminococcus] torques*, *Blautia obeum*, and *Blautia* sp. *KLE 1732* abundances associate in an additive manner with higher risk for statin-associated new-onset type-2 diabetes.

While statins increase the risk of new-onset diabetes, there seems to be potential for alleviating or aggravating this risk via the gut microbiota, although direct study of these associations in statin users is still scant. Thus, we set out to assess whether there are microbial components that associate with SANOD risk. Our aims for this study were the following: (1) to analyze cross-sectional associations of statin treatment with human gut microbiota diversity and composition at the population level. (2) To analyze whether these microbial associations are connected with incident T2D risk during follow-up. (3) To prospectively analyze possible shifts of SANOD risk in the presence of these microbial features.

METHODS

Resources needed for the replication of this study are listed in Major Resources Table in the Supplemental Material. Data described in the article are available from Finnish Institute for Health and Welfare Biobank based on a written application as instructed on the website of the Biobank (https://thl.fi/en/web/thl-biobank). The corresponding author had full access to all the data in the study and took responsibility for their integrity and analysis. The data are not freely available because they contain information that could compromise research participant privacy/consent. Code used in the analyses is openly available at 10.5281/zenodo.7503450. FINRISK 2002 was approved by the Ethical Committee on Epidemiology and Public Health of the Helsinki and Uusimaa Hospital District (decision number 87/2001), and the participants gave informed consent. The study was conducted according to the World Medical Association's Declaration of Helsinki on ethical principles.[27]

Cohort and Sample Selection

The FINRISK surveys, conducted every 5 years by the Finnish Institute for Health and Welfare from 1972 until 2012, were designed to provide population-level information on risk factors for noncommunicable diseases, health behavior, and their changes in adult Finns between the ages 25 and 74.[28]

This study utilized data from the FINRISK 2002 cohort, which consisted of 8738 individuals. The sample for the current study consisted of 5755 individuals after exclusions. The following exclusions were used: individuals with missing fecal samples, missing data on statin use, no permission for register access from the participant, pregnant women and individuals who had used antibiotics within 6 months before the baseline

investigation. A flowchart for the sample selection is shown in Figure S1. The antibiotics used in this filtering included antibacterials for systemic use (ATC: J01), antimycotics for systemic use (ATC: J02), antimycobacterials (ATC: J04), and antivirals for systemic use (ATC: J05).

Prospective analyses for SANOD risk in statin users were done in 2 subsamples. The first subsample included T2D-free individuals at baseline and individuals who used statins at baseline or started using them during follow-up (n=1823). The second subsample for SANOD-analyses consisted of individuals who used statins and were T2D-free at baseline and those who never used statins and were T2D-free at baseline (n=3922). Flowcharts for subsamples' sample selections are displayed in Figure S1.

Follow-Up

In Finland, each resident is assigned a unique personal identity number at birth or after immigration, which ensures reliable linkage to electronic health records. The Finnish national health registers ensure practically 100% coverage of all major health events and deaths over an individual's lifetime. The participants of the current study have been followed through December 31, 2019 (17+ years from baseline). Only those few participants who had moved permanently abroad before the disease event or December 31, 2019 were lost to follow-up. The reliability of the Finnish health care registers has been documented.[29,30]

Baseline Data Collection, Covariates, and Other Variables

The FINRISK 2002 study utilized several different data collection methods. Individuals filled in a questionnaire at home answering various socioeconomical and lifestyle-related questions. They also participated in a health examination where the questionnaire was inspected, anthropometric features and blood pressure were measured, and a blood sample was drawn by trained health care practitioners. Fecal samples were collected from willing donors. Medical history and information on drug use were collected from national registries and linked with pseudonymized FINRISK study identifiers with the help of the previously mentioned national personal identity numbers.

Statins included HMG-CoA reductase inhibitors (ATC: C10AA) and combinations of various lipid modifying agents (ATC: C10BA) as listed in the ATC/DDD Index of the WHO.[31] Covariates used included common gut microbiota perturbing factors that were baseline age, sex, smoking, medications, prevalent diseases, and body mass index (BMI).[32,33] An individual was flagged as a smoker if they had smoked within 6 months of baseline investigation. Adjusted medications included dummy variables for use of blood pressure medication (reported on the questionnaire), psychotropics (ATC: N05 and N06), and metformin use (ATC: A10BA02). An individual was flagged as a medication user if they had a registered purchase within 4 months before baseline investigation, and at least 3 separate purchase events. Adjusted prevalent diseases were coronary heart disease, ischemic stroke, and T2D. Metformin and prevalent T2D were excluded from adjustment in prospective analyses of SANOD risk. BMI was calculated as weight (kg) divided by height (m) squared. All continuous covariates were converted into *Z* scores.

Fecal Samples

Those willing to donate a fecal sample were given a sampling kit alongside detailed instructions and a prepaid postal parcel for returning the samples via mail. Samples were collected into 50-mL Falcon tubes without a stabilizing solution and mailed to the laboratory under typical Finnish winter conditions between Monday and Thursday. The samples were frozen immediately upon reception and stored unthawed at −20 °C until sequencing in 2017.

Sequencing was performed at the University of California, San Diego. Microbial DNA was extracted using a MagAttract PowerSoil kit (Qiagen). Library preparation was carried out using a miniaturized version of the Kapa HyperPlus Illumina-compatible preparation kit (Kapa Biosystems). Samples were normalized to 5 ng with an Echo 550 acoustic liquid handling robot followed by enzymatic fragmentation, end-repair, and adapter-ligation reactions performed with a Mosquito HV liquid handling robot (TTP Labtech, Inc). PCR-amplified and barcoded libraries were quantified using PicoGreen assay.

Sequencing adapters were based on the iTru protocol[34] and sequencing itself was performed with an Illumina Hi-Seq 4000 (Illumina Inc.) for 150-bp reads. Average read count per sample was 900 000 reads. After removing human DNA reads by mapping them against the reference genome assembly GRCh38 using Bowtie2,[35] the reads were classified taxonomically using SHOGUN v1.0.5[36] by mapping them against bacterial and archaeal genomic sequences found in the NCBI RefSeq database Version 82 (May 8, 2017).

To quantify differences in community composition, we used Shannon index for alpha-diversity.[37] Aitchison distance for beta-diversity,[38] and centered log-ratio (CLR) transformed relative abundances for species-specific differential abundance comparisons. Relative abundances were calculated by scaling the raw counts to the total sum of reads. Alpha-diversity was calculated from unfiltered species-level raw counts. To exclude very rare taxa, the species-specific analyses were filtered to a core microbiota that included species with a minimum of 1% prevalence and a minimum of 0.01% relative abundance. The Aitchison distances were also broken down into principal components, of which the first 2 were chosen to be used in downstream analyses. Finally, all microbiota features were converted into *Z* scores before analysis when they were used in a predictor capacity.

Final preparation and analysis of the microbial counts data was done in R version 4.1.3.[39] Central packages were phyloseq (1.38.0), vegan (2.3.2), and microbiome (1.16.0).[40–42]

Traditional Statistical Methods

Covariate effects in statistical models were evaluated in a stepwise manner. We report the results for demographically adjusted and fully adjusted models. Fully adjusted models included demographic variables (ie, baseline age and sex) and variables for BMI, smoking, medications, and prevalent conditions for cardiovascular disease, T2D, and stroke. The variables for metformin and prevalent T2D were removed from covariates in the analyses for SANOD. To account for the large amount of tests in our analyses, we corrected all *P* values of the analyses for multiple testing by using the Benjamini-Hochberg false discovery rate procedure.[43] The significance threshold was set at an false discovery rate–corrected q value of 0.05.

