Streptococcus species abundance in the gut is linked to subclinical coronary atherosclerosis in 8973 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

27 [8]Clinical Microbiomics A/S, Copenhagen, Denmark

28 [9]The Novo Nordisk Foundation Center for Biosustainability, Technical University of
29 Denmark, Lyngby, Denmark

30 [10]Clinical Studies Sweden, Forum Söder, Region Skåne, Lund, Sweden

31 [11]Section for Clinical Mass Spectrometry, Danish Center for Neonatal Screening, Department
32 of Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark

33 [12]Department of Cariology, Faculty of Odontology, Malmö University, Malmö, Sweden

34 [13]Department of Dental Medicine, Karolinska Institutet, Solna, Sweden

35 [14]Department of Internal Medicine, Skånes University Hospital, Malmö, Sweden

36 [15]Department of Medical Sciences, Clinical Physiology, Uppsala University, Uppsala,
37 Sweden

38 [16]Department of Medical Sciences, Clinical Epidemiology, Uppsala University, Uppsala,
39 Sweden

40 [17]Department of Molecular and Clinical Medicine, Institute of Medicine, Sahlgrenska
41 Academy, University of Gothenburg, Gothenburg, Sweden

42 [18]Department of Clinical Physiology, Sahlgrenska University Hospital, Region Västra
43 Götaland, Gothenburg, Sweden

44 [19]The George Institute for Global Health, University of New South Wales, Sydney, Australia

45 [20]School of Health and Social Studies, Dalarna University, Falun, Sweden

46 [21]The Wallenberg Laboratory/Department of Molecular and Clinical Medicine, Institute of
47 Medicine, Gothenburg University and the Department of Cardiology, Sahlgrenska University
48 Hospital, Gothenburg, Sweden

49 [22]Department of Cardiology, Clinical Sciences, Lund University and Skåne University
50 Hospital, Lund, Sweden

51 [23]Wallenberg Center for Molecular Medicine and Lund University Diabetes Center, Lund
52 University, Lund, Sweden

53 **Short title:** Gut streptococci are linked to atherosclerosis

54 Correspondence to: Tove Fall, PhD, Professor in Molecular Epidemiology

55 Department of Medical Sciences, Molecular Epidemiology

56 Uppsala University, EpiHubben, MTC-huset

57 75185, Uppsala, Sweden.

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

59 **Word Count**

60 Abstract: 346

61 Main text: 4951

ABSTRACT

BACKGROUND: Gut microbiota have been implicated in atherosclerotic disease, but their relation with subclinical coronary atherosclerosis is unclear. This study aimed to identify associations between the gut microbiome and computed tomography-based measures of coronary atherosclerosis, and to explore relevant clinical correlates.

METHODS: We conducted a cross-sectional study of 8973 participants aged 50 to 65 without overt atherosclerotic disease from the population-based Swedish Cardiopulmonary BioImage Study (SCAPIS). Coronary atherosclerosis was measured using coronary artery calcium score (CACS) and coronary computed tomography angiography (CCTA). Gut microbiota species abundance and functional potential were assessed with shotgun metagenomics sequencing of stool samples, and their association with coronary atherosclerosis was evaluated with multivariable regression models adjusted for cardiovascular risk factors. Associated species were evaluated for association with inflammatory markers, metabolites, and corresponding species in saliva.

RESULTS: The mean age of the study sample was 57.4 years, and 53.7% were female. Coronary artery calcification was detected in 40.3% of participants, and 5.4% had at least one stenosis with more than 50% occlusion. Sixty-four species were associated with CACS independent of cardiovascular risk factors, with the strongest associations observed for *Streptococcus anginosus* and *S. oralis* subsp. *oralis* ($P<1x10^{-5}$). Associations were largely similar across CCTA-based measurements. Out of the 64 species, 19 species, including streptococci and other species commonly found in the oral cavity, were associated with high-sensitivity C-reactive protein plasma concentrations and 16 with neutrophil counts. Oral species in the gut were negatively associated with plasma indole propionate and positively associated with plasma secondary bile acids and imidazole propionate. Five species correlated with the same species in saliva and were associated with worse dental health in the Malmö Offspring Dental Study. Microbial functional potential of dissimilatory nitrate reduction, anaerobic fatty acid beta-oxidation and amino acid degradation was associated with CACS.

CONCLUSIONS: This study provides evidence of an association of a gut microbiota composition characterized by increased abundance of *Streptococcus* spp. and other species commonly found in the oral cavity with coronary atherosclerosis and systemic inflammation. Further longitudinal and experimental studies are warranted to explore the potential implication of a bacterial component in atherogenesis.

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

CLINICAL PERSPECTIVE

WHAT IS NEW?

- Shotgun metagenomics identified associations between gut species and subclinical atherosclerosis assessed with computed tomography-derived coronary artery calcium score (CACS) in 8973 participants, with an overrepresentation of the *Streptococcus* and *Oscillobacter* genera.
- The relative abundance of CACS-associated oral species detected in fecal samples was negatively associated with indole propionate, while positively associated with secondary bile acids and imidazole propionate.
- Gut *Streptococcus* spp. were positively 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 microbiota composition, 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.

