Streptococcus species abundance in the gut is linked to subclinical coronary atherosclerosis in 8,973 participants from the SCAPIS cohort

Sergi Sayols-Baixeras, PhD[1,2*], Koen F. Dekkers, MSc[1*], Gabriel Baldanzi, MD[1], Daniel Jönsson, DDS, PhD[3,4,5], Ulf Hammar, BSc[1], Yi-Ting Lin, MD, PhD[1,6], Shafqat Ahmad, PhD[1,7], Diem Nguyen, PhD[1], Georgios Varotsis, DVM[1], Sara Pita, MSc[8,9], Nynne Nielsen, MSc[8], Aron C. Eklund, PhD[8], Jacob B. Holm, PhD[8], H. Bjørn Nielsen, PhD[8], Ulrika Ericson, PhD[3], Louise Brunkwall, PhD[3,10], Filip Ottosson, PhD[3,11], Anna Larsson, MD[3], Dan Ericson, DDS, PhD[12], Björn Klinge, DDS, PhD[5,13], Peter M. Nilsson, MD, PhD[3,14], Andrei Malinovschi, MD, PhD[15], Lars Lind, MD, PhD[16], Göran Bergström, MD, PhD[17,18], Johan Sundström, MD, PhD[16,19], Johan Ärnlöv, MD, PhD[6,20], Gunnar Engström, MD, PhD[3], J. Gustav Smith, MD, PhD[21,22,23], Marju Orho-Melander, PhD[3†] and Tove Fall, DVM, PhD[1†]

*Shared first authorship
†Shared senior authorship

[1]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

[2]CIBER Cardiovascular diseases (CIBERCV), Instituto de Salud Carlos III, Madrid, Spain

[3]Department of Clinical Sciences in Malmö, Lund University, Malmö, Sweden

[4]Public Dental Service of Skåne, Lund, Sweden

[5]Department of Periodontology, Faculty of Odontology, Malmö University, Malmö, Sweden

[6]Division of Family Medicine and Primary Care, Department of Neurobiology, Care Science and Society, Karolinska Institutet, Huddinge, Sweden

[7]Preventive Medicine Division, Harvard Medical School, Brigham and Women's Hospital, Boston, MA, United States of America

[8]Clinical Microbiomics A/S, Copenhagen, Denmark

[9]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark

[10]Clinical Studies Sweden, Forum Söder, Region Skåne, Lund, Sweden

[11]Section for Clinical Mass Spectrometry, Danish Center for Neonatal Screening, Department of Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark

[12]Department of Cariology, Faculty of Odontology, Malmö University, Malmö, Sweden.

[13]Department of Dental Medicine, Karolinska Institutet, Solna, Sweden.

[14]Department of Internal Medicine, Skånes University Hospital, Malmö, Sweden

[15]Department of Medical Sciences, Clinical Physiology, Uppsala University, Uppsala, Sweden.

[16]Department of Medical Sciences, Clinical Epidemiology, Uppsala University, Uppsala, Sweden

[17]Department of Molecular and Clinical Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[18]Department of Clinical Physiology, Sahlgrenska University Hospital, Region Västra Götaland, Gothenburg, Sweden

[19]The George Institute for Global Health, University of New South Wales, Sydney, Australia

[20]School of Health and Social Studies, Dalarna University, Falun, Sweden

[21]The Wallenberg Laboratory/Department of Molecular and Clinical Medicine, Institute of Medicine, Gothenburg University and the Department of Cardiology, Sahlgrenska University Hospital, Gothenburg, Sweden

[22]Department of Cardiology, Clinical Sciences, Lund University and Skåne University Hospital, Lund, Sweden

[23]Wallenberg Center for Molecular Medicine and Lund University Diabetes Center, Lund University, Lund, Sweden

Short title: Gut Streptococci are linked to atherosclerosis

Correspondence to: Tove Fall, PhD, Professor in Molecular Epidemiology, Dept.

Department of Medical Sciences, Molecular Epidemiology

Uppsala University, EpiHubben, MTC-huset,

75185, Uppsala, Sweden.

E-mail: tove.fall@medsci.uu.se

Word Count

Abstract: 313

Main text: 4982

ABSTRACT

BACKGROUND: The gut microbiota are implicated in atherosclerotic disease, but the relation to subclinical coronary atherosclerosis and to the oral microbiome is unclear. The aim of the study was to identify associations between gut microbial species and computed tomography-based measures of asymptomatic coronary atherosclerosis in a large population-based cohort and to explore relevant clinical correlates.

METHODS: We conducted a cross-sectional study of 8973 participants 50 to 65 years old without known cardiovascular disease from the Swedish Cardiopulmonary BioImage Study (SCAPIS). Coronary atherosclerosis was assessed by coronary artery calcium score (CACS) and coronary computed tomography angiography (CCTA). Gut microbiota were identified by deep shotgun metagenomics sequencing of stool samples. Multiple regression models adjusted for cardiovascular risk factors were applied, and findings were validated in a case-control study of patients (n=373) with symptomatic atherosclerotic cardiovascular disease. CACS-associated species were assessed for association with carotid plaques, corresponding species in saliva samples, inflammatory markers and infection, and enrichment of metabolic pathways.

RESULTS: Gut bacterial associations with CACS were enriched for *Streptococcus* and *Oscillobacter* genera and for bacteria carrying genes linked to amino acid and carbohydrate degradation functions. In the main model, 67 gut microbial species were associated with CACS; associations were strongest for *Streptococcus anginosus* and *Streptococcus oralis* subsp. *oralis*. These two and 13 other associations were replicated in the validation study. Associations were largely similar across CCTA-based measurements. Five CACS-associated species correlated with the same species colonizing the mouth in the Malmö Offspring Dental study (MODS), which were in turn associated with worse dental health. Nineteen CACS-associated species were associated with high-sensitivity C-reactive protein concentrations in the same direction as with CACS, and 16 were associated with neutrophil counts.

CONCLUSIONS: This study provides robust evidence of an association of a higher abundance in the gut microbiome of several *Streptococcus* spp. and other species commonly found in the mouth with subclinical atherosclerosis and biomarkers of systemic inflammation and infection. Our findings call for a re-evaluation of the hypothesis of an infectious component in atherogenesis with further longitudinal and experimental studies.

Key Words: Atherosclerosis, coronary artery calcium score, tomography, metagenomics, gut microbiota, *Streptococcus.*

CLINICAL PERSPECTIVE

WHAT IS NEW?

- Shotgun metagenomics identified robust associations between gut species and subclinical atherosclerosis assessed with computed tomography-derived coronary artery calcium score (CACS) in 8973 participants, showing an overrepresentation of the *Streptococcus and Oscillobacter* genera.
- We observed an enrichment for bacterial genes involved in amino acid and carbohydrate degradation functions.
- Gut *Streptococcus* spp. were robustly associated with circulating biomarkers of systemic inflammation and infection response, and with the same species located in the mouth, which were in turn associated with oral pathologies.

WHAT ARE THE CLINICAL IMPLICATIONS?

- We describe the link between gut metagenomics species, especially species commonly found in the mouth, with subclinical coronary atherosclerosis and biomarkers of inflammation in the largest cardiovascular and metagenomics study to date.
- The effects of gut and oral *Streptococcus* spp. on risk for coronary artery disease merit further longitudinal and experimental studies.

Non-standard Abbreviation and Acronyms

CACS	Coronary artery calcium score
CCTA	Coronary computed tomography angiography
CVD	Cardiovascular disease
GMM	Gut metabolic module
GSEA	Gene-set enrichment analysis
hsCRP	High-sensitivity C-reactive protein
ln	Natural logarithm
MODS	Malmö Offspring Dental Study
MOS	Malmö Offspring Study
SCAPIS	Swedish CArdioPulmonary bioImage Study
SIS	Segment involvement score
TSEA	Taxon-set enrichment analysis
VGS	Viridans streptococci group

Atherosclerotic cardiovascular disease (CVD) is a major cause of death and disability.[1] Atherosclerotic plaque formation is characterized by progressive accumulation of lipids, fibrous elements, calcium minerals, and inflammatory molecules in the subendothelial space. The underlying mechanisms are incompletely understood; however, it has been proposed that the gut microbiota composition could accelerate atherosclerotic development through transmission of bacteria into the circulation, resulting in direct infection of plaques or systemic inflammation associated with infection at other sites.[2] Because the gut and oral microbiomes are connected, and dental health is associated with endothelial dysfunction and atherosclerotic disease,[3-5] the gut microbiome could also act as a sentinel for oral pathogenic bacteria. Experimental studies suggest that oral challenge with *Streptococcus* spp. accelerates plaque growth and macrophage invasion,[6] and an overlap between bacteria found in oral cavity, fecal samples and carotid atherosclerotic plaque bacteria in patients has been shown.[7] Certain gut bacteria could affect atherogenesis by modulating host metabolism or by interacting with dietary components to produce beneficial or harmful molecules.[2] The composition of the gut microbiota has been linked to cardiovascular risk factors, such as obesity, insulin resistance, type 2 diabetes, although the causal relations are unclear.[8] Studies of symptomatic coronary atherosclerosis identified differences in the abundance of more than 500 gut species in cases and controls.[9-11] However, symptomatic cases and controls often differ in medical treatment and lifestyle factors, which makes findings prone to bias. Thus, the earlier phases of the atherogenesis merit study in a larger population-based sample.

To avoid the limitations of previous studies and gain insight into the relation of the gut microbiota and atherosclerosis, we sought to identify associations between gut microbiota composition and computed tomography-based measures of asymptomatic coronary atherosclerosis in a large cohort of middle-aged Swedes from the Swedish CArdioPulmonary bioImage Study (SCAPIS). The findings were validated in a case-control study of symptomatic

atherosclerotic disease in China. We further assessed associations between atherosclerosis-associated gut bacterial species and biomarkers of inflammation and infection, plasma metabolites, and the abundance of corresponding bacterial species in the oral cavity.

METHODS

Study design and participants

The study design is illustrated in **Figure 1**. The primary data source was SCAPIS, a population-based study focusing on cardiovascular and respiratory disease in 30 154 subjects 50–64 years of age from six sites in Sweden.[12] We studied participants with high-quality fecal metagenomics data from the study sites in Uppsala (n=4541) and Malmö (n=4432), after excluding subjects with missing information on country of birth (n=39), coronary artery calcium score (CACS) (n=356) or who had self-reported CVD (n=451). All participants gave written informed consent. This study was conducted in accordance with the Declaration of Helsinki and was approved by the regional ethics committee (Umeå: Dnr 2010-228-31M; Uppsala: Dnr 2018/315).