Univariate linear regression was used to assess alpha-diversity associations. Analysis of similarities was used to apply a statistical test to assess whether Aitchison distances between statin users and nonusers differed.[44] Permutational multivariate ANOVA was used, with the number of permutations set to 999, to determine the proportion of variance each gut microbiota feature explains in the Aitchison-distances between samples.[45] A visual inspection of the ordination of the first 2 principal components of Aitchison-distances indicated possible clustering between groups for statin users versus nonusers, which was tested for statistical significance with the factorfit-function from the vegan R-package version 2.6.2.[42] These principal components were also used as predictors in downstream analyses of SANOD risk. General multivariate regression was used for species-specific associations, as implemented in the MaAsLin2 R-package Version 1.10.0.[46] Cox regression was used in all prospective analyses.

The assumption of permutational multivariate analysis of variance for homogeneity of variances was checked. The proportionality of hazards assumption for each Cox model was tested with a Schoenfeld test, and the results were verified by inspecting residual plots.

Machine Learning Methods

For applicable analyses where the final tested outcome was binary (ie, analysis for SANOD), we used extreme gradient boosting utilizing XGBoost R library (Version 0.90.0.2).[47] Samples were randomly partitioned into a training set (70%) and a validation set (30%). Training set was used for discovery and validation set for evaluation. Prediction models we developed and tested using simple cross-validation with 5 iterations of the training set. The optimal models were then tested in the validation set, where their performance was assessed.

Incident T2D was set as the label. Included covariates were the same sets of variables as in the fully adjusted traditional counterpart of the analysis. Conventional covariates were assessed with and without the microbiome features. Microbial features were preselected using the same filtering thresholds as in traditional taxa-specific analyses. Before ML-step, microbial relative abundances were CLR-transformed. For the final assessment, we extracted the top 25 species contributing to information gain in the validation set and compared their overlap with the results of the traditional analysis results.

Resampling strategy included 5-fold cross-validation, which was performed in a stratified manner (based on participant sex). Division to training and testing sets was performed before hyperparameter optimization, so there was no data leakage. Hyperparameters were optimized among 10 randomized grids of the parameter space. Both, tree-based and linear models were assessed. Other optimized hyperparameters included minimum weight, number of leaves of each tree, subsample ratio of the training sample, and percentage of features used for building each tree. The following ranges were considered for hyperparameters: Eta: 0.001 to 0.3; gamma: 0.1 to 5; max_depth: 2 to 10; scale_pos_weight: 1 to 12; min_child_weight: 1 to 10; subsample: 0.2 to 0.8; nrounds: 50 to 5000; and colsample_bytree: 0.2 to 0.9. Binary logistic regression was used as an objective function. Model performance was evaluated using binary classification error rate. Error rate was calculated as $\frac{n_{wrong\ \ cases}}{n_{all\ \ cases}}$. For the purpose of prediction, values >0.5 were considered as positive instances, and values <0.5 as negative instances. The number of decision trees for the final models were set to 100 (default) and learning rate to 0.1. These parameters were not optimized and remained the same for all runs.

The values for optimized parameters of the models were the following. For analysis of all covariates and microbiota learner was tree-based, for example, consisted of tree ensemble. Maximal number of leaves for each tree (max_depth) was set to 5, and minimal weight (min_child_weight) to 1.04. Subsample ratio of the training sample (parameter subsample) was set to 0.705, and percentage of features used for building each tree (colsample_bytree) to 0.706. For the analysis of all covariates without microbiota, the learner was of linear type (gblinear), for example, the model consisted of weighted sum of linear functions. Subsample ratio of the training sample (parameter subsample) was set to 0.711.

RESULTS

Descriptive Statistics

Baseline statin users were significantly older and consisted of more men compared with nonusers. Smoking was less frequent among statin users. Statin users were also more frequently users of antihypertensive drugs, metformin, and psychotropics. Systolic blood pressure was higher in statin users. Cholesterol levels (ie, total-, non-HDL-, LDL-, and HDL [high-density lipoprotein] cholesterols) and apolipoprotein Apo B and Apo A1 levels were all lower in statin users. Plasma triglyceride levels were higher in statin users than in nonusers. The prevalence for coronary heart disease, stroke, and T2D were greatly elevated in statin users when compared with nonusers. A detailed description of the main sample can be found in the Table. The descriptive statistics for subsamples are listed in Table S1.

Cross-Sectional Associations With Statin Use

Microbiota alpha-diversity was not significantly higher in statin users versus nonusers (mean Shannon-diversity: 3.47±0.42 versus 3.43±0.42; q=0.14). For beta-diversity, statin use associated significantly with the first 2 PCs of Aitchison distances (q=0.001, Figure 1). Additionally, a significant analysis of similarities result (analysis of similarities: R=0.07; q=0.01) indicated that statin users and nonusers had compositionally differing gut microbiotas. Permutational multivariate analysis of variance tests yielded significant results in both demographically and fully adjusted configurations (R^2=0.02%; q=0.02, fully adjusted), including a significant interaction between baseline age and statin use (R^2=0.02%; q=0.03, fully adjusted), meaning that baseline age and statin use both individually and together may be associated with compositionally different microbiotas irrespective of other microbiota influencing covariates. However, a significant dispersion test (F=10.14; *P*=0.004) signaled that these

Table. Descriptive Statistics for the Main Sample Divided Into Statin Users and Nonusers at Baseline

	Statin users	Statin nonusers	*P* value
n	393 (6.8%)	5362 (93.2%)	-
Baseline age, y	61.06±7.68*	48.81±12.79*	4.44×10^{-113}*
Men, n	220 (56.0%)*	2523 (47.1%)*	6.26×10^{-4}*
Current smoking, n	55 (14.0%)*	1309 (24.4%)*	3.52×10^{-6}*
Antihypertensive drugs use, n	200 (50.9%)*	666 (12.4%)*	3.45×10^{-94}*
Metformin use, n	33 (8.4%)*	52 (1.0%)*	4.87×10^{-32}*
Psychotropics use, n	48 (12.2%)*	340 (6.3%)*	7.40×10^{-6}*
BMI, kg/m^2	28.89±4.35*	26.84±4.64*	8.44×10^{-18}*
SBP, mm Hg	143.91±20.91*	135.68±20.04*	2.73×10^{-13}*
DBP, mm Hg	80.32±11.16	79.43±11.19	0.13
Total cholesterol, mmol/l	5.11±0.97*	5.65±1.06*	5.31×10^{-24}*
Non–HDL cholesterol, mmol/l	3.74±0.94*	4.14±1.09*	3.86×10^{-15}*
LDL-cholesterol, mmol/l	2.94±0.72*	3.43±0.90*	4.14×10^{-32}*
HDL-cholesterol, mmol/l	1.37±0.36*	1.51±0.41*	8.14×10^{-13}*
Triglycerides, mmol/l	1.71±1.07*	1.40±0.93*	5.04×10^{-8}*
Apo B, g/L	0.95±0.20*	1.00±0.25*	9.78×10^{-6}*
Apo A1, g/L	1.49±0.27*	1.54±0.28*	3.28×10^{-4}*
CRP, mg/L	2.30±3.55	2.38±4.45	0.69
Prevalent CHD, n	100 (25.5%)*	57 (1.1%)*	2.07×10^{-180}*
Prevalent stroke, n	23 (5.9%)*	44 (0.8%)*	2.80×10^{-19}*
Prevalent diabetes, n	50 (12.7%)*	109 (2.03%)*	9.60×10^{-36}*

Descriptive statistics for the main sample after exclusions (n=5755). Values for continuous variables are means±SD and for categorical variables the number of observations and their proportion in their respective population in parentheses. Groups were compared using Welch *t* test for continuous variables and for categorical variables using χ^2 test. BMI indicates body mass index; DBP, diastolic blood pressure; CHD, coronary heart disease; CRP, C-reactive protein; HDL, high-density lipoprotein; LDL, low-density lipoprotein; and SBP, systolic blood pressure.