114 **Non-standard Abbreviation and Acronyms**

115 CACS	Coronary artery calcium score
116 CCTA	Coronary computed tomography angiography
117 CVD	Cardiovascular disease
118 GMM	Gut metabolic module
119 hsCRP	High-sensitivity C-reactive protein
120 ln	Natural logarithm
121 MGS	Metagenomics species
122 MODS	Malmö Offspring Dental Study
123 MOS	Malmö Offspring Study
124 SCAPIS	Swedish CArdioPulmonary bioImage Study
125 SIS	Segment involvement score
126 VGS	Viridans streptococci group

127 Atherosclerotic cardiovascular disease (CVD) is a major cause of death and disability. [1] The
128 microbial community of the gastrointestinal tract, also known as the gut microbiota, is
129 hypothesized to affect the progression of atherosclerosis through three potential mechanisms.
130 Firstly, microbial metabolites can interfere with the host metabolism, including lipid
131 metabolism. [2] The composition of the gut microbiota has been linked to metabolic disorders
132 such as obesity, insulin resistance and type 2 diabetes, although the causal relation and direction
133 are unclear. [3] Secondly, translocation of live bacteria or bacterial structural components (e.g.,
134 endotoxins) into the bloodstream can contribute to low-grade systemic inflammation, thereby
135 exacerbating the atherosclerosis process. [2] Thirdly, the discovery of bacteria within
136 atherosclerotic plaques has led to the proposal that bacteria might directly infect plaques and
137 accelerate atherosclerosis progression. [4] Case-control studies of symptomatic coronary
138 atherosclerotic disease have identified differences in the abundance of more than 500 gut
139 species. [5-7] However, dissimilarities in medical treatment and lifestyle factors in cases and
140 controls make these studies prone to bias. Thus, studies of individuals without overt coronary
141 disease in large population-based samples incorporating biomarkers of metabolism and
142 inflammation are warranted.

143 Recent evidence supports that oral species are commonly transmitted to the gut, [8] indicating
144 that the gut and oral microbiota are connected rather than being two separate microbial
145 communities. Causative species of dental caries, plaque-dwelling bacteria and endocarditis-
146 associated species including species from the viridans streptococci group (VGS) are reported
147 to be transmitted at a high rate. Additionally, an overlap is reported between oral bacteria such
148 as *Veillonella* spp. and *Streptococcus* spp. found in oral cavity, fecal samples and carotid
149 atherosclerotic plaque. [9] Given the association between dental health and endothelial
150 dysfunction and atherosclerotic disease, [10-12] the gut may serve as a niche and/or entry route for

151 oral pathogenic bacteria passing to the blood. However, there is a lack of studies assessing the
152 association of atherosclerosis-related gut species with the oral microbiome.
153 To avoid the limitations of previous studies and gain insight into the relation of the gut
154 microbiota and atherosclerosis, we sought to identify associations between the gut microbiome
155 and computed tomography-based measures of subclinical coronary atherosclerosis in a large
156 cohort of middle-aged participants from the Swedish CArdioPulmonary bioImage Study
157 (SCAPIS). We further assessed associations between atherosclerosis-associated gut bacterial
158 species and biomarkers of inflammation and infection, plasma metabolites, and the abundance
159 of corresponding bacterial species in the oral cavity.

160 **METHODS**

161 **Data Availability**

162 De-identified metagenomic sequencing data can be accessed from the ENA under accession
163 number PRJEB51353 (https://www.ebi.ac.uk/ena/browser/view/prjeb51353). Access to the
164 pseudonymized data requires ethical approval from the Swedish Ethical Review Board and
165 approval from the SCAPIS Data access board (https://www.scapis.org/data-access/). The
166 source code and the summary data underlying all figures used to generate the results for the
167 analysis are available at https://github.com/MolEpicUU/GUTSY_CACS.

168 **Study Design and Participants**

169 The study design is illustrated in **Figure 1**. The primary data source was SCAPIS, a population-
170 based study focusing on cardiovascular and respiratory disease in 30154 subjects 50–64 years
171 of age from six sites in Sweden. [13] Fecal metagenomics data was available for 4839 participants
172 from Uppsala and 4980 from Malmö after removal of 10 samples for failing in the quality
173 control. After excluding subjects with missing information on country of birth (n=39), coronary
174 artery calcium score (CACS) (n=356) or who had self-reported CVD (n=451) including

175 atherosclerotic CVD (n=119), 4541 participants from Uppsala and 4432 from Malmö
176 remained.

177 The association between gut species and their counterparts in saliva samples and with oral
178 health was investigated in participants of the Malmö Offspring Dental Study (MODS, n=831,
179 mean age 52.9 years), a substudy of the family-based Malmö Offspring Study (MOS). [10,14]
180 Salivary metagenomics, oral examination data, and complete data on age and sex were
181 available for 650 subjects; informative fecal metagenomics data from MOS were available for
182 435 of those subjects.

183 All participants gave written informed consent. The study was conducted in accordance
184 with the Declaration of Helsinki and was approved by the regional ethics committees (SCAPIS:
185 Dnr 2010-228-31M and Dnr 2018/315; MODS: Dnr 2013/761; and MOS: Dnr 2012/594).