To validate the findings, we analyzed publicly available data from a case-control study[10] of 210 patients with atherosclerotic cardiovascular disease, most with stable angina (mean age ± SD 61.0 ± 9.72 years), and 163 controls (60.4 ± 10.0 years); the subjects were of Han Chinese origin and were recruited at the Medical Research Center of Guangdong General Hospital, after excluding 12 participants with missing age or sex information. Disease status, age, and sex were assessed with the *curatedMetagenomicData* R package.[13]

The association between gut species and their counterparts in saliva samples and with oral health was investigated in participants of the Malmö Offspring Dental Study (MODS, n=831, mean age 52.9 years), a substudy of the family-based Malmö Offspring Study (MOS).[4, 14] Salivary metagenomics, oral examination data, and complete data on age and sex were available for 650 subjects; informative fecal metagenomics data from MOS were available for 435 of

those subjects. Both studies were approved by the regional ethics committee in Lund (MODS: Dnr 2013/761) (MOS: Dnr 2012/594).

Atherosclerosis Measurements

Cardiac imaging in SCAPIS included coronary computed tomography angiography (CCTA) as described previously.[15] Briefly, computed tomography was done with a dedicated dual-source scanner and a Stellar Detector (Somatom Definition Flash, Siemens Medical Solutions). Noncontrast images for calcium scoring were obtained from all subjects, followed by contrast-enhanced images from subjects without contraindications. CACS was determined using the syngo.via calcium scoring software (Volume Wizard; Siemens). The area of calcification was summed for the whole coronary artery tree to a CAC score according to Agatston.[16] Radiologists inspected the images for any coronary atherosclerosis, segment involvement score, Duke prognostic coronary artery disease index modified for SCAPIS,[15] and occlusive atherosclerosis (≥50% stenosis in at least one vessel) (see **Supplemental Methods)**. We included individuals with valid readings in at least four proximal segments. Carotid atherosclerosis were categorized as absent, unilateral, or bilateral[12] from two-dimensional gray-scale ultrasound images obtained with a Siemens Acuson S2000 ultrasound scanner and a 9L4 linear transducer (both from Siemens, Germany) using a standardized protocol and the Mannheim consensus.[17]

Metagenomics

DNA was extracted from fecal samples (SCAPIS and MOS) and saliva samples (MODS) at Clinical Microbiomics (Copenhagen, Denmark). The analytical pipeline is described in detail in **Supplemental Methods**. In brief, libraries of fragmented DNA were sequenced with the Illumina Novaseq 6000 platform. Nonhost reads and data from other cohorts were used to build separate nonredundant gene catalogs for fecal and saliva samples. Metagenomic species and corresponding signature gene were identified.[18] In all three datasets and in the validation study,

species abundance was estimated by mapping reads to the signature gene sets and normalization for effective gene length. For the validation study, FASTQ files were downloaded from the European Nucleotide Archive (Project PRJEB21528).[10] The taxonomical information for the two catalogues was mapped with the NCBI RefSeq database.[19]

Species abundance was transformed with the formula ln(x+1), where x denotes the relative abundance of each species, followed by z-transformation to set the mean to 0 and SD to 1. These transformations were applied to species in all four analytical datasets before statistical analysis.

Alpha (Shannon diversity index, inverse Simpson index, and Chao1) and beta diversity (Bray-Curtis dissimilarity) were estimated with the R v4.1.1 package *vegan* v2.5-7 using rarefied data.[20] Rarefied data were also used to define individuals with above/less than median of each species. A functional potential profile was determined for each species assigning genes to gut metabolic modules (GMM),[21] which includes 103 metabolic pathways that represent cellular enzymatic processes.

Clinical Chemistry, Hematology and Metabolomics

Venous samples were analyzed for high-density and low-density lipoprotein cholesterol, triglycerides, and high-sensitivity C-reactive protein (hsCRP). Blood cells were counted by standard methods. Enriched metabolite groups in species-metabolite associations in SCAPIS were accessed from the GUTSY atlas (https://gutsyatlas.serve.scilifelab.se/) from Dekkers et al.[22] Only the metabolite subclasses with ≥5 metabolites and species with >100 carriers were included.

Oral Health Phenotypes

Dental examinations of MODS participants were done by five trained dentists. Both surfaces with caries and fillings were recorded on all teeth, counting four to five surfaces per tooth. Surfaces with caries were detected using standard clinical criteria aided by mirror, probe (Hu-

Friedy EXD57), and bite-wing radiographs combining initial and manifest lesions. Filled surfaces included both fillings and crowns. Gingival inflammation was recorded as percentage of bleeding on probing with a Hu-Friedy PCPUNC157 probe, excluding wisdom teeth, and counting six surfaces per tooth.

Other Phenotypes

In SCAPIS and MOS, data on country of birth, smoking, physical activity, and diet were collected with validated standardized questionnaires. Blood pressure and body mass index were recorded. Data on medications were acquired from the drug prescription register, self-report, plasma measurements, and other sources. Relatedness in SCAPIS from kinship analysis of genotype data, assigning first-degree relatives individuals to the same family and in MOS on family information in the study design phase. Sociodemographic information in the validation study was collected with questionnaires.[10] Collected phenotypes are described in detail in **Supplemental Methods**.

Statistical Analysis

Association between Gut Microbiota Diversity and Atherosclerosis The association of alpha diversity with coronary atherosclerosis was assessed by fitting linear models with ln of (CACS+1) as the dependent and alpha diversity indices as the independent variable. Two sets of covariates were selected as fixed effects in the model. The basic model was adjusted for age, sex, country of birth, study center and metagenomics extraction plate. The main model was further adjusted for smoking; physical activity; energy-adjusted carbohydrate, protein, and fiber intake; systolic and diastolic blood pressure; total cholesterol; high- and low-density lipoprotein cholesterol; ln(triglycerides); diabetes; body mass index; and self-reported medication for dyslipidemia, hypertension, and diabetes. Correlation of covariates and alpha diversity measurements are reported in **Figure S1**. Beta diversity was assessed by a distance-based multivariate analysis of variance of ln of (CACS+1) as the independent variable using the

dmanova function in the GUniFrac v1.3 package. These analyses were adjusted for the same two set of covariates as in the alpha diversity analyses.

Association between Gut Microbiota Species and CACS. A series of linear multivariable regressions done with the ln of (CACS+1) as the dependent variable and the abundance of each species as the independent variable. Multiple testing was accounted for using the Benjamini-Hochberg False Discovery Rate (FDR) at 5%.[23] Extensive simulation studies before the analysis supported the use of linear regression to maximize power and interpretability and to minimize false-positive findings (**Supplemental Methods, Figure S2**). The models were fitted for each species separately; covariates from the basic model were used as fixed effects. A Taxon Set Enrichment Analysis (TSEA) based on the *fgsea* v1.18.0 R package, was used[24] to determine whether certain genera were enriched for associations, with multiple testing accounted for after combining the results for positive and negative regression coefficients. To determine whether these results were driven by certain genera, TSEA was applied repeatedly, leaving out in scheme, each time, all species of a genus from the analysis. A Gene Set Enrichment Analysis (GSEA), also based on the *fgsea* v1.18.0 R package[24], was used to assess whether associations were enriched in species carrying genes for specific metabolic models (GMMs). In both analyses, *P* values were ranked separately for positive and negative regression coefficients.

Species with a q-value <0.05 in the basic model were tested adjusting for the main model covariates, again accounting for multiple testing. To determine whether linear associations were caused by influential observations, unscaled dfbeta values were calculated using the *influence.measures* R function. The linear regression estimates were deemed unreliable if an influential observation had higher dfbeta values than the regression coefficient and the $P<0.05$ after removing the influential observation. Findings not caused by few influential observations are referred to as "CACS-associated species". Effect estimate modification by sex was assessed

for CACS-associated species by entering an interaction term between sex and each of the model variables and extracting the interaction coefficient of sex with species.

Validation. In three separate analyses, we excluded three groups of participants: those in whom proton-pump inhibitors were measured in plasma, those with inflammatory bowel disease, and those treated with antibiotics during the year before the baseline visit. Analyses for CACS-associated species were repeated with the Shannon diversity index as a covariate to test if findings were driven by global changes in the gut microbiota composition rather than change in individual species. Lastly, the associations from the main model were tested in linear mixed models accounting for first degree-relatedness among 437 participants from 222 families using the *lmerTest* v3.1-3 R package.

In the external validation case-control study, Firth-type logistic regressions adjusted for age and sex were fitted with atherosclerotic cardiovascular disease status as the binary outcome and available CACS-associated species as the independent variable using the *logistf* v1.24 R package. Significant associations ($P<0.05$) with consistent effect estimate directions were considered validated.

Association between CACS-Associated Species and CCTA-Based Measurements and Carotid Plaques. We assessed the association of CACS-associated species with presence of the CCTA-based measures of any coronary atherosclerotic plaque or coronary stenosis >50% in at least one vessel in a series of logistic regressions using the main model covariates. The association between CACS-associated species and segment involvement score, modified Duke index, and number of carotid vessels with atherosclerosis plaques were assessed with ordinal regression models adjusted using the main model covariates using the *MASS* v7.3-54 R package. Categories 4, 5 and 6 of the modified Duke index were merged.

Associations between CACS-Associated Species and Plasma Metabolites

The enrichment results for 63 metabolite subclasses ($P<0.05$) in the positive associations of single species with plasma metabolites were downloaded. To determine whether CACS-associated and non-associated species differed in the frequency of enrichments for certain metabolite subclasses, a series of exact logistic regressions was applied, one for each metabolite subclass, including all species; "enriched for" was dependent variable and CACS-associated (yes/no) was independent variable. Species positively and negatively associated with CACS were analyzed separately, and a FDR of 0.05 was applied.

Associations between CACS-Associated Species and Markers of Systemic Inflammation and Infection. A series of linear regressions was fitted to assess the association between CACS-associated species (independent variable) and ln-transformed hsCRP, and ln-transformed counts of neutrophils and leukocytes, separately, as the dependent variable. The models were adjusted for the same covariates as the main model.

Associations between Gut and Oral CACS-Associated Species and between Oral CACS-Associated Species and Oral Health Phenotypes. To investigate the association between CACS-associated species in the gut and the corresponding species in saliva, MOS/MODS data was analyzed with a of partial Spearman correlations from a mixed model based on rank-transformed data (details in Supplemental Methods). The models were adjusted for age, sex, and metagenomics extraction plate as fixed effects and family relatedness as a random effect. The oral species associated (q-value <0.05) with the corresponding CACS-associated gut species were assessed for association with three oral health phenotypes: filled surfaces (split by deciles in 10 categories), caries, and gingival inflammation. Filled surfaces and caries were assessed using ordinal mixed regressions using the v2019.12-10 R package, while gingival inflammation were assessed with a linear mixed model. These regression models were adjusted for age, sex, smoking, education, Silness-Löe plaque index, as a marker of oral hygiene, and

the activity in the hour immediately before the dental examination (e.g., eating, brushing teeth, smoking) as fixed effects and family relatedness as a random effect.