*P values smaller than the significance threshold (P<0.05).

results might be also due to a heterogenous dispersion effect between groups. No interaction between sex and statin use was detected.

In species-specific analyses, we identified 26 species that associated significantly with statin use in the demographically adjusted model (Table S2) and 13 species in the fully adjusted model (Figure 2; Table S2). Of these, the strongest positive association with statins in the fully adjusted model was for *Clostridium sartagoforme* (β=0.37; SE=0.13; q=0.02) and the strongest negative association was for *Bacteroides cellulosilyticus* (β=−0.31; SE=0.11; q=0.02).

Prospective Statin Type-2 Diabetes Risk Analyses

Type-2 Diabetes Risk Analyses in Statin Users

Incidence of T2D was significantly higher among statin users in a χ^2 test (16.3% versus 5.2%; $q=1.66\times10^{-40}$). However, alpha-diversity did not associate with incident T2D risk in neither the demographically nor fully adjusted Cox models. In beta-diversity analyses PC1 had a statistically significant positive (higher risk) association with incident T2D in both the demographically (HR, 1.30 [1.15–1.46]; $q=2.10\times10^{-4}$) and the fully adjusted (HR, 1.20 [1.07–1.35]; $q=9.62\times10^{-3}$) models (Figure S2; Table S3).

Species-specific analyses of the gut microbiota with the demographically adjusted model revealed 24 species with significant associations (Figure S2; Table S3). The strongest association with lower T2D risk was observed for *Butyrivibrio crossotus* (HR, 0.80 [0.71–0.90]; q=0.01) and strongest association with higher risk was observed for *Bacteroides vulgatus* (HR, 1.34 [1.18–1.52]; $q=1.91\times10^{-3}$). The fully adjusted model identified 1 species, namely *B. vulgatus* (HR, 1.29 [1.14–1.46]; q=0.03).

Additionally, of the 24 significant species found in Cox regression models 4 were found in the top 25 microbiome features of a fitted XGBoost model, 3 of which were in the top 3. These species, in descending order of information gain, were *Bacteroides stercoris*, *[Clostridium] citroniae*, *Clostridium saccharobutylicum*, and *B. vulgatus*.

Type-2 Diabetes Risk Analyses in a Combined Sample

To determine whether the significant gut microbiota features observed in incident T2D analyses in statin users modified the inherent T2D risk associated with statin use, we assessed their influence on statin HR in a sample where statin nonusers were added and follow-up statin users were removed. Follow-up statin users were left out due to their inclusion causing a major violation of the proportionality of hazards assumption. These combined-sample analyses showed no significant associations between alpha-diversity and incident T2D. Neither did alpha-diversity shift SANOD risk (interpreted as $\Delta HR_{statins}$ of >±0.05).

For beta-diversity, PC1 displayed a significant association with incident T2D in the demographically adjusted (HR, 1.20 [1.05–1.37]; q=0.03) but not in the fully adjusted model (Figure 3; Table S4). No significant interactions were detected. SANOD risk increased by 0.07 from 4.31 to 4.38 in the demographically adjusted model once PC1 was added as a covariate.

In species-specific analyses we found 14 of the original 24 significant species that were associated with incident T2D risk in the demographically adjusted model (Figure 3; Table S4). The fully adjusted model found no significantly associated species. The strongest association with lower risk was observed for *Ruminococcus champanellensis* (HR, 0.78 [0.69–0.88]; $q=1.31\times10^{-3}$, demographically adjusted) and the strongest association with higher risk for *Coprococcus comes* (HR, 1.25 [1.08–1.44]; $q=7.69\times10^{-3}$, demographically adjusted).

Out of the 14 significantly associated species, seven modified SANOD risk by at least ±0.05, of which in 3

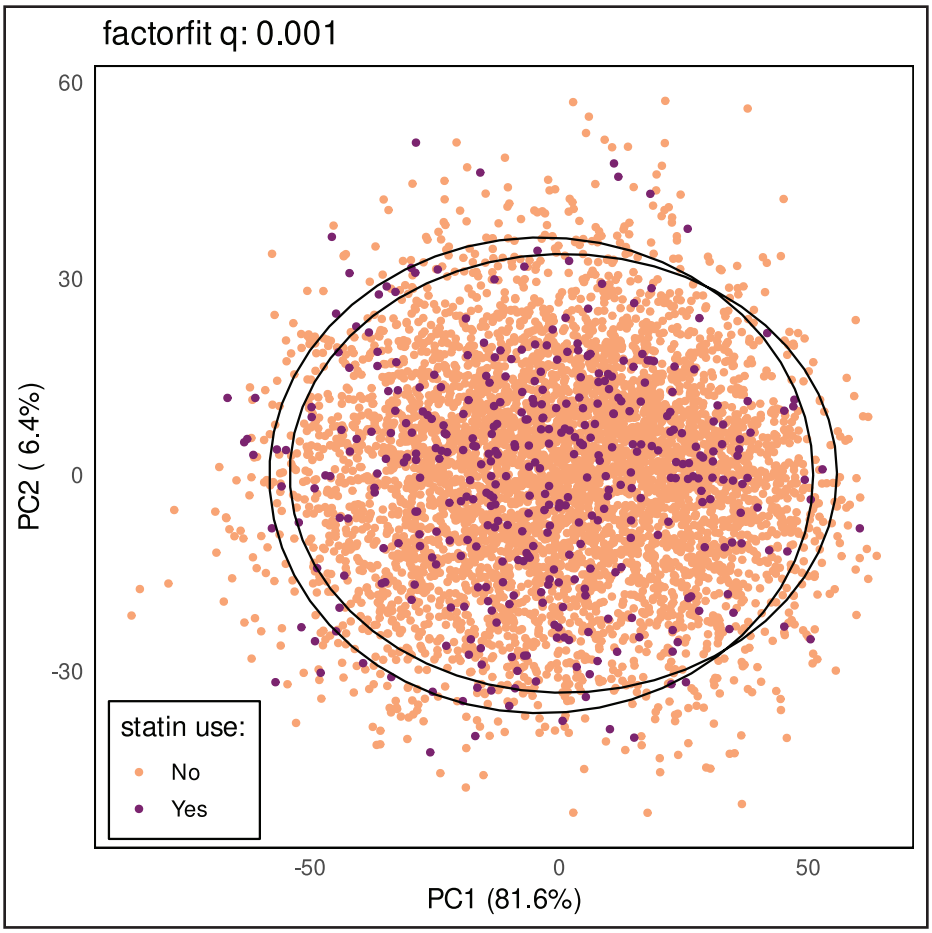

Figure 1. Ordinations of beta-diversity principal components 1 and 2 for statin users and nonusers.
Ordination plot based on the first 2 principal components based on Aitchison distances. Statin users are indicated with purple. The false discovery rate corrected *P* values (Q value) for a factorfit-test between the groups is displayed above the plot. Both ellipsoids depict the spread of 95% of the respective groups according to Student *t*-distribution. The amount of variance explained by each respective principal component is marked on their axis labels. PC indicates principal component.

cases the direction of $\Delta HR_{statins}$ was in harmony with the direction of the HR of the species itself (ie, a species that associated with a higher/lower risk also had a positive/negative $\Delta HR_{statins}$, respectively). These 3 species were *[Ruminococcus] torques* (HR, 1.17 [1.03–1.32]; q=0.03, $\Delta HR_{statins}$, +0.11), *Blautia obeum* (HR, 1.22 [1.06–1.41]; q=0.01, $\Delta HR_{statins}$, +0.06), and *Blautia* sp. *KLE 1732* (HR, 1.24 [1.09–1.41]; q=0.01, $\Delta HR_{statins}$, +0.05). Diabetes-free survival curves for the demographically adjusted microbial features that were significant with a harmonious $\Delta HR_{statins}$ are displayed in Figure 4.