186 **Atherosclerosis Measurements**

187 Cardiac imaging in SCAPIS included coronary computed tomography angiography (CCTA) as
188 described previously. [15] Briefly, computed tomography was performed with a dedicated dual-
189 source scanner and a Stellar Detector (Somatom Definition Flash, Siemens Medical Solutions).
190 Non-contrast images for calcium scoring were obtained from all subjects, followed by contrast-
191 enhanced images from subjects without contraindications. CACS was determined using the
192 syngo.via calcium scoring software (Volume Wizard; Siemens). The area of calcification was
193 summed for the whole coronary artery tree to a CACS according to Agatston. [16] Radiologists
194 inspected the images for any coronary atherosclerosis, segment involvement score (SIS), Duke
195 prognostic coronary artery disease index modified for SCAPIS, [15] and occlusive atherosclerosis
196 (≥50% stenosis in at least one vessel) (see **Supplemental Methods)**. We included individuals
197 with valid readings in at least four proximal segments. Carotid atherosclerosis was categorized
198 as absent, unilateral, or bilateral[13] from two-dimensional gray-scale ultrasound images obtained

199 with a Siemens Acuson S2000 ultrasound scanner and a 9L4 linear transducer (both from
200 Siemens, Germany) using a standardized protocol and the Mannheim consensus. [17]

201 **Metagenomics**

202 DNA was extracted from fecal samples (SCAPIS and MOS) and saliva samples (MODS) at
203 Clinical Microbiomics (Copenhagen, Denmark). The analytical pipeline is described in detail
204 in **Supplemental Methods**. In brief, libraries of fragmented DNA were sequenced with the
205 Illumina Novaseq 6000 platform. This technique generated an average of 26.3 million read
206 pairs per sample in SCAPIS-Malmö and MOS, 25.3 million read pairs in SCAPIS-Uppsala,
207 and 26.3 million read pairs in MODS. Non-host reads and data from these and other cohorts
208 were used to build separate nonredundant gene catalogues for fecal and saliva samples.
209 Metagenomic species and corresponding signature gene were identified. [18] In all three datasets,
210 species abundance was estimated by mapping reads to the signature gene sets and
211 normalization for effective gene length. The taxonomic information for the two catalogues was
212 mapped with the NCBI RefSeq database. [19]

213 Alpha (Shannon diversity index, inverse Simpson index, and Chao1) and beta diversity
214 (Bray-Curtis dissimilarity) were estimated with the R v4.1.1 package *vegan* v2.5-7 using
215 rarefied data.[20] Rarefied data were also used to define individuals with above/less than median
216 abundance of each species. The functional potential profile of the gut microbiota was
217 determined by assigning genes to gut metabolic modules (GMM),[20] which includes 103
218 metabolic pathways that represent cellular enzymatic processes. The GMM abundance was
219 estimated using the *Omixer-RPM* v0.3.226 R package, with a minimum module coverage
220 threshold set at 66.6%. Before conducting statistical analysis, the abundance data for each
221 species in all analytical datasets and for each GMM were transformed with the formula ln(x+1),
222 where x denotes the relative abundance of each species or GMM, followed by z-transformation
223 to set the mean to 0 and SD to 1.

224 **Clinical Chemistry, Hematology and Metabolomics**

225 Venous plasma samples were analyzed for high-density lipoprotein (HDL) and low-density
226 lipoprotein (LDL) cholesterol, triglycerides, and high-sensitivity C-reactive protein (hsCRP).
227 Blood cells were counted by standard methods. CACS-associated species-plasma metabolite
228 associations in SCAPIS with q-value <0.05 were accessed from the GUTSY atlas
229 (https://gutsyatlas.serve.scilifelab.se/) from Dekkers et al. [21]

230 **Oral Health Phenotypes**

231 Dental examinations of MODS participants were done by five dentists. Both surfaces with
232 caries and fillings were recorded on all teeth, counting four to five surfaces per tooth. Surfaces
233 with caries were detected using standard clinical criteria aided by mirror, probe (Hu-Friedy
234 EXD57), and bite-wing radiographs combining initial and manifest lesions. Filled surfaces
235 included both fillings and crowns. Gingival inflammation was recorded as percentage of
236 bleeding on probing, excluding wisdom teeth, and counting six surfaces per tooth.

237 **Other Phenotypes**

238 In SCAPIS and MOS, data on country of birth, smoking, physical activity, and diet were
239 collected with validated standardized questionnaires. Blood pressure and body mass index
240 (BMI) were recorded. Data on medications were acquired from the drug prescription register,
241 self-report, plasma measurements, and other sources. Relatedness in SCAPIS was based on
242 kinship analysis of genotype data, assigning first-degree relatives individuals to the same
243 family, and relatedness in MODS/MOS was based on family information from the study design
244 phase. Collected phenotypes are described in detail in **Supplemental Methods**.

245 **Statistical Analysis**

246 *Association between Gut Microbiota Diversity and Atherosclerosis* The association of alpha
247 diversity with coronary atherosclerosis was assessed by linear mixed models with ln(CACS+1)

248 as the dependent and alpha diversity indices as the independent variables. To address the
249 potential influence of family relatedness, the first degree-relatedness among 423 participants
250 from 208 families was accounted for with a random effect variable by using the *lmerTest* v3.1-
251 3 R package. Furthermore, two sets of covariates were modeled as fixed effects. The basic
252 model was adjusted for age, sex, country of birth, study center, and metagenomics extraction
253 plate. The main model was further adjusted for smoking; physical activity; energy-adjusted
254 carbohydrate, protein, and fiber intake; systolic and diastolic blood pressure; total cholesterol;
255 high- and low-density lipoprotein cholesterol; ln(triglycerides); diabetes; BMI; and self-
256 reported medication for dyslipidemia, hypertension, and diabetes. Correlation of covariates and
257 alpha diversity measurements are reported in **Figure S1**.