Data Availability

Deidentified metagenomic sequencing data can be accessed in the ENA under accession number PRJEB51353 (https://www.ebi.ac.uk/ena/browser/view/prjeb51353). Access to the remaining data requires ethical approval from the Swedish Ethical Review Board and approval from the SCAPIS Data access board (https://www.scapis.org/data-access/).

Code Availability

The source code and the summary data underlying all figures used to generate the results for the analysis are available at https://github.com/MolEpicUU/GUTSY_CACS.

RESULTS

Gut Microbiota Composition and Richness Are Associated with Asymptomatic Atherosclerosis and Attenuated by Adjustment for Lifestyle Factors, Diet and Medication

The study included 8973 subjects with no self-reported history of symptomatic CVD (**Table 1, Table S1**). The Shannon diversity index—a measure of overall species richness and evenness within each sample—was inversely associated with CACS (n=8973, β=-0.44, P=4.5×10^{-4}) in the basic model adjusting only for age, sex, country of birth and technical variables but not in the main model with extensive adjustment for cardiovascular risk factors (n=8118, β=–0.17, P=0.21). Alternative alpha diversity indices were similarly associated with CACS (**Figure S3a**). Beta diversity—a pair-wise indicator of dissimilarity of samples—differed across the CACS categories in both models, although the main model was attenuated (r^2_{basic}=0.0005; P_{basic} = 1×10^{-27} vs r^2_{main}=0.0002; P_{main}=0.036) (**Figure S3b**). These findings encouraged us to investigate species-specific associations.

Species Associated with Subclinical Atherosclerosis Are Enriched in Species from *Streptococcus* and *Oscillibacter* Genera

The taxonomical profiling of fecal samples identified 1985 species with an average of 325 species per sample (range: 11–696); *Firmicutes* (72%) and *Bacteroidetes* (20%) were the dominant phyla (**Figure S4**). In the basic model including 8973 individuals, the relative abundance of 78 species associated positively with CACS and that of 30 species was associated negatively. (**Figure 2, Table S2**). TSEA revealed an overrepresentation of the *Streptococcus* and *Oscillibacter* genera in the positive CACS associations (**Table S3**). Seven GMMs were found enriched by GSEA in bacterial species positively associated with CACS (**Table S4**); two were involved in amino acid degradation (threonine degradation I and II), two in carbohydrate degradation (trehalose and fructose degradation), and one each in *Bifidobacterium* shunt, glycerol degradation, and propionate production. One GMM was enriched in species negatively

associated with CACS (homoacetogenesis). We tested whether these functions were due to the overrepresentation of specific genera (Figure S5). This analysis revealed that the trehalose degradation enrichment was strongly attenuated when we removed *Streptococcus* spp. For propionate production, *Oscillibacter* was the most important genus, and for homoacetogenesis, *Romboutsia.* Thus, species associated with CACS share, at least partially, some specific bacterial functions related to amino acid and carbohydrate degradation; moreover, *Streptococcus* spp. were the most important contributors to the enrichment for trehalose degradation genes.

Sixty-seven Species Were Associated with CACS Independent of Cardiovascular Risk Factors

In the main model, which included 8118 individuals with complete data on covariates including age, sex, country of birth, metagenomics extraction plate, smoking, physical activity, diet, blood pressure, lipids, diabetes, body mass index, and self-reported cardiovascular medication, 68 species remained associated with CACS (54 positively and 14 negatively) (Figure 2). The relative abundance of these 68 species across CACS groups is shown in Figure 3. However, the association of *Muribaculaceae* sp. (MGS ID: HG3A.1967) was deemed nonrobust, as the removal of an influential observation abolished the association (Table S5). The relative abundance of the remaining 67 CACS-associated species across CACS groups is shown in Figure S6. *Streptococcus anginosus*, *S. oralis* subsp. *oralis*, *Escherichia coli*, *Eubacteriales* sp. (HG3A.1354) and *Intestinimonas* sp. (HG3A.1149) were positively associated with CACS with the lowest *P* values. Comparison of the main clinical characteristics in participants with high and low abundance of CACS-associated species (cut-off median) showed that those with higher abundance of the two most strongly associated species, *S. anginosus,* and *S. oralis* subsp. *oralis* had in general more cardiovascular risk factors (Table S6). Furthermore, there were indications that the associations between *S. agalactiae*, *Rothia mucilaginosa*, two *Eubacteriales* spp.

(HG3A.0511and HG3A.0854), and an *Oscillibacter* sp. (HG3A.0243) and CACS were modified by sex (interaction $P<0.05$, Table S5).

Validation

Sensitivity analyses — exclusion of individuals who used proton pump inhibitors, had inflammatory bowel disease, or were treated with antibiotics in the previous 12 months from the analyses, as well as adjusting for Shannon diversity index, or accounting for relatedness — showed largely similar effect estimates (Figure S7). Overall, these findings indicate that the associations identified in the main analysis were robust. We attempted to replicate the findings by re-analyzing public shotgun metagenomics data from a case-control study[10] (Table S7) of subjects from a different geographical area and different microbiota composition (Figure S4). Of the 67 species associated with CACS in the main model, we identified 48 species using the same gene sequence signatures as in the SCAPIS study; and of these, 15 including five *Streptococcus* spp., were associated with atherosclerotic cardiovascular disease with effect estimate in the same direction as in SCAPIS (Figure 3).

Other Atherosclerosis Measurements

Next, we assessed the association between CACS-associated species and four CCTA measurements (presence of coronary atherosclerosis plaque, coronary stenosis >50%, modified Duke index score, and segment involvement score) and the number of carotid vessels with atherosclerotic plaque. Twenty-seven CACS-associated species were also associated ($P<0.05$) with any coronary atherosclerotic plaque, 6 with coronary stenosis >50%, 38 with the modified Duke index score, 42 with the segment involvement score, and 5 with carotid plaque, all in the same direction as with CACS (Figure 3). The correlation among atherosclerosis measurements is shown in Figure S8.

Metabolomics

Compared to the 1464 non-CACS-associated species, the 54 species positively associated with CACS in the main model were enriched in positive associations with several metabolite subclasses, including food/plant components (e.g., 2-piperidinone), aminosugar metabolism (e.g., N-acetylneuraminic acid), sterols (e.g., 7-HOCA), tobacco metabolites, xanthine metabolism and cardiovascular drugs (e.g., enalpril, metoprolol). As an example, 14% of CACS-associated species had metabolite associations enriched for aminosugar metabolites, while only 0.6% of non-CACS-associated species had these enrichments (q-value 8×10^{-6}). The 13 species negatively associated with CACS were enriched in positive associations with metabolites from the vitamin A metabolism and sphingomyelin groups (**Table S8**).

Oral Species in the Gut Flora are Associated with Markers of Systemic Inflammation and Infection

We assessed associations between CACS-associated species and hsCRP (n=8117) and counts of leukocytes (n=8098) and neutrophils (n=8092) in multivariable models. Of 54 species positively associated with CACS ($P<0.05$), 13 were also positively associated with hsCRP, 10 with leukocyte counts and 11 with neutrophil counts. Species associated with all three markers were dominated by species common in the oral microbiome[25] (e.g., *S. parasanguinis, S. oralis* subsp. *oralis, R. mucilaginosa, Ligilactobacillus salivarius*, and *S. anginosus*). Among the 13 species negatively associated with CACS, 6 were negatively associated with hsCRP, 6 with leukocyte counts, and 5 with neutrophil counts (**Figure 3**).

CACS-Associated Species from Fecal Samples Correlate with Corresponding Species in Saliva

Several of the CACS-associated species are normally considered as oral species. We therefore investigated the correlation between the abundance of CACS-associated species in the fecal and saliva samples from the 435 participants in MODS (**Table S9**) who underwent a thorough dental examination within 4 to 12 months after fecal sampling in the MOS. Six of the CACS-

associated species were identified in the saliva samples, of which five (*R. mucilaginosa*, *S. parasanguinis*, *S. gordonii*, *Bifidobacterium dentium*, *S. anginosus*) were positively associated with the abundance of the corresponding species in the fecal sample (rho=0.15–0.46) (**Table S10**). In up to 639 participants from MODS, three species (*S. anginosus, S. parasanguinis,* and *Rothia mucilaginosa*) were positively associated with both caries and filled surfaces, and three (*S. anginosus, S. gordonii and B. dentium*) with gingivitis (**Table S11**).

DISCUSSION

Gut bacteria have been proposed to affect the development and progression of atherosclerosis through infections local or distal to the atherosclerotic plaque or through production of metabolites affecting atherogenesis.[2, 26] The association of gut microbiota with coronary atherosclerosis has previously been studied only in symptomatic patients, who are often under treatment, resulting in high risk of bias. Here, detailed image-based measurements of coronary artery atherosclerosis and deep characterization of the gut microbiome by shotgun metagenomics in the large population-based SCAPIS cohort excluding participants with earlier cardiovascular disease were used to identify 67 species associated with CACS independent of risk factors. Several of these species were associated with circulating markers of inflammation and infection, and with the corresponding bacteria species in saliva samples, which were in turn associated with worse oral health. Thus, studies are merited to continue to investigate the role of these 67 species in atherogenesis.

We assessed the association of the CACS-associated species with other measures of coronary atherosclerosis from the CCTA as well as with carotid atherosclerosis measured by ultrasound. In general, we found high agreement of associations of species with CACS, modified Duke index, any atherosclerosis, and segment involvement score; however, the associations with having an occlusion >50% of at least one vessel were weaker, perhaps because this was a rather rare phenotype in our cohort, with low power. Likewise, only a few of the

species were associated with carotid plaque, perhaps because the measurement of carotid plaque was not very detailed or because associations differ in different vascular beds.

The *Streptococcus* genus was enriched for in associations between species and CACS, consistent with observations in previous case-control studies of atherosclerotic cardiovascular disease.[10, 11, 27] Specifically, in both SCAPIS and the validation study, *S. anginosus*, *S. oralis* subsp. *oralis*, *S. parasanguinis*, *S. gordonii* and *S. agalactiae* were associated with coronary atherosclerosis. These species all belong to the heterogenous viridans streptococci group (VGS)[28], except for *S. agalactiae* (a β-hemolytic non-VGS). In the mouth, VGS can enter the bloodstream through mucosal barrier injuries from daily dental care activities and dental procedures.[29] The gut could also be a site for expansion of these oral bacteria, and could potentially also act as an entry point. VGS can infect the valves and the coronary vessels, accounting for 20% of infective endocarditis cases, and have been isolated from samples of human coronary atherosclerotic plaque.[30-32] Studies in animal models support a causal link between *Streptococcus* species and atherogenesis.[6, 33] For instance, in atherosclerosis-prone mice, oral challenge with *S. sanguinis* elicited pro-inflammatory responses in the aorta and accelerated atherosclerosis.[6] VGS species are early colonizers, and may contribute to or initiate biofilm formations on different surfaces of the body. Biofilms are syntrophic polymicrobial communities that facilitate bacterial survival in aerobic environments and have been found in atherosclerotic lesions;[34] however, the relevance of biofilms in atherogenesis is unclear. In model systems, VGS has been shown to be able to invade human aortic endothelial cells and stimulate atherosclerosis-related pro-inflammatory cytokines.[35] In our study, the abundance of CACS-associated *Streptococcus* sp. in the gut associated strongly and positively with hsCRP, leukocytosis, and neutrophilia, which could have been triggered by low-grade bacteremia. In the current study, species carrying trehalose degradation genes were enriched for in associations with CACS, and *Streptococcus* spp. were an important contributor to this enrichment. Trehalose

degradation capacity has been linked to virulence, evasion from the immune system, resistance to antimicrobial, cell wall integrity and protection from starvation, and osmotic stress.[36, 37] Some *Streptococcus* spp. can catalyze different types of carbohydrates, including most common dietary sugars (fructose, glucose, and sucrose), forming compounds involved in virulence processes including biofilm formation.[38] Current knowledge of the involvement of these bacteria in cardiovascular pathologies is summarized in **Table S12**.