DISCUSSION

Our study is the first to assess the impact of statin medication on the human gut microbiota in an epidemiological setting with a large population-based sample using shotgun sequencing. It supports previous findings on the connections between statin treatment and a differing gut microbiota composition. Additionally, we uncovered novel associations between the inherent T2D risk associated with statin use and specific gut microbiota features that associate with risk in an additive manner. Especially notable are the associations with increased risk for statin-associated new-onset type-2 diabetes of *[Ruminococcus] torques*, *B. obeum*, and *Blautia* sp. *KLE 1732*.

Our observations that statin use associated with a compositionally differing gut microbiota and was approaching significance in its association with higher alpha-diversity, are in line with previous findings, although it is important to note that that our results for beta-diversity might have been influenced by heterogenous dispersion of sample variances between statin users and nonusers.[1,20,48,49] Higher richness and diversity of the gut microbiota has been associated inversely with BMI and triglyceride levels as well as positively with HDL-cholesterol in a cohort of 893 individuals.[50] However, on the species level our study gives mixed results between beneficial and detrimental taxa. Statin use was associated negatively with multiple cellulolytic and polysaccharide utilizing symbionts such as *B. cellulosilyticus*[51] and

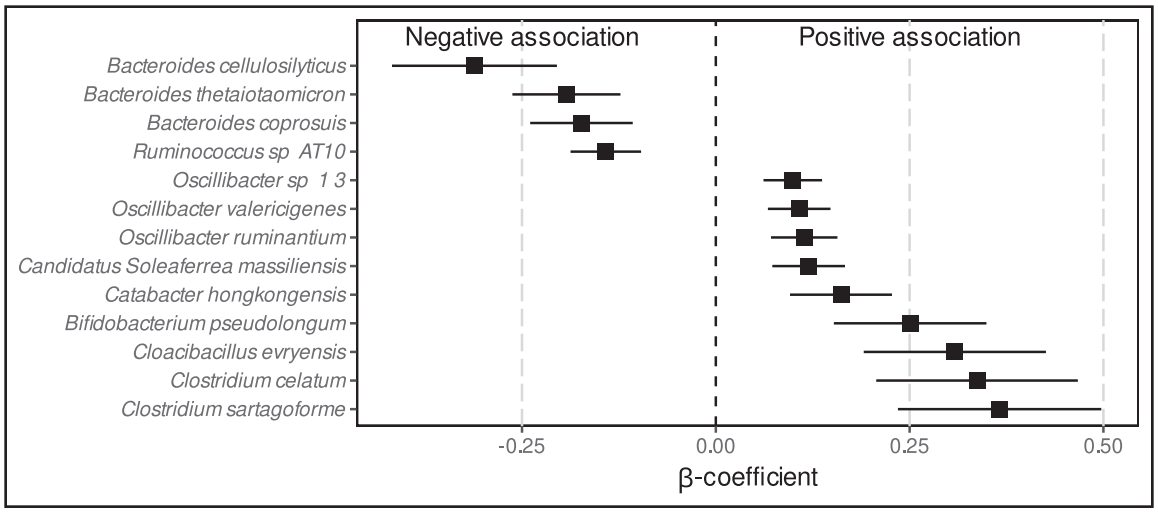

Figure 2. Statistically significant associations of bacterial species with statin medication.
Forest plot displaying statistically significant regression coefficients along with their SEs (q<0.05) of individual species with statin use in general multivariate regression models that were adjusted for baseline age, sex, body mass index, smoking, antihypertensive drugs, metformin, psychotropics, prevalent coronary heart disease, prevalent ischemic stroke, and prevalent type-2 diabetes. All continuous predictors were converted to *Z* scores and individual species relative abundances were centered log-ratio-transformed. The effect size is the scaled unit change of centered log-ratio transformed relative abundances of the species. Q values: false discovery rate corrected *P* values.

Bacteroides thetaiotaomicron.[52,53] It was also positively associated with both potential pathogens and symbionts such as multiple *Clostridium* species and *Bifidobacterium pseudolongum*, a potentially therapeutic bacterium for metabolic disorders.[54] These results might be a feature of our cross-sectional setting, where perceived positive associations of statin medication on the gut microbiota are difficult to disentangle from the negative underlying conditions of the participants needing statin medication.

Our prospective results that showed an association between gut microbiota composition of statin users and higher risk for incident T2D, irrespective of covariate effects, indicate that the gut microbiota plays a role in incident T2D risk. Noteworthy are the identified multiple species that modified the future risk of developing SANOD over a lengthy 17+ years long follow-up time. Many of these same species were identified as being connected with overall T2D risk in a similar manner in a study by Ruuskanen et al[55] in this same cohort. Especially noteworthy were the *Blautia* and *Ruminococcus* species in our results, which were associated with higher risk for future SANOD. These same 2 species were among the top microbial features contributing to higher incident T2D risk in Ruuskanen et al's analyses in an eastern Finnish subsample of 3871 individuals. *Blautia* and *Ruminococcus* have been identified as 2 of the few genera that consistently across studies associate with a higher T2D risk.[19] Our study adds to this knowledge by suggesting that *Blautia*, along with *[Ruminococcus] torques*, might also have additive links with the inherent risk of statin medication for developing T2D. Unfortunately, our findings did not survive full adjustment, hinting that dysbiosis of the microbiota may also play a role. Alternatively, as explained above, adjustment for obesity and factors correlated with it may be overadjustment, that is, adjusting for factors that are along the same biological pathway.