258 Beta diversity was assessed by a distance-based multivariate analysis of variance of
259 ln(CACS+1) as the independent variable using the *dmanova* function in the *GUniFrac* v1.3
260 package. These analyses were adjusted for the same two set of covariates as used in the alpha
261 diversity analyses. However, in contrast to the alpha diversity analyses, only one participant
262 from each family cluster was included, rather than by treating family relatedness as a random
263 effect variable.

264 *Association of Species and Bacterial Functions with CACS.* Extensive simulations were
265 conducted before the analysis and supported the use of linear regression to minimize false-
266 positive findings (**Supplemental Methods, Figure S2**). Hence, a series of linear mixed
267 multivariable regressions were performed with the ln(CACS+1) as the dependent variable and
268 the abundance of each species and GMM as the independent variable. The models were fitted
269 for each species and each GMM separately; covariates from the basic model were used as fixed
270 effects and family relatedness as random effects. Multiple testing was accounted for using the
271 Benjamini-Hochberg False Discovery Rate (FDR) at 5%.[22] A Taxon Set Enrichment Analysis[23]

272 based on the *fgsea* v1.18.0 R package based on *P* value ranking separately for positive and
273 negative regression coefficients was used to determine whether certain genera were enriched.
274 Species with a q-value <0.05 in the basic model were assessed adjusting for main model
275 covariates, again accounting for multiple testing. To determine whether associations were
276 caused by influential observations, unscaled dfbeta values were calculated using the *influence*
277 and *dfbetas* R functions. The linear regression estimates were deemed unreliable if the
278 exclusion of the observations with highest absolute dfbeta value resulted in a change of the
279 direction of the regression coefficient or in a $P \geq 0.05$. Findings not caused by an influential
280 observation are referred to as "CACS-associated species". Effect estimate modification by sex
281 was assessed for CACS-associated species by entering an interaction term between sex and
282 each of the fixed effect variables in the model and extracting the interaction coefficient of sex
283 with species.

284 *Sensitivity Analyses*. In three separate analyses, we excluded three groups of participants: those
285 in whom proton-pump inhibitors were measured in plasma, those with inflammatory bowel
286 disease, and those treated with antibiotics during the year before the baseline visit.
287 Furthermore, the analyses for CACS-associated species were repeated with the Shannon
288 diversity index as a covariate to test if findings were driven by global changes in the gut
289 microbiota composition rather than change in individual species. Lastly, the associations from
290 the main model were tested in an ordinal mixed model using CACS categories as the outcome
291 (CACS: 0; 1-100; 101-400; >400) using Stata v15.

292 *Association between CACS-Associated Species and CCTA-Based Measurements and Carotid*
293 *Plaques.* We assessed the association of CACS-associated species with presence of any
294 coronary atherosclerotic plaque or coronary stenosis >50% in at least one vessel in a series of
295 logistic mixed regressions. The association between CACS-associated species and SIS,

296 modified Duke index, and number of carotid vessels with atherosclerosis plaques were assessed
297 with ordinal mixed regression models. Categories 4, 5 and 6 of the modified Duke index were
298 merged due to low number of observations. The logistic and ordinal mixed models were fitted
299 using Stata v15 and were adjusted for the main model covariates.

300 *Clustering of Plasma Metabolites associated with CACS-Associated Species.* For each CACS-
301 associated species, the three strongest associations with plasma metabolites were selected
302 based on their p-value. For the resulting unique subset of 63 metabolites hierarchical clustering
303 was performed based on the Euclidian distance on the partial Spearman correlation coefficients.

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

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

320 adjusted for age, sex, smoking, education, Silness-Löe plaque index, as a marker of oral
321 hygiene, and the activity in the hour immediately before the dental examination (e.g., eating,
322 brushing teeth, smoking) as fixed effects and family relatedness as a random effect.

323 **RESULTS**

324 **Gut Microbiota Composition and Richness Are Associated with Subclinical**
325 **Atherosclerosis and Attenuated by Adjustment for Lifestyle Factors, Diet and Medication**

326 The study included 8973 subjects with no self-reported history of CVD. Among them, 40.3%
327 had measurable coronary artery calcification and 5.4% manifested at least one stenosis with
328 >50% occlusion (**Table 1**, **Table S1**). The Shannon diversity index—a measure of overall
329 species richness and evenness within each sample—was inversely associated with CACS
330 (n=8972, β=-0.17, P=5.1×10^{-4}) in the basic model. The basic model was adjusted for age, sex,
331 country of birth, and technical variables as fixed effects, and family relatedness as a random
332 effect. However, the association was attenuated and non-significant in the main model with
333 more extensive adjustment (n=8118, β=-0.03, P=0.53). The covariates primarily attenuating
334 the association were triglycerides and BMI, followed by medication, blood pressure traits,
335 smoking, HDL cholesterol, and physical activity (Figure S3). Alternative alpha diversity
336 indices were similarly associated with CACS (**Figure 2a**). Beta diversity—a measure of the
337 amount of variation in species composition—differed among CACS values in both models,
338 although the main model was attenuated (r^2_{basic}=0.0005; P_{basic} = 2×10^{-25} vs r^2_{main}=0.0002;
339 P_{main}=0.009) mainly by the same covariates that attenuated the relationship between alpha
340 diversity and CACS (**Figure 2b**, **Figure S3**). These models were not adjusted for family
341 relatedness, but only one of the participants in each family cluster was included instead (n_{basic}=
342 8757, n_{main}= 7921). These findings encouraged us to investigate species-specific associations.