The abundance of CACS-associated species might be reduced by antimicrobial treatment. However, clinical trials showed that anti-infective therapies, targeted at chlamydial colonization of the coronary plaques, seemed harmful in patients with established coronary heart disease.[39] In the CLARICOR study, mortality was increased in patients with stable coronary artery disease treated with clarithromycin.[40] Possible explanations to why these antibiotics were ineffective in reducing coronary events include resistance to antibiotics due to biofilm formation,[41] recolonization after antibiotic treatment, the need for treatment earlier in atherosclerosis process, or simply that these bacteria are not causal to atherosclerosis. In the current study, participants with low abundance of CACS-associated *Streptococcus* spp. were not more exposed to antibiotics over the past 12 months than those with higher abundance. We also found that the abundance of *Streptococcus* spp. were associated with coronary artery calcification even after exclusion of participants treated with antibiotics.

The gut microbiota composition might also contribute to atherogenesis through alteration of the host metabolism. We found that the 54 species positively associated with CACS in the main model were enriched for positive associations with metabolites from several metabolite subclasses—including food/plant components (e.g., 2-piperidinone), aminosugar metabolism (e.g., N-acetylneuraminic acid), sterols (e.g., 7-HOCA), tobacco metabolites, cardiovascular drugs (e.g., enalapril, metoprolol) and xanthine metabolism —compared to the 1464 non-CACS-associated species. Several metabolites in these classes are reported risk markers for

CVD.[42, 43] The findings of diet and tobacco metabolite associations with these species, emphasizes the association of the gut microbiome composition with cardiovascular risk factors, and the importance of accounting for such potential confounders in species-CVD associations, as was done in the current study. We also found that CACS-associated species were enriched in associations with cardiovascular drugs (e.g., enalapril, metoprolol). This finding could be explained by the observed correlation of CACS-associated bacteria with cardiovascular risk factors or as effects of the bacteria (if these species cause CVD).

Our study has some limitations. First, few participants had high levels of subclinical atherosclerosis, reducing statistical power, However, our cohort is at least seven times larger than those analyzed previously. Second, microbial composition can vary extensively throughout the gastrointestinal tract. Fecal samples contain microbial populations from the distal colon, and less so from other sites such as the small intestine. Therefore, we could not identify associations of species that are not well represented in fecal samples. Third, we were unable to validate two-thirds of our findings in the validation cohort. As the validation was performed in a small case-control study of symptomatic disease, we could not determine whether the inability to validate some of our findings reflects the lack of a true association, insufficient statistical power, differences in study design, or cross-country differences. However, several of the top findings were replicated, in spite of these differences, which should be regarded as a strength. Fourth, our study does not take into consideration the different interactions among bacterial species such as synergistic effects in the relationship with coronary atherosclerosis. Fifth, the cross-sectional study design prevented causal inferences. Finally, our data on antibiotic treatment does not capture antibiotics provided in inpatient care; however, in this age group, we do not expect a large number of hospital-treated patients. In future studies, different causal inference methods should be used to determine whether the identified species are causally related to atherosclerosis development.

In conclusion, by combining data from a large population-based cohort study and highly accurate bioimaging to evaluate subclinical coronary atherosclerosis, we found that the abundance of several species in the gut were associated with coronary atherosclerosis, biomarkers of inflammation, and their oral counterparts. If causal, these species might contribute to atherogenesis by direct infection or by altering host metabolism. Future studies will show whether these species can be used as potential biomarkers or treatment targets.

Sources of Funding

The main financial support for the study was from the European Research Council [RC-2018-STG-801965 (T.F.), ERC-CoG-2014-649021 (M.O.-M.), and ERC-STG-2015-679242 (J.G.S.)]. Funding was also provided by the Swedish Research Council [VR 2019-01471 (T.F.), 2018-02784 (M.O.-M.), 2018-02837 (M.O.-M.), 2019-01015 (J.Ä.), 2020-00243 (J.Ä.), 2019-01236 (G.E.), 2017-02554 (J.G.S.), and EXODIAB 2009-1039 (M.O.-M.)]; the Swedish Heart-Lung Foundation [Hjärt-Lungfonden 20190505 (T.F.), 20200711 (M.O.-M.), 20180343 (J.Ä.), 2019-0526 (J.G.S.), 20200173 (G.E.)]; the A.L.F. governmental grant [2018-0148 (M.O.-M.)]; the Novo Nordisk Foundation [NNF20OC0063886 (M.O.-M.)]; the Swedish Diabetes Foundation [(DIA 2018-375 (M.O.-M.)]; the Swedish Foundation for Strategic Research [LUDC-IRC 15-0067 (M.O.-M.)]; Formas [2020-00989 (S.A.)]; Göran Gustafsson Foundationfoundation [2016 (T.F.)]; and Axel and Signe Lagerman's Foundation (T.F.).

The Swedish CArdioPulmonary bioImage Study (SCAPIS) was funded mainly by the Swedish Heart-Lung Foundation. The study was also funded by the Knut and Alice Wallenberg Foundation, the Swedish Research Council and VINNOVA (Sweden's innovation agency), the University of Gothenburg and Sahlgrenska University Hospital, Karolinska Institutet and Stockholm County Council, Linköping University and University Hospital, Lund University and Skåne University Hospital, Umeå University and University Hospital, and Uppsala University and University Hospital.

The Malmö Offspring Dental Study (MODS) was funded mainly by Oral Health Related Research by Region Skåne (OFRS 422361, OFRS 512951, OFRS 567711, OFRS 655561, OFRS 752071, OFRS 853031, OFRS 931171 and OFRS968144). The Malmö Offspring Study (MOS) was funded by the Research Council of Sweden [VR 521-2013-2756 (P.M.N.)], the Swedish Heart and Lung Foundation [Hjärt-Lungfonden 20150427 (P.M.N.)], and by funds ("ALF") from the local Region Skåne County Council (P.M.N.). In addition, funding was obtained from Ernhold Lundströms Stiftelse (L.B.).

The computations and data handling were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at Uppsala University (SNIC CENTRE), which is partially funded by the Swedish Research Council through grant agreement no. 2018-05973. The computations were done using resources provided by SNIC through the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project SNIC sens2019512.

Disclosures

The authors declare the following competing interests: Ms Pita, Ms Nielsen, Drs Eklund, Holm, and Nielsen are employees of Clinical Microbiomics A/S, where samples were processed and DNA extraction and estimations of relative abundance of the metagenomics species were done. Dr Ärnlöv has received lecture fees from Novartis and AstraZeneca and has served on advisory boards for AstraZeneca and Boerhinger Ingelheim, all unrelated to the present paper. Dr Nilsson has received lecture fees from Novartis, Novo Nordisk, Amgen, and Boerhinger Ingelheim. The other authors declare no competing interests.

Supplemental Material

Supplemental Methods

Tables S1-S12

Figures S1-S8

Table 1 Descriptive characteristics of participants in the SCAPIS study with metagenomics data and no previous CVD event

	Malmö	**Uppsala**	**Observations**
N	4541	4432	
Age (years)*	57.3 (4.3)	57.6 (4.4)	8973
Female, n (%)	2480 (54.6%)	2336 (52.7%)	8973
CACS†‡	0.0 [0.0;27.0]	0.0 [0.0;15.0]	8973
Clinical CACS, n (%)‡:			8973
0	2574 (56.7%)	2781 (62.7%)	
1-100	1379 (30.4%)	1187 (26.8%)	
101-400	398 (8.8%)	319 (7.2%)	
>400	190 (4.2%)	145 (3.3%)	
Any atherosclerosis, n (%)	1781 (45.9%)	1650 (40.8%)	7924
Stenosis > 50%, n (%)	224 (5.8%)	263 (6.5%)	7924
Modified Duke index >1, n (%)	1104 (28.4%)	968 (24.0%)	7924
Segment involvement score >4, n (%)	379 (9.8%)	283 (7.0%)	7924
Carotid arteries with identified plaques, n (%):			8955
None	1862 (41.1%)	1959 (44.2%)	
Unilateral	1388 (30.7%)	1391 (31.4%)	
Bilateral	1275 (28.2%)	1080 (24.4%)	
Country of birth, n (%):			8973
Scandinavia	3575 (78.7%)	4001 (90.3%)	
Non-Scandinavian Europe	629 (13.9%)	165 (3.7%)	
Asia	237 (5.2%)	170 (3.8%)	
Other	100 (2.2%)	96 (2.2%)	
Smoking, n (%):			8638
Never smoker	1948 (44.0%)	2469 (58.6%)	
Former smoker	1681 (38.0%)	1351 (32.1%)	
Current smoker	795 (18.0%)	394 (9.3%)	
Physical activity in leisure time, n (%):			8464
Sedentary	594 (13.8%)	425 (10.2%)	
Moderate exercise	2107 (48.9%)	1873 (45.1%)	
Moderate but regular exercise	1114 (25.8%)	1357 (32.7%)	
Regular exercise and training	498 (11.5%)	496 (11.9%)	
Triglycerides, mmol/L†	1.1 [0.8;1.5]	1.1 [0.8;1.5]	8955
LDL cholesterol, mmol/L *‡	3.6 (0.9)	3.6 (0.9)	8951
HDL cholesterol, mmol/L†‡	1.6 [1.3;2.0]	1.4 [1.2;1.7]	8952
Total cholesterol, mmol/L *	5.5 (1.0)	5.7 (1.0)	8954
Systolic blood pressure, mmHg*‡	122.3 (16.4)	124.9 (15.9)	8969
Diastolic blood pressure, mmHg*‡	74.7 (9.7)	76.9 (9.8)	8969