The taxonomic statuses of some *Ruminococcus* species, including *[Ruminococcus] torques*, have been disputed and have since been renamed as members of the novel *Mediterraneibacter* genus.[56] Findings of previous studies on *Ruminococcus* and T2D are mixed. The majority of studies has reported consistently associations with increased risk between T2D and the *Ruminococcus* genus, although when moving from the genus-level to the species-level results are inconclusive.[19] As noted above, many *Ruminococcus* species have been renamed as members of the novel *Mediterraneibacter* genus and thus, species-specific variation is a probable explanation for the mixed results. Additionally, varying outcomes have been hypothesized to be also due to heterogeneity of the assessed exposures (ie, effects of metformin treatment versus bariatric surgery, etc.).[19] Indeed, Ruuskanen et al[55] found that *R. champanellensis*, the species with the strongest association with lower incident T2D-risk in our study, was negatively associated with T2D risk in their eastern Finnish subsample. It is noteworthy, however, that this was only the case if participants who developed T2D in the first 2 years were excluded. Also, a human gut enterotype where *Ruminococcus* was among the top contributors has been noted to be more prevalent in individuals with lower BMI.[20] In our results, multiple members of the *Ruminococcus* genus associated with a lower risk for incident T2D by themselves but shifted SANOD risk toward higher risk. Given the available information, detailed study of the relation of genera *Ruminococcus* and *Mediterraneibacter* and their

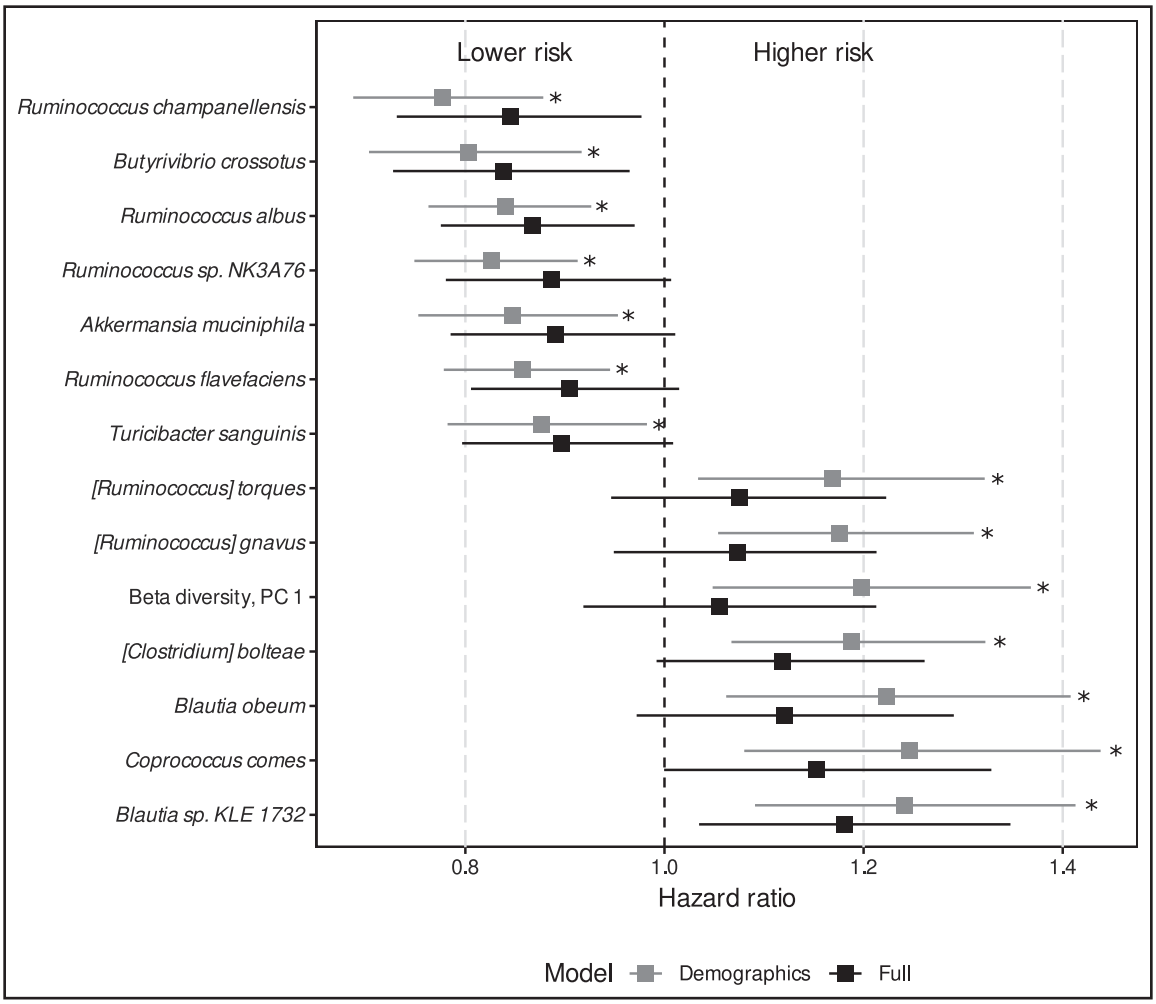

Figure 3. Sample-wide associations of gut microbiota features with type-2 diabetes (T2D) risk.
Forest plot displaying associations of gut microbiota features with incident T2D risk in the whole sample using Cox regression. HR is depicted with a point and the 95% CI by bars. Results from demographically adjusted models (baseline age and sex) are indicated by gray and those from fully adjusted models (demographic covariates, body mass index, smoking, antihypertensive drugs, psychotropics, prevalent coronary heart disease, and prevalent ischemic stroke) are indicated by black. All continuous predictors were converted into *Z* scores before analysis and species relative abundances were centered log-ratio–transformed before *Z*-conversion. Beta-diversity was quantified using Aitchison distance. Statistically significant results ($q<0.05$) are indicated with an asterisk (*). Q values: false discovery rate corrected *P* values. PC indicates principal component.

underlying species with SANOD risk and overall T2D risk is warranted.

In this study, we utilized shallow-shotgun sequenced whole metagenomic data, which gives us far better resolution for taxa identification and classification, as opposed to the often-used 16S ribosomal RNA-based sequencing methods, especially on the species level.[57] This in turn increases confidence in species-specific results. Another major strength of our study is the large and population-based study sample, which is still a rare feature for microbiota studies. This, coupled with a 17+-years long follow-up period, gives us the sufficient statistical power to uncover prospective microbiota-phenotype associations that might not be possible to detect with smaller sample sizes. It is important to note that many of the associations we detected started to emerge only after the 5-year mark of follow-up, similar to findings of Ruuskanen et al.[55] Another strength of our study is the reliable end point data enabled by Finnish nationwide health care registers that ensure practically a 100% coverage of all major health events.[29]

Despite its strengths, our study has also limitations. The cross-sectional results do require confirmation and clarification in prospective settings. Another limitation, stemming from the largely cross-sectional nature of the FINRISK studies, is the lack of longitudinal microbiome data. Also, due to being register-based purchase data, compliance with prescribed medications is unknown. Additionally, although statins can be evaluated broadly as one overarching drug group, different statin types may display differing effects. Also, various drugs including ezetimibe, fibrates, ACE-inhibitors, and angiotensin receptor blockers have