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

345 The taxonomic profiling of fecal samples identified the prevalence of 1985 species, with 411
346 species having a prevalence <1%. The average number of species per sample was 325, with a
347 range of 11 to 696. The dominant phyla were Firmicutes (72%) and Bacteroidetes (20%)

348 (**Figure 3**). In the basic model, including 8973 individuals, the relative abundance of 78 species
349 (73 of which had a prevalence ≥1%) showed a positive association with CACS, while 29
350 species (all of which had a prevalence ≥1%) were negatively associated. (**Table S2**). Taxon-
351 set enrichment analysis revealed an overrepresentation of the *Streptococcus* and *Oscillibacter*
352 genera in the positive CACS associations (**Table S3**).

353 **Sixty-four Species Were Associated with CACS Independent of Cardiovascular Risk**
354 **Factors**

355 In the main model, which included 8118 individuals with complete data on covariates, 67
356 species remained associated with CACS (54 positively and 13 negatively) (**Figure 4**).
357 Covariates included age, sex, country of birth, metagenomics extraction plate, smoking,
358 physical activity, energy-adjusted carbohydrate, protein, fiber intake, blood pressure, lipids,
359 diabetes, BMI, self-reported cardiovascular medication, and family relatedness. However, the
360 associations of *Peptoniphilus harei*, *Muribaculaceae* sp. (MGS identifier: HG3A.1967) and
361 *Eubacteriales* sp. (HG3A.0270) were deemed non-robust, as the removal of single influential
362 observations abolished the association (**Table S4**). Among the remaining 64 CACS-associated
363 species, 51 had a positive association with CACS, while 13 had negative associations (**Figure**
364 **4**, **Figure 5**). The abundance of bacteria across CACS groups and the previously excluded
365 SCAPIS participants with self-reported atherosclerotic CVD is shown in **Figure S4**. The lowest
366 *P* values were observed for *Streptococcus anginosus*, *S. oralis* subsp. *oralis*, *Escherichia coli*,
367 *Eubacteriales* sp. (HG3A.1354) and *Intestinimonas* sp. (HG3A.1149), all positively associated
368 with CACS. Other associated species included *S. parasanguinis*, *S. gordonii*, and
369 *S. agalactiae*. Comparison of the main clinical characteristics in participants with high and low
370 abundance of CACS-associated species (cut-off median) showed that those with higher
371 abundance of the two most strongly associated species, *S. anginosus,* and *S. oralis* subsp. *oralis*
372 had in general more cardiovascular risk factors (**Table S5**). Furthermore, there were indications

373 that the associations between *S. agalactiae*, *Rothia mucilaginosa*, two *Eubacteriales* spp.
374 (HG3A.0511 and HG3A.0854), and an *Oscillibacter* sp. (HG3A.0243) and CACS were
375 modified by sex (interaction $P<0.05$, **Table S4**). After adjusting for multiple testing the
376 associations were no longer significant.

377 **Sensitivity Analyses**

378 The exclusion of individuals who used proton-pump inhibitors, had inflammatory bowel
379 disease, or were treated with antibiotics in the previous 12 months before the baseline visit, as
380 well as adjusting for Shannon diversity index showed largely similar effect estimates as the
381 main model (**Figure S5**). Applying ordinal mixed models showed consistent effect direction
382 and $P<0.05$ for 61 out of 64 species (**Table S4**). Overall, these findings indicate that the
383 associations identified in the main analysis were reliable.

384 **Gut Metabolic Modules Related to Dissimilatory Nitrate Reduction, Anaerobic Fatty**
385 **Acid Beta-oxidation and Amino Acid Degradation are Positively Associated with**
386 **Subclinical Atherosclerosis.**

387 In the basic model (n = 8973), 36 GMMs were positively associated while 4 had a negative
388 association with CACS (**Table S6**). In the main model (n=8118), 11 GMMs were found to be
389 positively associated with CACS, and none were negatively associated with CACS (**Table S7**).
390 None of these associations were driven by an influential observation. These 11 GMMs are
391 involved in various metabolic pathways, including amino acid degradation (e.g., asparagine,
392 proline, and valine degradation), anaerobic fatty acid beta-oxidation, enzymatic reactions on
393 lactate consumption, propionate production, and nitrogen metabolism.

394 **Other Atherosclerosis Measurements**

395 Next, we assessed the association between CACS-associated species and four CCTA
396 measurements (presence of coronary atherosclerosis plaque, coronary stenosis >50%, modified

397 Duke index, and SIS) and the number of carotid vessels with atherosclerotic plaque adjusted
398 for the covariates in the main model. Twenty-five CACS-associated species were also
399 associated with any coronary atherosclerotic plaque, 5 with coronary stenosis >50%, 39 with
400 the modified Duke index, 44 with the SIS, and 5 with carotid plaque (P<0.05), all in the same
401 direction as with CACS (**Figure 5**). The correlation among atherosclerosis measurements is
402 shown in **Figure S6**.