Body mass index, kg/m^{2}‡	27.2 (4.6)	27.0 (4.4)	8973
hsCRP, mmol/L†‡	1.2 [0.6;2.4]	1.2 [0.6;2.3]	8955
Neutrophil count, 10^9/L†	3.0 [2.4;3.8]	2.8 [2.3;3.5]	8930
Leukocyte count,10^9/L†	5.6 [4.7;6.7]	5.3 [4.6;6.3]	8937
Total energy intake, kcal/day†	1579.4 [1221.1;2066.8]	1613.0 [1270.7;2048.3]	8845
Carbohydrate intake, E%†‡	44.6 [39.9;49.0]	44.5 [40.3;48.5]	8845
Protein intake, E%†‡	16.8 [14.7;19.0]	17.0 [15.3;18.9]	8845
Fat intake, E%†‡	36.7 [32.7;40.6]	36.5 [32.7;40.3]	8845
Fiber intake, g/1000kcal†	11.7 [9.0;14.6]	11.8 [9.3;14.5]	8845
Diabetes, n (%)	198 (4.5%)	171 (4.1%)	8552
Inflammatory bowel disease, n (%)	43 (1.0%)	52 (1.2%)	8499
Medication for dyslipidemia, n (%)	292 (6.7%)	274 (6.5%)	8550
Medication for high blood pressure, n (%)	834 (19.2%)	764 (18.2%)	8544
Medication for diabetes, n (%)	165 (3.8%)	140 (3.3%)	8544
Proton-pump inhibitor, n (%)	161 (4.6%)	122 (2.8%)	7906
Narrow spectrum antibiotics treatment, n (%)	579 (12.8%)	550 (12.4%)	8973
Broad spectrum antibiotics treatment, n (%)	505 (11.1%)	400 (9.0%)	8973

*Mean (standard deviation)

†Median [interquartile range]

‡CACS indicates coronary artery calcium score; LDL, low-density lipoprotein; HDL, high-density lipoprotein; BMI, body mass index; hsCRP, high-sensitivity C-reactive protein, E%, percentage of total energy intake.

Figures

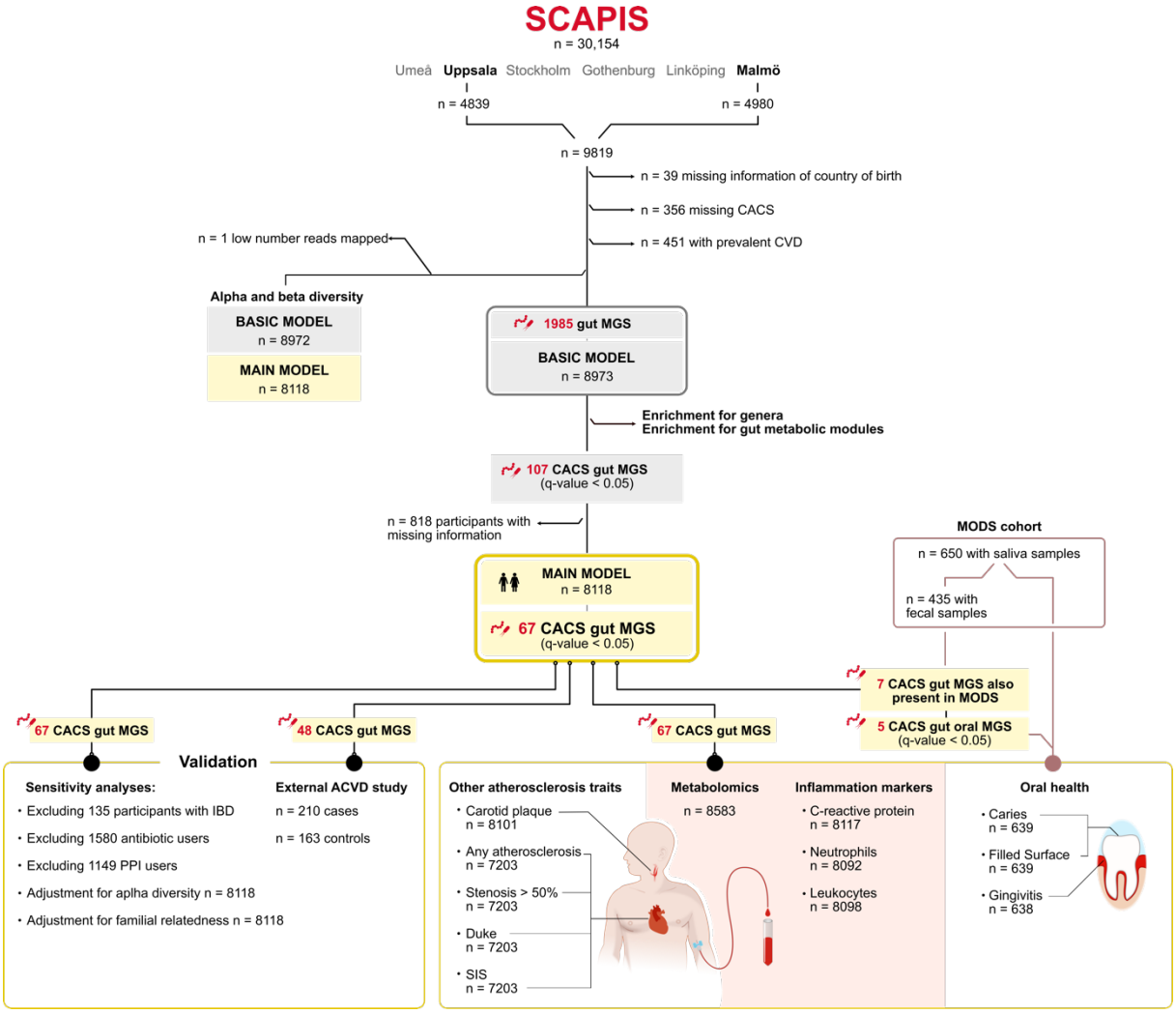

Figure 1. Flowchart of the overall design of the study.

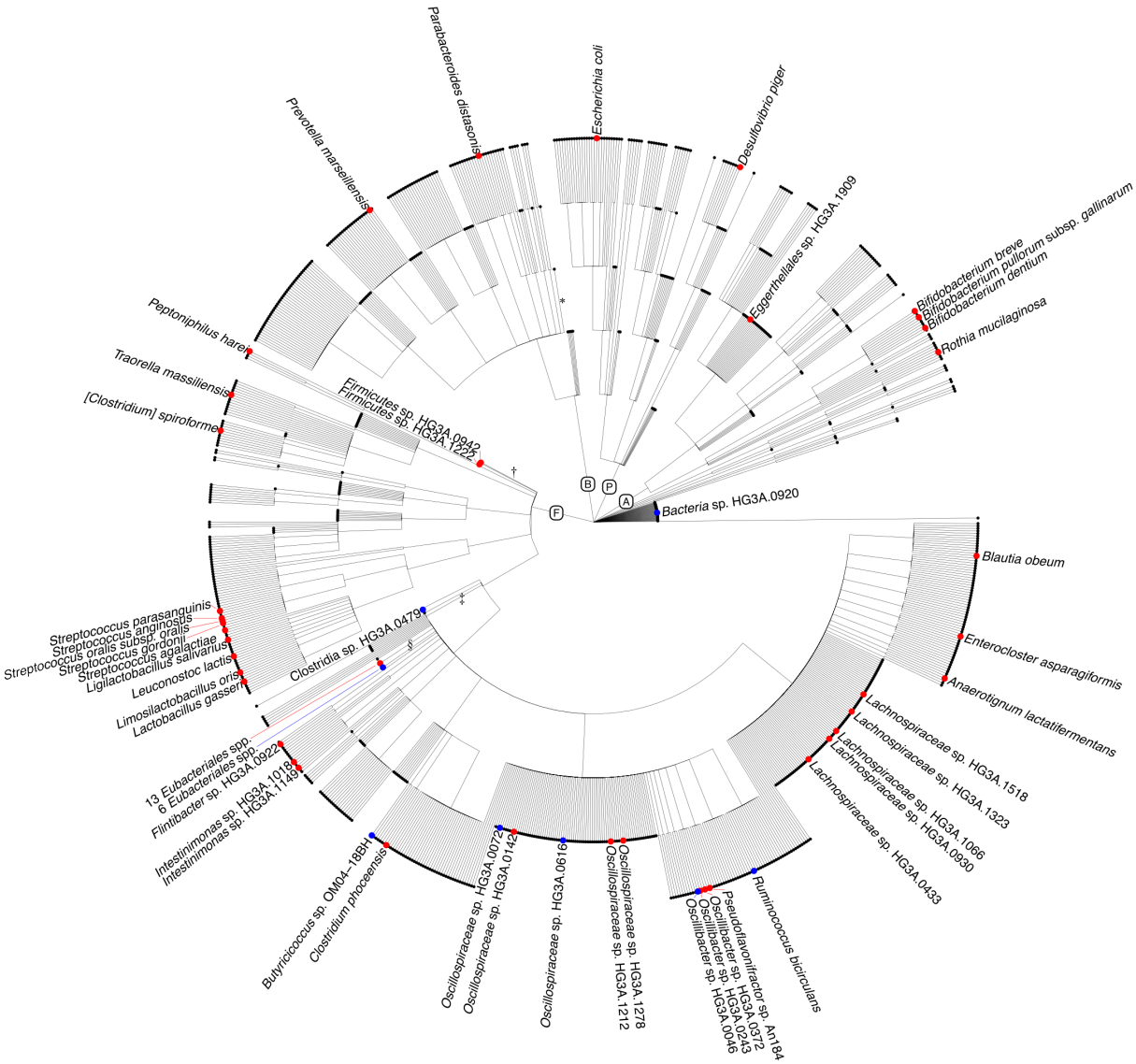

Figure 2. Cladogram of species investigated in the study.

Circles represent metagenomic species to the lowest taxonomic level annotated. Dots indicate the species positively (red) and negatively (blue) associated with CACS in the full model (false discovery rate <5%). A: Actinobacteria; P: Proteobacteria; B: Bacteroides; F: Firmicutes. The following unclassified non-CACS-associated metagenomics species were collapsed: *99 species with lowest annotation to the order Bacteroidales, †95 species with lowest annotation to phylum Firmicutes, ‡163 species with lowest annotation to the class Clostridia, and §649 species with lowest annotation to the order Eubacteriales.