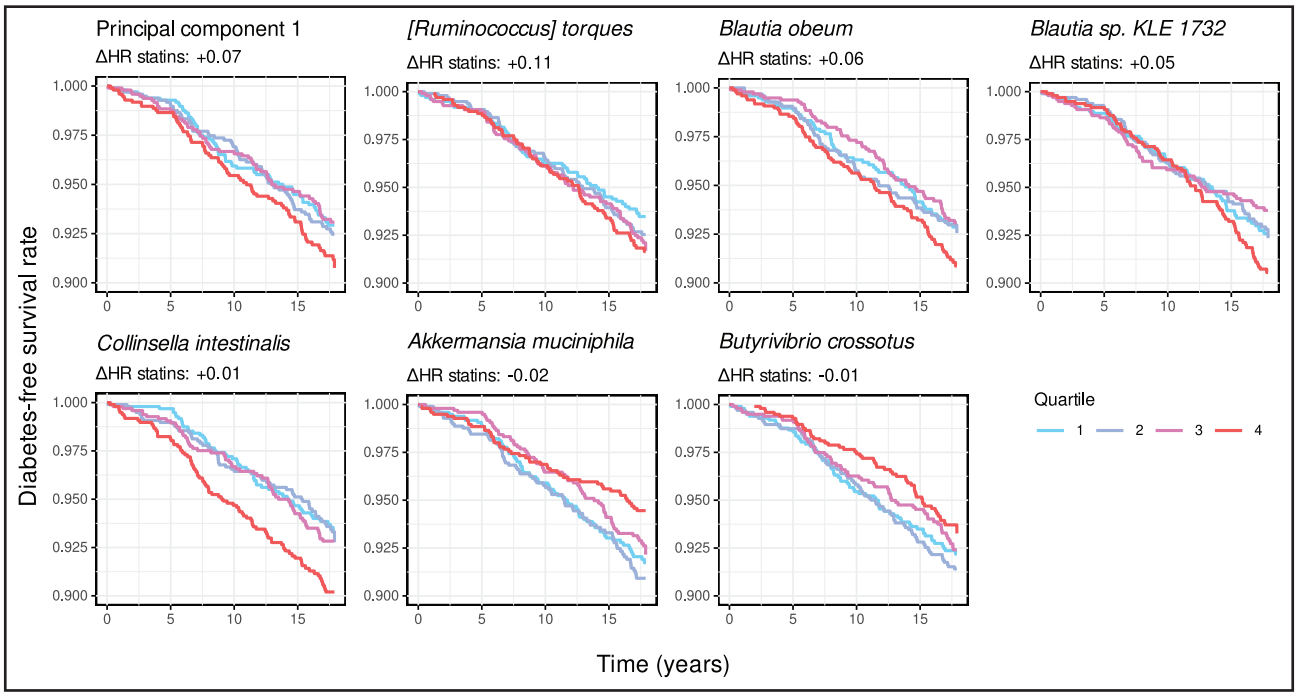

Figure 4. Sample-wide associations of bacterial features that modify statin-associated type-2 diabetes (T2D) risk with diabetes-free survival.

Demographically adjusted (baseline age and sex) Kaplan-Meier survival curves in the combined sample (n=3922) of statistically significant (q<0.05) gut microbiota features for which the direction of the change in statin-associated T2D risk after their inclusion in the model was in harmony with the direction of their individual HR for incident T2D (ie, a positive $\Delta HR_{statins}$ for a feature with an individual HR of >1, etc.). The sample was divided into quartiles in ascending order based on the values of each plot's respective feature (scaled measures of principal component value for principal component 1 (PC1) and centered log-ratio transformed relative abundances for individual species). Q1 (lowest quartile) is colored light blue, Q2 dark blue, Q3 purple, and Q4 red. The *x* axis depicts follow-up time in years and the *y* axis T2D-free survival rate. Beta-diversity components were based on Aitchison distances. Q values: false discovery rate corrected *P* values.

been noted to ameliorate insulin sensitivity, and thus may act in favor of reducing incident T2D when taken in combination with statins.[58] It is therefore necessary to supplement our findings with analyses for each commonly used statin type separately and in combination with the listed drug groups. Finally, since geographic location, culture, ethnicity, and many other environmental factors strongly influence the gut microbiota, care should be taken when generalizing these results to other populations.[59]

In conclusion, our study found novel associations between statin use and the gut microbiota. We identified microbiota features that potentially influence the risk for statin-associated new-onset type-2 diabetes, thus offering tools to study and understand pathogenesis of the disease and help defining a health-promoting gut microbiota profile in the future. Our study makes a meaningful contribution to the collective understanding of how statin medication and the gut microbiota interact with each other and suggests more precise ways to hopefully help refine clinical practice in a way that decreases side effects associated with statin use.

ARTICLE INFORMATION

Received April 18, 2023; accepted October 31, 2023.

Affiliations

Finnish Institute for Health and Welfare (THL), Helsinki, Finland (K.K., O.K., B.J., P.J., T.N., A.S.H., V.S.). Department of Computing, University of Turku, Finland (M.O.R., L.L.). Department of Pediatrics (R.S., C.B., R.K.), Department of Bioengineering (R.S., R.K.), Department of Medicine and Pharmacology (M.J.), Department of Computer Science and Engineering (R.K.), and Center for Microbiome Innovation (R.K.), University of California San Diego, La Jolla. Cambridge Baker Systems Genomics Initiative, Baker Heart and Diabetes Institute, Melbourne, Australia (G.M., M.I.). Department of Infectious Diseases, Central Clinical School, Monash University, Melbourne, Australia (G.M.). Cambridge Baker Systems Genomics Initiative, Department of Public Health and Primary Care, University of Cambridge, United Kingdom (M.I.). Department of Medicine, Turku University Hospital and University of Turku, Finland (T.N.). Institute for Molecular Medicine Finland, FiMM-HiLIFE, Helsinki, Finland (A.S.H.).

Acknowledgments

The authors thank Jon Sanders (Department of Pediatrics, University of California San Diego, La Jolla, CA) for his contribution in preparing the final microbiome data. Additionally, we would like to acknowledge the valuable contribution of the peer-reviewers, whose feedback resulted in a substantially improved article. We also thank the participants and staff of the FINRISK 2002 study.

The authors' responsibilities were as follows: K. Koponen, O. Kambur, T. Niiranen, A.S. Havulinna, R. Knight and V. Salomaa designed the research; R. Salido, C. Brennan, M. Jain, G. Meric, and M. Inouye conducted research; O. Kambur, B. Joseph, A.S. Havulinna, and P. Jousilahti provided essential materials; K. Koponen and O. Kambur performed statistical analysis; K. Koponen had primary responsibility for the final content; K. Koponen, O. Kambur, B. Joseph, M.O. Ruuskanen, L. Lahti, T. Niiranen, A.S. Havulinna, and V. Salomaa wrote the article. All authors read and approved the final article.

Sources of Funding

None of the entities funding this work were involved in its design, implementation, interpretation, writing, or editing. V. Salomaa was supported by the Finnish

Foundation for Cardiovascular Research; and by the Juho Vainio Foundation. T. Niiranen was supported by the Finnish Foundation for Cardiovascular Research; the Emil Aaltonen Foundation; and the Sigrid Juselius Foundation. M.O. Ruuskanen, T. Niiranen, L. Lahti, and A.S. Havulinna were supported by the Research Council of Finland (No. 338818 to MOR; No. 321351 and No. 354447 to T. Niiranen; No. 295741 to L. Lahti; No. 321356 to A.S. Havulinna). M. Inouye was supported by the Munz Chair of Cardiovascular Prediction and Prevention at the Baker Heart and Diabetes Institute. Additional support was provided by Illumina Inc. and Janssen Pharmaceutica through their sponsorship of the Center for Microbiota Innovation at University of California San Diego.

Disclosures

V. Salomaa has ongoing research collaboration with Bayer Ltd (Unrelated to the present study). T. Niiranen has received a speaker's honorarium from Servier. M. Jain holds a position and equity in Sapient Bioanalytics, LLC for work unrelated to the current article. The other authors report no conflicts.