403 **CACS-Associated Species are Linked to Numerous Plasma Metabolites and Clustered**
404 **into Distinct Groups Based on their Strongest Associations**

405 We assessed whether the 60 CACS-associated species available in the GUTSY Atlas[21]
406 describing species-metabolite associations were associated with plasma metabolites. All 60
407 species had at least one association with a plasma metabolite and there were 2500 associations
408 (1478 positive and 1022 negative) between CACS-associated species and metabolites in total
409 (**Table S8**). Based on their strongest associations with metabolites the species clustered in
410 distinct groups, with one group containing species commonly found in the oral microbiome[24]
411 (e.g., all CACS-associated *Streptococcus* spp., *R. mucilaginosa, Bifidobacterium dentium*, and
412 *Ligilactobacillus salivarius*). Amongst others, this group was negatively associated with a set
413 of uncharacterized molecules, and with indole propionate and positively associated with
414 secondary bile acids and imidazole propionate. The remaining species were dominated by
415 *Eubacteriales* sp., and included both CACS-positive and negatively associated species and had
416 an inverse metabolic pattern compared to the oral microbiome set, with the exception of
417 tobacco-related metabolites, which were positively associated with CACS-positively
418 associated species and negatively associated with CACS-negatively associated species (**Figure**
419 **S7**).

420 **Oral Species in the Gut Microbiota are Associated with Markers of Systemic**
421 **Inflammation and Infection**

422 We assessed associations between CACS-associated species and plasma hsCRP (n=8117) and
423 blood counts of leukocytes (n=8098), and neutrophils (n=8092) in linear mixed models. Of 54
424 species positively associated with CACS, 13 were also positively associated with hsCRP, 10
425 with leukocyte counts, and 11 with neutrophil counts ($P<0.05$). Species associated with all
426 three markers were dominated by species common in the oral microbiome[24] including the
427 streptococci species most strongly associated with CACS: *S. anginosus*, *S. oralis* subsp. *oralis*,
428 and *S. parasanguinis.* Among the 13 species negatively associated with CACS, 6 were
429 negatively associated with hsCRP, 6 with leukocyte counts, and 5 with neutrophil counts
430 (**Figure 5**).

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

433 Several of the CACS-associated species are normally considered as oral species. [24] We
434 therefore investigated the correlation between the abundance of CACS-associated species in
435 the fecal and saliva samples from the 435 participants in MODS (**Table S9**) who underwent a
436 thorough dental examination within 4 to 12 months after fecal sampling in the MOS. Six of the
437 CACS-associated species were identified in the saliva samples, of which five (*S. anginosus,*
438 *S. parasanguinis*, *S. gordonii*, *R. mucilaginosa*, and *B. dentium*) were positively associated
439 with the abundance of the corresponding species in the fecal sample (rho=0.15–0.46), while
440 *S. oralis* subsp. *oralis* was not (**Table S10**). In up to 639 participants from MODS, three species
441 (*S. anginosus, S. parasanguinis,* and *R. mucilaginosa*) were positively associated with both
442 caries and filled surfaces, and three (*S. anginosus, S. gordonii and B. dentium*) with gingivitis
443 (**Table S11**).

444 **DISCUSSION**

445 Gut bacteria have been proposed to affect the development and progression of atherosclerosis
446 through infections local or distal to the atherosclerotic plaque or through production of
447 atherogenic metabolites.[2] The association of gut microbiota with coronary atherosclerosis has
448 previously been studied only in symptomatic patients, who are often under treatment, resulting
449 in high risk of bias as medication may result in large shifts of the composition. Here, detailed
450 image-based measurements of coronary artery atherosclerosis and deep characterization of the
451 gut microbiome by shotgun metagenomics of individuals without earlier cardiovascular disease
452 in the large population-based SCAPIS cohort were used to identify 64 species associated with
453 CACS independent of risk factors. Several of these species were associated with circulating
454 markers of inflammation and infection, and with the corresponding bacteria species in saliva
455 samples, which were in turn associated with worse oral health. Thus, studies are merited to
456 continue to investigate the role of these 64 species in atherogenesis.

457 The *Streptococcus* genus was associated with CACS, consistent with observations in
458 previous case-control studies of symptomatic atherosclerotic cardiovascular disease.[6,7,25]
459 Specifically, *S. anginosus*, *S. oralis* subsp. *oralis*, *S. parasanguinis*, *S. gordonii*, and
460 *S. agalactiae* were associated with coronary atherosclerosis. Current knowledge of the
461 involvement of these bacteria in cardiovascular pathologies is summarized in **Table S12**. These
462 species all belong to the VGS,[26] except for *S. agalactiae* (a β-hemolytic non-VGS). VGS can
463 enter the bloodstream through mucosal barrier injuries from daily dental care activities and
464 dental procedures, and could also pass the gut barrier when injured.[27] VGS can infect the valves
465 and the coronary vessels, accounting for 20% of infective endocarditis cases.[4,9,28] VGS species
466 initiate and contribute to biofilm formations, which enhance bacterial survival and have been
467 found in atherosclerotic lesions with unclear significance.[29] Moreover, in model systems, VGS
468 can invade human aortic endothelial cells and stimulate atherosclerosis-related pro-

469 inflammatory cytokines. [30] Animal studies suggest a causal link between *Streptococcus* spp.
470 and atherogenesis. [31,32] In our study, the abundance of CACS-associated *Streptococcus* spp. in
471 the gut associated strongly and positively with hsCRP, leukocytosis, and neutrophilia, which
472 could have been triggered by low-grade bacteremia.