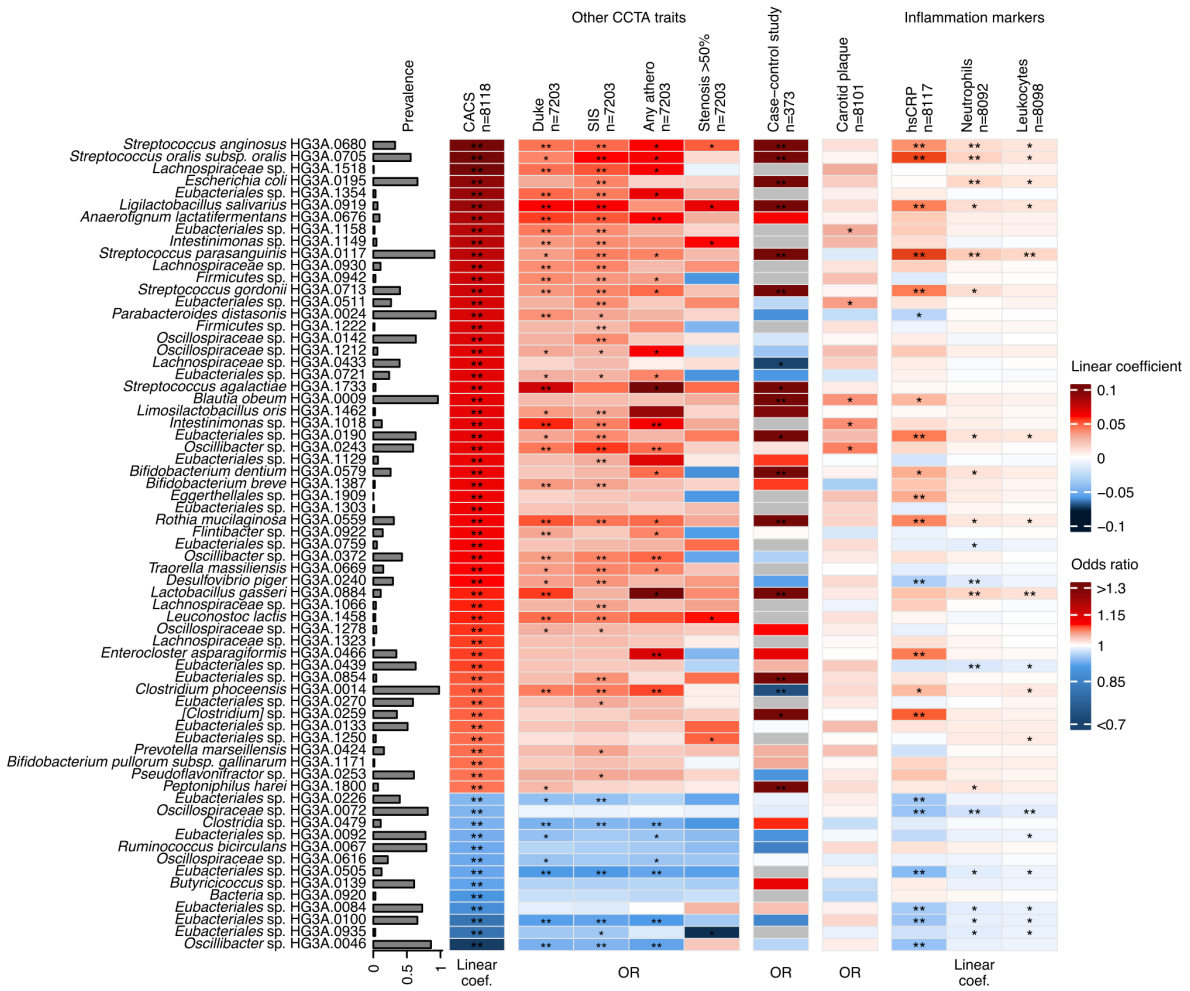

Figure 3. Associations between CACS-related gut species and alternate measurements of atherosclerosis and markers of inflammation and infection. Heatmap showing the associations of the 67 coronary artery calcium score (CACS)-associated species (FDR<0.05) with CACS, other coronary computed tomography angiography atherosclerosis traits, atherosclerotic cardiovascular disease (ASCVD) in re-analyzed data from the case-control study in Jie Z et al. (Nat Commun 8, 845 (2017), PMID: 29018189), and inflammation biomarkers. Associations with disease status in the case-control study is also presented as OR after adjusting for age and sex. Associations with CACS and with the inflammatory markers hsCRP, neutrophils and leukocytes counts are presented as linear regression coefficients after adjustment for the main model covariates (age, sex, country of birth, metagenomics extraction plate, smoking, physical activity, energy-adjusted carbohydrate, protein and fiber intake,

systolic and diastolic blood pressure, HDL, LDL, and total cholesterol, triglycerides, BMI, diabetes, and self-reported medication for dyslipidemia, hypertension, and diabetes). Associations with atherosclerosis traits (i.e., modified Duke index, segment involvement score (SIS), any atherosclerosis, and ≥50% stenosis) and presence of carotid plaques are presented as odds ratio (OR) after adjustment for main model covariates. Associations with FDR q-value <0.05 are marked with two asterisks (**) and those with P<0.05 are marked with one asterisk (*). Species not present in the case-control study are shown in gray. hsCRP: high-sensitivity C-reactive protein.

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton AZ, Benjamin EJ, Benziger CP, Bonny A, Brauer M, Brodmann M, Cahill TJ, Carapetis J, Catapano AL, Chugh SS, Cooper LT, Coresh J, Criqui M, DeCleene N, Eagle KA, Emmons-Bell S, Feigin VL, Fernández-Solà J, Fowkes G, Gakidou E, Grundy SM, He FJ, Howard G, Hu F, Inker L, Karthikeyan G, Kassebaum N, Koroshetz W, Lavie C, Lloyd-Jones D, Lu HS, Mirijello A, Temesgen AM, Mokdad A, Moran AE, Muntner P, Narula J, Neal B, Ntsekhe M, Moraes de Oliveira G, Otto C, Owolabi M, Pratt M, Rajagopalan S, Reitsma M, Ribeiro ALP, Rigotti N, Rodgers A, Sable C, Shakil S, Sliwa-Hahnle K, Stark B, Sundström J, Timpel P, Tleyjeh IM, Valgimigli M, Vos T, Whelton PK, Yacoub M, Zuhlke L, Murray C, Fuster V, Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton A, Benjamin EJ, Benziger CP, Bonny A, Brauer M, Brodmann M, Cahill TJ, Carapetis JR, Catapano AL, Chugh S, Cooper LT, Coresh J, Criqui MH, DeCleene NK, Eagle KA, Emmons-Bell S, Feigin VL, Fernández-Sola J, Fowkes FGR, Gakidou E, Grundy SM, He FJ, Howard G, Hu F, Inker L, Karthikeyan G, Kassebaum NJ, Koroshetz WJ, Lavie C, Lloyd-Jones D, Lu HS, Mirijello A, Misganaw AT, Mokdad AH, Moran AE, Muntner P, Narula J, Neal B, Ntsekhe M, Oliveira GMM, Otto CM, Owolabi MO, Pratt M, Rajagopalan S, Reitsma MB, Ribeiro ALP, Rigotti NA, Rodgers A, Sable CA, Shakil SS, Sliwa K, Stark BA, Sundström J, Timpel P, Tleyjeh II, Valgimigli M, Vos T, Whelton PK, Yacoub M, Zuhlke LJ, Abbasi-Kangevari M, Abdi A, Abedi A, Aboyans V, Abrha WA, Abu-Gharbieh E, Abushouk AI, Acharya D, Adair T, Adebayo OM, Ademi Z, Advani SM, Afshari K, Afshin A, Agarwal G, Agasthi P, Ahmad S, Ahmadi S, Ahmed MB, Aji B, Akalu Y, Akande-Sholabi W, Aklilu A, Akunna CJ, Alahdab F, Al-Eyadhy A, Alhabib KF, Alif SM, Alipour V, Aljunid SM, Alla F, Almasi-Hashiani A, Almustanyir S, Al-Raddadi RM, Amegah AK, Amini S, Aminorroaya A, Amu H, Amugsi DA, Ancuceanu R, Anderlini D, Andrei T, Andrei CL, Ansari-Moghaddam A, Anteneh ZA, Antonazzo IC, Antony B, Anwer R, Appiah LT, Arabloo J, Ärnlöv J, Artanti KD, Ataro Z, Ausloos M, Avila-Burgos L, Awan AT, Awoke MA, Ayele HT, Ayza MA, Azari S, B DB, Baheiraei N, Baig AA, Bakhtiari A, Banach M, Banik PC, Baptista EA, Barboza MA, Barua L, Basu S, Bedi N, Béjot Y, Bennett DA, Bensenor IM, Berman AE, Bezabih YM, Bhagavathula AS, Bhaskar S, Bhattacharyya K, Bijani A, Bikbov B, Birhanu MM, Boloor A, Brant LC, Brenner H, Briko NI, Butt ZA, Caetano dos Santos FL, Cahill LE, Cahuana-Hurtado L, Cámera LA, Campos-Nonato IR, Cantu-Brito C, Car J, Carrero JJ, Carvalho F, Castañeda-Orjuela CA, Catalá-López F, Cerin E, Charan J, Chattu VK, Chen S, Chin KL, Choi J-YJ, Chu D-T, Chung S-C, Cirillo M, Coffey S, Conti S, Costa VM, Cundiff DK, Dadras O, Dagnew B, Dai X, Damasceno AAM, Dandona L, Dandona R, Davletov K, De la Cruz-Góngora V, De la Hoz FP, De Neve J-W, Denova-Gutiérrez E, Derbew Molla M, Derseh BT, Desai R, Deuschl G, Dharmaratne SD, Dhimal M, Dhungana RR, Dianatinasab M, Diaz D, Djalalinia S, Dokova K, Douiri A, Duncan BB, Duraes AR, Eagan AW, Ebtehaj S, Eftekhari A, Eftekharzadeh S, Ekholuenetale M, El Nahas N, Elgendy IY, Elhadi M, El-Jaafary SI, Esteghamati S, Etisso AE, Eyawo O, Fadhil I, Faraon EJA, Faris PS, Farwati M, Farzadfar F, Fernandes E, Fernandez Prendes C, Ferrara P, Filip I, Fischer F, Flood D, Fukumoto T, Gad MM, Gaidhane S, Ganji M, Garg J, Gebre AK, Gebregiorgis BG, Gebregzabiher KZ, Gebremeskel GG, Getacher L, Obsa AG, Ghajar A, Ghashghaee A, Ghith N, Giampaoli S, Gilani SA, Gill PS, Gillum RF, Glushkova EV, Gnedovskaya EV, Golechha M, Gonfa KB, Goudarzian AH, Goulart AC, Guadamuz JS, Guha A, Guo Y, Gupta R, Hachinski V, Hafezi-Nejad N, Haile TG, Hamadeh RR, Hamidi