Supplemental Material

Figures S1 and S2
Tables S1–S4
Major resources table

REFERENCES

1. Liu Y, Song X, Zhou H, Zhou X, Xia Y, Dong X, Zhong W, Tang S, Wang L, Wen S, et al. Gut microbiome associates with lipid-lowering effect of rosuvastatin in vivo. *Front Microbiol*. 2018;9:530. doi: 10.3389/fmicb.2018.00530
2. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, Chapman MJ, De Backer GG, Delgado V, Ference BA, et al; ESC Scientific Document Group. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *Eur Heart J*. 2020;41:111–188. doi: 10.1093/eurheartj/ehz455
3. Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, Kirby A, Sourjina T, Peto R, Collins R, et al; Cholesterol Treatment Trialists' (CTT) Collaborators. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *Lancet*. 2005;366:1267–1278. doi: 10.1016/S0140-6736(05)67394-1
4. Yandrapalli S, Malik A, Guber K, Rochlani Y, Pemmasani G, Jasti M, Aronow WS. Statins and the potential for higher diabetes mellitus risk. *Expert Rev Clin Pharmacol*. 2019;12:825–830. doi: 10.1080/17512433.2019.1659133
5. Guber K, Pemmasani G, Malik A, Aronow WS, Yandrapalli S, Frishman WH. Statins and higher diabetes mellitus risk: incidence, proposed mechanisms, and clinical implications. *Cardiol Rev*. 2021;29:314–322. doi: 10.1097/CRD.0000000000000348
6. Wang HJ, Park JY, Kwon O, Choe EY, Kim CH, Hur KY, Lee M-S, Yun M, Cha BS, Kim Y-B, et al. Chronic HMGCR/HMG-CoA reductase inhibitor treatment contributes to dysglycemia by upregulating hepatic gluconeogenesis through autophagy induction. *Autophagy*. 2015;11:2089–2101. doi: 10.1080/15548627.2015.1091139
7. Ling Z, Shu N, Xu P, Wang F, Zhong Z, Sun B, Li F, Zhang M, Zhao K, Tang X, et al. Involvement of pregnane X receptor in the impaired glucose utilization induced by atorvastatin in hepatocytes. *Biochem Pharmacol*. 2016;100:98–111. doi: 10.1016/j.bcp.2015.11.023
8. Nakata M, Nagasaka S, Kusaka I, Matsuoka H, Ishibashi S, Yada T. Effects of statins on the adipocyte maturation and expression of glucose transporter 4 (SLC2A4): implications in glycaemic control. *Diabetologia*. 2006;49:1881–1892. doi: 10.1007/s00125-006-0269-5
9. Zhao S-P, Zhao W. Different effects of statins on induction of diabetes mellitus: an experimental study. *Drug Des Devel Ther*. 2015;9:6211–6223. doi: 10.2147/DDDT.S87979
10. Gustavsson J, Parpal S, Strålfors P. Insulin-stimulated glucose uptake involves the transition of glucose transporters to a caveolae-rich fraction within the plasma membrane: implications for type II diabetes. *Mol Med*. 1996;2:367–372. doi: 10.1007/BF03401634
11. Yada T, Nakata M, Shiraishi T, Kakei M. Inhibition by simvastatin, but not pravastatin, of glucose-induced cytosolic Ca^{2+} signalling and insulin secretion due to blockade of L-type Ca^{2+} channels in rat islet β-cells: simvastatin on β-cell $[Ca2^{2+}]_i$ and insulin release. *Br J Pharmacol*. 1999;126:1205–1213. doi: 10.1038/sj.bjp.0702397
12. Xia F, Xie L, Mihic A, Gao X, Chen Y, Gaisano HY, Tsushima RG. Inhibition of cholesterol biosynthesis impairs insulin secretion and voltage-gated calcium channel function in pancreatic β-cells. *Endocrinology*. 2008;149:5136–5145. doi: 10.1210/en.2008-0161
13. Curry L, Almukhtar H, Alahmed J, Roberts R, Smith PA. Simvastatin inhibits L-type Ca2+-channel activity through impairment of mitochondrial function. *Toxicol Sci*. 2019;169:543–552. doi: 10.1093/toxsci/kfz068
14. Brault M, Ray J, Gomez Y-H, Mantzoros CS, Daskalopoulou SS. Statin treatment and new-onset diabetes: a review of proposed mechanisms. *Metabolism*. 2014;63:735–745. doi: 10.1016/j.metabol.2014.02.014
15. Heintz-Buschart A, Wilmes P. Human gut microbiome: function matters. *Trends Microbiol*. 2018;26:563–574. doi: 10.1016/j.tim.2017.11.002
16. Cresci GA, Bawden E. Gut microbiome: what we do and don't know. *Nutr Clin Pract*. 2015;30:734–746. doi: 10.1177/0884533615609899
17. Fan Y, Pedersen O. Gut microbiota in human metabolic health and disease. *Nat Rev Microbiol*. 2021;19:55–71. doi: 10.1038/s41579-020-0433-9
18. Roessler J, Leistner DM, Landmesser U, Haghikia A. Modulatory role of gut microbiota in cholesterol and glucose metabolism: potential implications for atherosclerotic cardiovascular disease. *Atherosclerosis*. 2022;359:1–12. doi: 10.1016/j.atherosclerosis.2022.08.018
19. Gurung M, Li Z, You H, Rodrigues R, Jump DB, Morgun A, Shulzhenko N. Role of gut microbiota in type 2 diabetes pathophysiology. *EBioMedicine*. 2020;51:102590. doi: 10.1016/j.ebiom.2019.11.051
20. Vieira-Silva S, Falony G, Belda E, Nielsen T, Aron-Wisnewsky J, Chakaroun R, Forslund SK, Assmann K, Valles-Colomer M, Nguyen TTD, et al; MetaCardis Consortium. Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. *Nature*. 2020;581:310–315. doi: 10.1038/s41586-020-2269-x
21. Tsalamandris S, Antonopoulos AS, Oikonomou E, Papamikroulis G-A, Vogiatzi G, Papaioannou S, Deftereos S, Tousoulis D. The role of inflammation in diabetes: current concepts and future perspectives. *Eur Cardiol*. 2019;14:50–59. doi: 10.15420/ecr.2018.33.1
22. Chambers ES, Viardot A, Psichas A, Morrison DJ, Murphy KG, Zac-Varghese SEK, MacDougall K, Preston T, Tedford C, Finlayson GS, et al. Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. *Gut*. 2015;64:1744–1754. doi: 10.1136/gutjnl-2014-307913
23. Vrieze A, Van Nood E, Holleman F, Salojärvi J, Kootte RS, Bartelsman JFWM, Dallinga-Thie GM, Ackermans MT, Serlie MJ, Oozeer R, et al. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology*. 2012;143:913–6.e7. doi: 10.1053/j.gastro.2012.06.031
24. Kootte RS, Levin E, Salojärvi J, Smits LP, Hartstra AV, Udayappan SD, Hermes G, Bouter KE, Koopen AM, Holst JJ, et al. Improvement of insulin sensitivity after lean donor feces in metabolic syndrome is driven by baseline intestinal microbiota composition. *Cell Metab*. 2017;26:611–619.e6. doi: 10.1016/j.cmet.2017.09.008
25. Kim J, Lee H, An J, Song Y, Lee C-K, Kim K, Kong H. Alterations in gut microbiota by statin therapy and possible intermediate effects on hyperglycemia and hyperlipidemia. *Front Microbiol*. 2019;10:1947. doi: 10.3389/fmicb.2019.01947
26. Henriksbo BD, Lau TC, Cavallari JF, Denou E, Chi W, Lally JS, Crane JD, Duggan BM, Foley KP, Fullerton MD, et al. Fluvastatin causes NLRP3 inflammasome-mediated adipose insulin resistance. *Diabetes*. 2014;63:3742–3747. doi: 10.2337/db13-1398
27. World Medical Association. World Medical Association declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310:2191. doi: 10.1001/jama.2013.281053
28. Borodulin K, Tolonen H, Jousilahti P, Jula A, Juolevi A, Koskinen S, Kuulasmaa K, Laatikainen T, Männistö S, Peltonen M, et al. Cohort profile: the national FINRISK study. *Int J Epidemiol*. 2018;47:696–696i. doi: 10.1093/ije/dyx239
29. Pajunen P, Koukkunen H, Ketonen M, Jerkkola T, Immonen-Räihä P, Kärjä-Koskenkari P, Mähönen M, Niemelä M, Kuulasmaa K, Palomäki P, et al. The validity of the Finnish Hospital Discharge Register and Causes of Death Register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil*. 2005;12:132–137. doi: 10.1097/00149831-200504000-00007
30. Sund R. Quality of the Finnish Hospital Discharge Register: a systematic review. *Scand J Public Health*. 2012;40:505–515. doi: 10.1177/1403494812456637
31. WHO Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index 2022. 2022. Accessed November 20, 2023. https://www.whocc.no/atc_ddd_index/
32. Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, Mujagic Z, Vila AV, Falony G, Vieira-Silva S, et al; LifeLines Cohort Study. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science*. 2016;352:565–569. doi: 10.1126/science.aad3369
33. Park J, Kato K, Murakami H, Hosomi K, Tanisawa K, Nakagata T, Ohno H, Konishi K, Kawashima H, Chen Y-A, et al. Comprehensive analysis of gut