473 The abundance of CACS-associated species might be reduced by antimicrobial treatment.
474 However, clinical trials showed that anti-infective therapies, targeted at chlamydial
475 colonization of the coronary plaques, seemed harmful in patients with established coronary
476 heart disease. [33] In the CLARICOR study, mortality was increased in patients with stable
477 coronary artery disease treated with clarithromycin. [34] Possible explanations to why these
478 antibiotics were ineffective in reducing coronary events include resistance to antibiotics due to
479 biofilm formation, [35] recolonization after antibiotic treatment, the need for treatment earlier in
480 atherosclerosis process, or simply that these bacteria are not causal to atherosclerosis. In the
481 current study, participants with low abundance of CACS-associated *Streptococcus* spp. were
482 not more exposed to antibiotics over the past 12 months than those with higher abundance. We
483 also found that the abundance of *Streptococcus* spp. was associated with coronary artery
484 calcification even after exclusion of participants treated with antibiotics.

485 We assessed the association of the CACS-associated species with other measures of
486 coronary atherosclerosis from the CCTA as well as with carotid atherosclerosis measured by
487 ultrasound. In general, we found high agreement of associations of species with CACS,
488 modified Duke index, any atherosclerosis, and SIS. However, the associations with having an
489 occlusion >50% of at least one vessel were weaker, perhaps because this was a rather rare
490 phenotype in our cohort, with low power. Likewise, only a few of the species were associated
491 with carotid plaque, perhaps because the measurement of carotid plaque was not very detailed
492 or because associations differ in different vascular beds.

493 The gut microbiota composition might also contribute to atherogenesis through alteration
494 of the host metabolism. In the current study, all CACS-associated species were associated with
495 at least one plasma metabolite and clustered in distinct groups based on their associations with
496 the metabolome. One group, which contained all the species commonly found in saliva, was
497 positively associated with several microbiota-derived metabolites such as primary and
498 secondary bile acids, and with omeprazole and metformin. Furthermore this group was
499 negatively associated with the microbially-derived tryptophan metabolite indole propionate, a
500 metabolite which has been found inversely associated with atherosclerotic coronary disease in
501 humans and reduced progression of atherosclerosis in mice. [36] Bacteria of this group were also
502 found positively associated with imidazole propionate, a microbially-derived metabolite from
503 histidine, reported to impair glucose metabolism. [37] Several tobacco metabolites, i.e. 3-
504 hydroxycotinine glucuronide, cotinine N-oxide and norcotinine, were positively associated
505 with species positively associated with CACS and negatively associated with species
506 negatively associated with CACS. Together, these findings suggest an interplay between
507 CACS-associated species and microbial metabolite production, drug intake and smoking
508 behavior.

509 Another aspect of the microbiome is the functional metabolic potential, which can be
510 summarized into metabolic modules based on known functions of bacterial genes. [20] We found
511 eleven modules positively associated with CACS, with the strongest associations noted for
512 dissimilatory nitrate reduction and anaerobic fatty acid beta-oxidation. Dietary nitrate is
513 absorbed in the intestines and excreted back into the gastrointestinal tract by the salivary glands
514 at high concentrations, forming the so-called entero-salivary circulation of nitrate. Denitrifying
515 oral bacteria are important for converting nitrate (NO_3^-) into nitrite (NO_2^-), which can be
516 further converted into nitric oxide (NO) in the acidic pH of the stomach. The generation of NO
517 contributes to the gastric protection by increasing blood flow and mucus thickness. [38] The

518 nitrate and nitrite that is not converted to NO in the stomach is absorbed into the bloodstream
519 and tissues, where they might have act as reservoir for NO, which appear to be important for
520 vasodilation under hypoxia[39] and modulation of mitochondrial respiration. [40] However, the
521 dissimilatory nitrate reduction pathway converts nitrate to ammonia, and having an increased
522 activity of this pathway in the colon could potentially inhibit the potential positive
523 cardiovascular effects of nitrate. We found that genes related to anaerobic fermentation of fatty
524 acids were associated with CACS. Such genes are rather uncommon in the gut microbiota and
525 the genetic background of oxidation of fatty acids under anaerobic conditions is not well
526 characterized. We found an association of five amino acid breakdown modules with CACS.
527 This could be driven by differences in diet not captured by our covariates.

528 Our study has some limitations. First, although our cohort is at least seven times larger than
529 those analyzed previously, few participants had high levels of subclinical atherosclerosis,
530 reducing statistical power. Second, microbial composition can vary extensively throughout the
531 gastrointestinal tract. Fecal samples contain microbial populations from the distal colon, and
532 less so from other sites such as the small intestine. Therefore, we could not identify associations
533 of species that are not well represented in fecal samples. Third, our study does not take into
534 consideration the different interactions among bacterial species such as synergistic effects in
535 the relationship with coronary atherosclerosis. Fourth, the cross-sectional study design limits
536 causal inference. Finally, our data on antibiotic treatment does not capture antibiotics provided
537 in inpatient care; however, in this age group, we do not expect a large number of hospital-
538 treated patients. In future studies, different causal inference methods should be used to
539 determine whether the identified species are causally related to atherosclerosis development.