S, Hankey GJ, Hargono A, Hartono RK, Hashemian M, Hashi A, Hassan S, Hassen HY, Havmoeller RJ, Hay SI, Hayat K, Heidari G, Herteliu C, Holla R, Hosseini M, Hosseinzadeh M, Hostiuc M, Hostiuc S, Househ M, Huang J, Humayun A, Iavicoli I, Ibeneme CU, Ibitoye SE, Ilesanmi OS, Ilic IM, Ilic MD, Iqbal U, Irvani SSN, Islam SMS, Islam RM, Iso H, Iwagami M, Jain V, Javaheri T, Jayapal SK, Jayaram S, Jayawardena R, Jeemon P, Jha RP, Jonas JB, Jonnagaddala J, Joukar F, Jozwiak JJ, Jürisson M, Kabir A, Kahlon T, Kalani R, Kalhor R, Kamath A, Kamel I, Kandel H, Kandel A, Karch A, Kasa AS, Katoto PDMC, Kayode GA, Khader YS, Khammarnia M, Khan MS, Khan MN, Khan M, Khan EA, Khatab K, Kibria GMA, Kim YJ, Kim GR, Kimokoti RW, Kisa S, Kisa A, Kivimäki M, Kolte D, Koolivand A, Korshunov VA, Koulmane Laxminarayana SL, Koyanagi A, Krishan K, Krishnamoorthy V, Kuate Defo B, Kucuk Bicer B, Kulkarni V, Kumar GA, Kumar N, Kurmi OP, Kusuma D, Kwan GF, La Vecchia C, Lacey B, Lallukka T, Lan Q, Lasrado S, Lassi ZS, Lauriola P, Lawrence WR, Laxmaiah A, LeGrand KE, Li M-C, Li B, Li S, Lim SS, Lim L-L, Lin H, Lin Z, Lin R-T, Liu X, Lopez AD, Lorkowski S, Lotufo PA, Lugo A, M NK, Madotto F, Mahmoudi M, Majeed A, Malekzadeh R, Malik AA, Mamun AA, Manafi N, Mansournia MA, Mantovani LG, Martini S, Mathur MR, Mazzaglia G, Mehata S, Mehndiratta MM, Meier T, Menezes RG, Meretoja A, Mestrovic T, Miazgowski B, Miazgowski T, Michalek IM, Miller TR, Mirrakhimov EM, Mirzaei H, Moazen B, Moghadaszadeh M, Mohammad Y, Mohammad DK, Mohammed S, Mohammed MA, Mokhayeri Y, Molokhia M, Montasir AA, Moradi G, Moradzadeh R, Moraga P, Morawska L, Moreno Velásquez I, Morze J, Mubarik S, Muruet W, Musa KI, Nagarajan AJ, Nalini M, Nangia V, Naqvi AA, Narasimha Swamy S, Nascimento BR, Nayak VC, Nazari J, Nazarzadeh M, Negoi RI, Neupane Kandel S, Nguyen HLT, Nixon MR, Norrving B, Noubiap JJ, Nouthe BE, Nowak C, Odukoya OO, Ogbo FA, Olagunju AT, Orru H, Ortiz A, Ostroff SM, Padubidri JR, Palladino R, Pana A, Panda-Jonas S, Parekh U, Park E-C, Parvizi M, Pashazadeh Kan F, Patel UK, Pathak M, Paudel R, Pepito VCF, Perianayagam A, Perico N, Pham HQ, Pilgrim T, Piradov MA, Pishgar F, Podder V, Polibin RV, Pourshams A, Pribadi DRA, Rabiee N, Rabiee M, Radfar A, Rafiei A, Rahim F, Rahimi-Movaghar V, Ur Rahman MH, Rahman MA, Rahmani AM, Rakovac I, Ram P, Ramalingam S, Rana J, Ranasinghe P, Rao SJ, Rathi P, Rawal L, Rawasia WF, Rawassizadeh R, Remuzzi G, Renzaho AMN, Rezapour A, Riahi SM, Roberts-Thomson RL, Roever L, Rohloff P, Romoli M, Roshandel G, Rwegerera GM, Saadatagah S, Saber-Ayad MM, Sabour S, Sacco S, Sadeghi M, Saeedi Moghaddam S, Safari S, Sahebkar A, Salehi S, Salimzadeh H, Samaei M, Samy AM, Santos IS, Santric-Milicevic MM, Sarrafzadegan N, Sarveazad A, Sathish T, Sawhney M, Saylan M, Schmidt MI, Schutte AE, Senthilkumaran S, Sepanlou SG, Sha F, Shahabi S, Shahid I, Shaikh MA, Shamali M, Shamsizadeh M, Shawon MSR, Sheikh A, Shigematsu M, Shin M-J, Shin JI, Shiri R, Shiue I, Shuval K, Siabani S, Siddiqi TJ, Silva DAS, Singh JA, Mtech AS, Skryabin VY, Skryabina AA, Soheili A, Spurlock EE, Stockfelt L, Stortecky S, Stranges S, Suliankatchi Abdulkader R, Tadbiri H, Tadesse EG, Tadesse DB, Tajdini M, Tariqujjaman M, Teklehaimanot BF, Temsah M-H, Tesema AK, Thakur B, Thankappan KR, Thapar R, Thrift AG, Timalsina B, Tonelli M, Touvier M, Tovani-Palone MR, Tripathi A, Tripathy JP, Truelsen TC, Tsegay GM, Tsegaye GW, Tsilimparis N, Tusa BS, Tyrovolas S, Umapathi KK, Unim B, Unnikrishnan B, Usman MS, Vaduganathan M, Valdez PR, Vasankari TJ, Velazquez DZ, Venketasubramanian N, Vu GT, Vujcic IS, Waheed Y, Wang Y, Wang F, Wei J, Weintraub RG, Weldemariam AH, Westerman R, Winkler AS, Wiysonge CS, Wolfe CDA, Wubishet BL, Xu G, Yadollahpour A, Yamagishi K, Yan LL, Yandrapalli S, Yano Y,

Yatsuya H, Yeheyis TY, Yeshaw Y, Yilgwan CS, Yonemoto N, Yu C, Yusefzadeh H, Zachariah G, Zaman SB, Zaman MS, Zamanian M, Zand R, Zandifar A, Zarghi A, Zastrozhin MS, Zastrozhina A, Zhang Z-J, Zhang Y, Zhang W, Zhong C, Zou Z, Zuniga YMH, Murray CJL and Fuster V. Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology*. 2020;76:2982-3021.

2. Jonsson AL and Backhed F. Role of gut microbiota in atherosclerosis. *Nat Rev Cardiol*. 2017;14:79-87.
3. Ryden L, Buhlin K, Ekstrand E, de Faire U, Gustafsson A, Holmer J, Kjellstrom B, Lindahl B, Norhammar A, Nygren A, Nasman P, Rathnayake N, Svenungsson E and Klinge B. Periodontitis Increases the Risk of a First Myocardial Infarction: A Report From the PAROKRANK Study. *Circulation*. 2016;133:576-83.
4. Jonsson D, Orho-Melander M, Demmer RT, Engstrom G, Melander O, Klinge B and Nilsson PM. Periodontal disease is associated with carotid plaque area: the Malmo Offspring Dental Study (MODS). *J Intern Med*. 2020;287:301-309.
5. Tonetti MS, D'Aiuto F, Nibali L, Donald A, Storry C, Parkar M, Suvan J, Hingorani AD, Vallance P and Deanfield J. Treatment of periodontitis and endothelial function. *N Engl J Med*. 2007;356:911-20.
6. Hashizume-Takizawa T, Yamaguchi Y, Kobayashi R, Shinozaki-Kuwahara N, Saito M and Kurita-Ochiai T. Oral challenge with *Streptococcus sanguinis* induces aortic inflammation and accelerates atherosclerosis in spontaneously hyperlipidemic mice. *Biochemical and Biophysical Research Communications*. 2019;520:507-513.
7. Koren O, Spor A, Felin J, Fak F, Stombaugh J, Tremaroli V, Behre CJ, Knight R, Fagerberg B, Ley RE and Backhed F. Human oral, gut, and plaque microbiota in patients with atherosclerosis. *Proc Natl Acad Sci U S A*. 2011;108 Suppl 1:4592-8.
8. Chakaroun RM, Olsson LM and Backhed F. The potential of tailoring the gut microbiome to prevent and treat cardiometabolic disease. *Nat Rev Cardiol*. 2022.
9. Fromentin S, Forslund SK, Chechi K, Aron-Wisnewsky J, Chakaroun R, Nielsen T, Tremaroli V, Ji B, Prifti E, Myridakis A, Chilloux J, Andrikopoulos P, Fan Y, Olanipekun MT, Alves R, Adiouch S, Bar N, Talmor-Barkan Y, Belda E, Caesar R, Coelho LP, Falony G, Fellahi S, Galan P, Galleron N, Helft G, Hoyles L, Isnard R, Chatelier EL, Julienne H, Olsson L, Pedersen HK, Pons N, Quinquis B, Rouault C, Roume H, Salem J-E, Schmidt TSB, Vieira-Silva S, Li P, Zimmermann-Kogadeeva M, Lewinter C, Søndertoft NB, Hansen TH, Gauguier D, Gøtze JP, Køber L, Kornowski R, Vestergaard H, Hansen T, Zucker J-D, Hercberg S, Letunic I, Bäckhed F, Oppert J-M, Nielsen J, Raes J, Bork P, Stumvoll M, Segal E, Clément K, Dumas M-E, Ehrlich SD and Pedersen O. Microbiome and metabolome features of the cardiometabolic disease spectrum. *Nature Medicine 2022*. 2022;37:1-12.
10. Jie Z, Xia H, Zhong SL, Feng Q, Li S, Liang S, Zhong H, Liu Z, Gao Y, Zhao H, Zhang D, Su Z, Fang Z, Lan Z, Li J, Xiao L, Li J, Li R, Li X, Li F, Ren H, Huang Y, Peng Y, Li G, Wen B, Dong B, Chen JY, Geng QS, Zhang ZW, Yang H, Wang J, Wang J, Zhang X, Madsen L, Brix S, Ning G, Xu X, Liu X, Hou Y, Jia H, He K and Kristiansen K. The gut microbiome in atherosclerotic cardiovascular disease. *Nature Communications*. 2017;8:845.
11. Liu H, Chen X, Hu X, Niu H, Tian R, Wang H, Pang H, Jiang L, Qiu B, Chen X, Zhang Y, Ma Y, Tang S, Li H, Feng S, Zhang S and Zhang C. Alterations in the gut microbiome and metabolism with coronary artery disease severity. *Microbiome*. 2019;7:68.