microbiota of a healthy population and covariates affecting microbial variation in two large Japanese cohorts. *BMC Microbiol*. 2021;21:151. doi: 10.1186/s12866-021-02215-0
34. Glenn TC, Nilsen RA, Kieran TJ, Sanders JG, Bayona-Vásquez NJ, Finger JW, Pierson TW, Bentley KE, Hoffberg SL, Louha S, et al. Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). *PeerJ*. 2019;7:e7755. doi: 10.7717/peerj.7755
35. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–359. doi: 10.1038/nmeth.1923
36. Hillmann B, Al-Ghalith GA, Shields-Cutler RR, Zhu Q, Knight R, Knights D. SHOGUN: a modular, accurate and scalable framework for microbiome quantification. *Bioinformatics*. 2020;36:4088–4090. doi: 10.1093/bioinformatics/btaa277
37. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J*. 1948;27:379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
38. Aitchison J, Barceló-Vidal C, Martín-Fernández JA, Pawlowsky-Glahn V. Logratio analysis and compositional distance. *Math Geol*. 2000;32:271–275. doi: 10.1023/A:1007529726302
39. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
40. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 2013;8:e61217. doi: 10.1371/journal.pone.0061217
41. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, et al. Vegan: community ecology package. R package version 2.3.2. 2022. Accessed November 20, 2023. https://github.com/vegandevs/vegan
42. Lahti L, Shetty S, et al. Tools for microbiome analysis in R. 2022. Accessed November 20, 2023. https://microbiome.github.io/tutorials/
43. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc*. 1995;57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
44. Clarke KR. Non-parametric multivariate analyses of changes in community structure. *Austral Ecol*. 1993;18:117–143. doi: 10.1111/j.1442-9993.1993.tb00438.x
45. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol*. 2001;26:32–46. doi: 10.1111/j.1442-9993.2001.tb00081.x
46. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, Tickle TL, Weingart G, Ren B, Schwager EH, et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol*. 2021;17:e1009442. doi: 10.1371/journal.pcbi.1009442
47. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco California USA: ACM; 2016:785–794. Accessed November 20, 2023. https://dl.acm.org/doi/10.1145/2939672.2939785
48. Wilmanski T, Kornilov SA, Diener C, Conomos MP, Lovejoy JC, Sebastiani P, Orwoll ES, Hood L, Price ND, Rappaport N, et al. Heterogeneity in statin responses explained by variation in the human gut microbiome. *Med*. 2022;3:388–405.e6. doi: 10.1016/j.medj.2022.04.007
49. Khan TJ, Ahmed YM, Zamzami MA, Mohamed SA, Khan I, Baothman OAS, Mehanna MG, Yasir M. Effect of atorvastatin on the gut microbiota of high fat diet-induced hypercholesterolemic rats. *Sci Rep*. 2018;8:662. doi: 10.1038/s41598-017-19013-2
50. Fu J, Bonder MJ, Cenit MC, Tigchelaar EF, Maatman A, Dekens JAM, Brandsma E, Marczynska J, Imhann F, Weersma RK, et al. The gut microbiome contributes to a substantial proportion of the variation in blood lipids. *Circ Res*. 2015;117:817–824. doi: 10.1161/CIRCRESAHA.115.306807
51. Robert C, Chassard C, Lawson PA, Bernalier-Donadille A. Bacteroides cellulosilyticus sp nov, a cellulolytic bacterium from the human gut microbial community. *Int J Syst Evol Microbiol*. 2007;57:1516–1520. doi: 10.1099/ijs.0.64998-0
52. Sonnenburg JL, Xu J, Leip DD, Chen C-H, Westover BP, Weatherford J, Buhler JD, Gordon JI. Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science*. 2005;307:1955–1959. doi: 10.1126/science.1109051
53. Wexler HM. Bacteroides: the good, the bad, and the nitty-gritty. *Clin Microbiol Rev*. 2007;20:593–621. doi: 10.1128/CMR.00008-07
54. Bo T, Wen J, Zhao Y, Tian S, Zhang X, Wang D. Bifidobacterium pseudolongum reduces triglycerides by modulating gut microbiota in mice fed high-fat food. *J Steroid Biochem Mol Biol*. 2020;198:105602. doi: 10.1016/j.jsbmb.2020.105602
55. Ruuskanen MO, Erawijantari PP, Havulinna AS, Liu Y, Méric G, Tuomilehto J, Inouye M, Jousilahti P, Salomaa V, Jain M, et al. Gut microbiome composition is predictive of incident type 2 diabetes in a population cohort of 5,572 Finnish adults. *Diabetes Care*. 2022;45:811–818. doi: 10.2337/dc21-2358
56. Togo AH, Diop A, Bittar F, Maraninchi M, Valero R, Armstrong N, Dubourg G, Labas N, Richez M, Delerce J, et al. Description of Mediterraneibacter massiliensis, gen nov, sp nov, a new genus isolated from the gut microbiota of an obese patient and reclassification of Ruminococcus faecis, Ruminococcus lactaris, Ruminococcus torques, Ruminococcus gnavus and Clostridium glycyrrhizinilyticum as Mediterraneibacter faecis comb nov, Mediterraneibacter lactaris comb nov, Mediterraneibacter torques comb nov, Mediterraneibacter gnavus comb nov and Mediterraneibacter glycyrrhizinilyticus comb nov. *Antonie Van Leeuwenhoek*. 2018;111:2107–2128. doi: 10.1007/s10482-018-1104-y
57. Hillmann B, Al-Ghalith GA, Shields-Cutler RR, Zhu Q, Gohl DM, Beckman KB, Knight R, Knights D. Evaluating the information content of shallow shotgun metagenomics. *mSystems*. 2018;3:e00069–e00018. doi: 10.1128/mSystems.00069-18
58. Lim S, Oh PC, Sakuma I, Koh KK. How to balance cardiorenometabolic benefits and risks of statins. *Atherosclerosis*. 2014;235:644–648. doi: 10.1016/j.atherosclerosis.2014.06.001
59. Gupta VK, Paul S, Dutta C. Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. *Front Microbiol*. 2017;8:1162. doi: 10.3389/fmicb.2017.01162