12. Bergström G, Berglund G, Blomberg A, Brandberg J, Engström G, Engvall J, Eriksson M, Faire Ud, Flinck A, Hansson MG, Hedblad B, Hjelmgren O, Janson C, Jernberg T, Johnsson A, Johansson L, Lind L, Löfdahl CG, Melander O, Östgren CJ, Persson A, Persson M, Sandström A, Schmidt C, Söderberg S, Sundström J, Toren K, Waldenström A, Wedel H, Vikgren J, Fagerberg B and Rosengren A. The Swedish CArdioPulmonary BioImage Study: Objectives and design. *Journal of Internal Medicine*. 2015;278:645-659.
13. Pasolli E, Schiffer L, Manghi P, Renson A, Obenchain V, Truong DT, Beghini F, Malik F, Ramos M, Dowd JB, Huttenhower C, Morgan M, Segata N and Waldron L. Accessible, curated metagenomic data through ExperimentHub. *Nature Methods*. 2017;14:1023-24.
14. Brunkwall L, Jönsson D, Ericson U, Hellstrand S, Kennbäck C, Östling G, Jujic A, Melander O, Engström G, Nilsson J, Ohlsson B, Klinge B, Orho-Melander M, Persson M and Nilsson PM. The Malmö Offspring Study (MOS): design, methods and first results. *European Journal of Epidemiology*. 2021;36:103-116.
15. Bergström G, Persson M, Adiels M, Björnson E, Bonander C, Ahlström H, Alfredsson J, Angerås O, Berglund G, Blomberg A, Brandberg J, Börjesson M, Cederlund K, Faire Ud, Duvernoy O, Ekblom Ö, Engström G, Engvall JE, Fagman E, Eriksson M, Erlinge D, Fagerberg B, Flinck A, Gonçalves I, Hagström E, Hjelmgren O, Lind L, Lindberg E, Lindqvist P, Ljungberg J, Magnusson M, Mannila M, Markstad H, Mohammad MA, Nystrom FH, Ostenfeld E, Persson A, Rosengren A, Sandström A, Själander A, Sköld MC, Sundström J, Swahn E, Söderberg S, Torén K, Östgren CJ and Jernberg T. Prevalence of Subclinical Coronary Artery Atherosclerosis in the General Population. *Circulation*. 2021;144:916-929.
16. Ohnesorge B, Flohr T, Fischbach R, Kopp AF, Knez A, Schroder S, Schopf UJ, Crispin A, Klotz E, Reiser MF and Becker CR. Reproducibility of coronary calcium quantification in repeat examinations with retrospectively ECG-gated multisection spiral CT. *Eur Radiol*. 2002;12:1532-40.
17. Touboul PJ, Hennerici MG, Meairs S, Adams H, Amarenco P, Bornstein N, Csiba L, Desvarieux M, Ebrahim S, Hernandez RH, Jaff M, Kownator S, Naqvi T, Prati P, Rundek T, Sitzer M, Schminke U, Tardif JC, Taylor A, Vicaut E and Woo KS. Mannheim Carotid Intima-Media Thickness and Plaque Consensus (2004–2006–2011): An Update on Behalf of the Advisory Board of the 3rd and 4th Watching the Risk Symposium 13th and 15th European Stroke Conferences, Mannheim, Germany, 2004, and Brussels, Belgium, 2006. *Cerebrovascular diseases (Basel, Switzerland)*. 2012;34:290.
18. Nielsen HB, Almeida M, Juncker AS, Rasmussen S, Li J, Sunagawa S, Plichta DR, Gautier L, Pedersen AG, Chatelier EL, Pelletier E, Bonde I, Nielsen T, Manichanh C, Arumugam M, Batto JM, Santos MBQD, Blom N, Borruel N, Burgdorf KS, Boumezbeur F, Casellas F, Doré J, Dworzynski P, Guarner F, Hansen T, Hildebrand F, Kaas RS, Kennedy S, Kristiansen K, Kultima JR, Léonard P, Levenez F, Lund O, Moumen B, Paslier DL, Pons N, Pedersen O, Prifti E, Qin J, Raes J, Sørensen S, Tap J, Tims S, Ussery DW, Yamada T, Renault P, Sicheritz-Ponten T, Bork P, Wang J, Brunak S and Ehrlich SD. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature Biotechnology*. 2014;32:822-828.
19. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T,

Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD and Pruitt KD. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic acids research*. 2016;44:D733-D745.

20. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E and Wagner H. vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan. 2020.
21. Vieira-Silva S, Falony G, Darzi Y, Lima-Mendez G, Garcia Yunta R, Okuda S, Vandeputte D, Valles-Colomer M, Hildebrand F, Chaffron S and Raes J. Species–function relationships shape ecological properties of the human gut microbiome. *Nature Microbiology*. 2016;1:16088.
22. Dekkers KF, Sayols-Baixeras S, Baldanzi G, Nowak C, Hammar U, Nguyen D, Varotsis G, Brunkwall L, Nielsen N, Eklund AC, Bak Holm J, Nielsen HB, Ottosson F, Lin Y-T, Ahmad S, Lind L, Sundström J, Engström G, Smith JG, Ärnlöv J, Orho-Melander M and Fall T. An online atlas of human plasma metabolite signatures of gut microbiome composition. *Nature Communications*. 2022;13:5370.
23. Benjamini Y and Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995;57:289-300.
24. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN and Sergushichev A. Fast gene set enrichment analysis. *bioRxiv*. 2021:060012.
25. Caselli E, Fabbri C, D'Accolti M, Soffritti I, Bassi C, Mazzacane S and Franchi M. Defining the oral microbiome by whole-genome sequencing and resistome analysis: the complexity of the healthy picture. *BMC Microbiol*. 2020;20:120.
26. Ma J and Li H. The Role of Gut Microbiota in Atherosclerosis and Hypertension. *Frontiers in Pharmacology*. 2018;9:1082.
27. Liu S, Zhao W, Liu X and Cheng L. Metagenomic analysis of the gut microbiome in atherosclerosis patients identify cross-cohort microbial signatures and potential therapeutic target. *FASEB Journal*. 2020;34:14166-81.
28. Doern CD and Burnham C-AD. It's Not Easy Being Green: the Viridans Group Streptococci, with a Focus on Pediatric Clinical Manifestations. *Journal of Clinical Microbiology*. 2010;48:3829.
29. Lockhart PB, Brennan MT, Sasser HC, Fox PC, Paster BJ and Bahrani-Mougeot FK. Bacteremia associated with toothbrushing and dental extraction. *Circulation*. 2008;117:3118-25.
30. Koren O, Spor A, Felin J, Fåk F, Stombaugh J, Tremaroli V, Behre CJ, Knight R, Fagerberg B, Ley RE and Bäckhed F. Human oral, gut, and plaque microbiota in patients with atherosclerosis. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108:4592-4598.
31. Ott SJ, Mokhtari NEE, Musfeldt M, Hellmig S, Freitag S, Rehman A, Kühbacher T, Nikolaus S, Namsolleck P, Blaut M, Hampe J, Sahly H, Reinecke A, Haake N, Günther R, Krüger D, Lins M, Herrmann G, Fölsch UR, Simon R and Schreiber S. Detection of

diverse bacterial signatures in atherosclerotic lesions of patients with coronary heart disease. *Circulation*. 2006;113:929-37.
32. Viehman JA, Smith BJ, Nanjappa S, Li SK, Perez CO and Fernandes CR. Chapter 3 - Microbiology of endocarditis. In: A. Kilic, ed. *Infective Endocarditis*: Academic Press; 2022: 43-59.
33. Kesavalu L, Lucas AR, Verma RK, Liu L, Dai E, Sampson E and Progulske-Fox A. Increased atherogenesis during Streptococcus mutans infection in ApoE-null mice. *Journal of Dental Research*. 2012;91:255-60.
34. Lanter BB and Davies DG. Propionibacterium acnes recovered from atherosclerotic human carotid arteries undergoes biofilm dispersion and releases lipolytic and proteolytic enzymes in response to norepinephrine challenge In vitro. *Infection and Immunity*. 2015;83:3960-3971.
35. Nagata E, Toledo AD and Oho T. Invasion of human aortic endothelial cells by oral viridans group streptococci and induction of inflammatory cytokine production. *Molecular Oral Microbiology*. 2011;26:78-88.
36. Iturriaga G, Suárez R and Nova-Franco B. Trehalose Metabolism: From Osmoprotection to Signaling. *International Journal of Molecular Sciences*. 2009;10:3793-3810.
37. Tournu H, Fiori A and Van Dijck P. Relevance of Trehalose in Pathogenicity: Some General Rules, Yet Many Exceptions. *PLOS Pathogens*. 2013;9:e1003447.
38. Lemos JA, Palmer SR, Zeng L, Wen ZT, Kajfasz JK, Freires IA, Abranches J and Brady LJ. The Biology of *Streptococcus mutans*. *Microbiology spectrum*. 2019;7:7.1.03.
39. Sethi NJ, Safi S, Korang SK and Hróbjartsson A. Antibiotics for secondary prevention of coronary heart disease. *Cochrane Database Syst Rev*. 2019.
40. Jespersen CM, Als-Nielsen B, Damgaard M, Hansen JF, Hansen S, Helo OH, Hildebrandt P, Hilden J, Jensen GB, Kastrup J, Kolmos HJ, Kjoller E, Lind I, Nielsen H, Petersen L, Gluud C and Group CT. Randomised placebo controlled multicentre trial to assess short term clarithromycin for patients with stable coronary heart disease: CLARICOR trial. *BMJ*. 2006;332:22-7.
41. Sharma D, Misba L and Khan AU. Antibiotics versus biofilm: an emerging battleground in microbial communities. *Antimicrobial Resistance & Infection Control*. 2019;8:76.
42. Shah HS, Moreno LO, Morieri ML, Tang Y, Mendonca C, Jobe JM, Thacker JB, Mitri J, Monti S, Niewczas MA, Pennathur S and Doria A. Serum Orotidine: A Novel Biomarker of Increased CVD Risk in Type 2 Diabetes Discovered Through Metabolomics Studies. *Diabetes Care*. 2022;45:1882-1892.
43. Zhang L, Wei T-T, Li Y, Li J, Fan Y, Huang F-Q, Cai Y-Y, Ma G, Liu J-F, Chen Q-Q, Wang S-L, Li H, Alolga RN, Liu B, Zhao D-S, Shen J-H, Wang X-M, Zhu W, Li P and Qi L-W. Functional Metabolomics Characterizes a Key Role for *N*-Acetylneuraminic Acid in Coronary Artery Diseases. *Circulation*. 2018;137:1374-1390.
44. Christensen SE, Moller E, Bonn SE, Ploner A, Wright A, Sjolander A, Balter O, Lissner L and Balter K. Two new meal- and web-based interactive food frequency questionnaires: validation of energy and macronutrient intake. *J Med Internet Res*. 2013;15:e109.
45. Silness J and Loe H. Periodontal Disease in Pregnancy. Ii. Correlation between Oral Hygiene and Periodontal Condtion. *Acta Odontol Scand*. 1964;22:121-35.
46. Langmead B and Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods 2012 9:4*. 2012;9:357-359.

47. Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P, Collado MC, Rice BL, DuLong C, Morgan XC, Golden CD, Quince C, Huttenhower C and Segata N. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell*. 2019;176:649-662.e20.
48. Li D, Liu CM, Luo R, Sadakane K and Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics (Oxford, England)*. 2015;31:1674-1676.
49. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H and Lam TW. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods (San Diego, Calif)*. 2016;102:3-11.
50. Steinegger M and Söding J. Clustering huge protein sequence sets in linear time. *Nature communications*. 2018;9:2542.
51. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, Mering CV and Bork P. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*. 2019;47:D309-D314.
52. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P and Huerta-Cepas J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*. 2021;38:5825-5829.