540 In conclusion, by combining data from a large population-based cohort study and imaging
541 to evaluate subclinical coronary atherosclerosis, we found that the abundance of several species
542 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 [ERC-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.), 2022-01460 (SA) 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 Foundation [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 Swedish Research Council [VR, 521-2013-2756 (P.M.N.)], the Swedish Heart and Lung Foundation [Hjärt-Lungfonden 20150427 (P.M.N.)], and by A.L.F. 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 in project sens2019512 provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) and the Swedish National Infrastructure for Computing (SNIC) at Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) partially funded by the Swedish Research Council through grant agreements no. 2022-06725 and no. 2018-05973.

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-S7

590 **Table 1** Descriptive characteristics of participants in the SCAPIS study with metagenomics

591 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
Past year narrow spectrum antibiotics treatment, n (%)	579 (12.8%)	550 (12.4%)	8973

Past year broad spectrum antibiotics treatment, n (%)	505 (11.1%)	400 (9.0%)	8973

592 *Mean (standard deviation)

593 †Median [interquartile range]

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

Figures

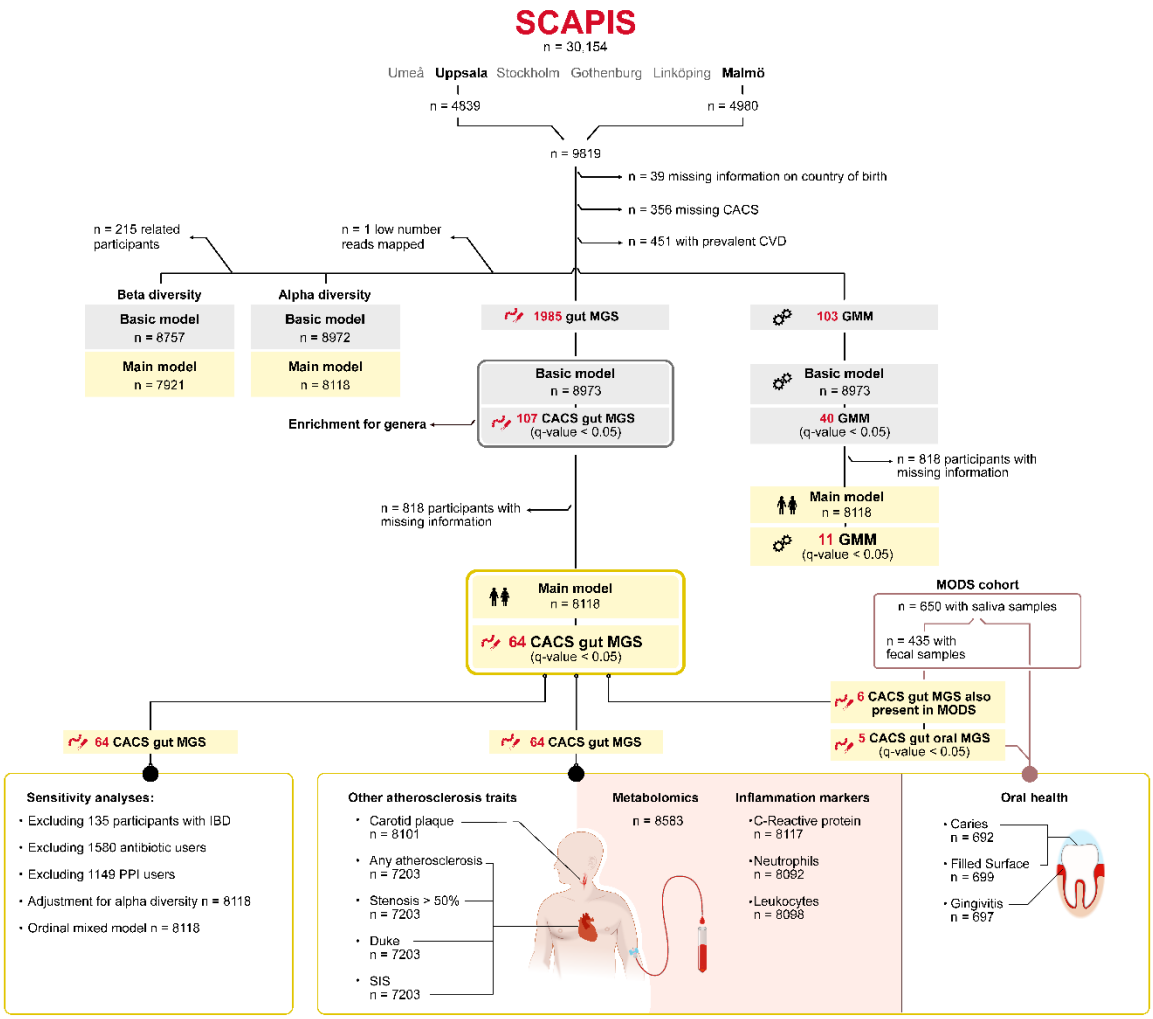


Figure 1. Flowchart of the overall design of the study.

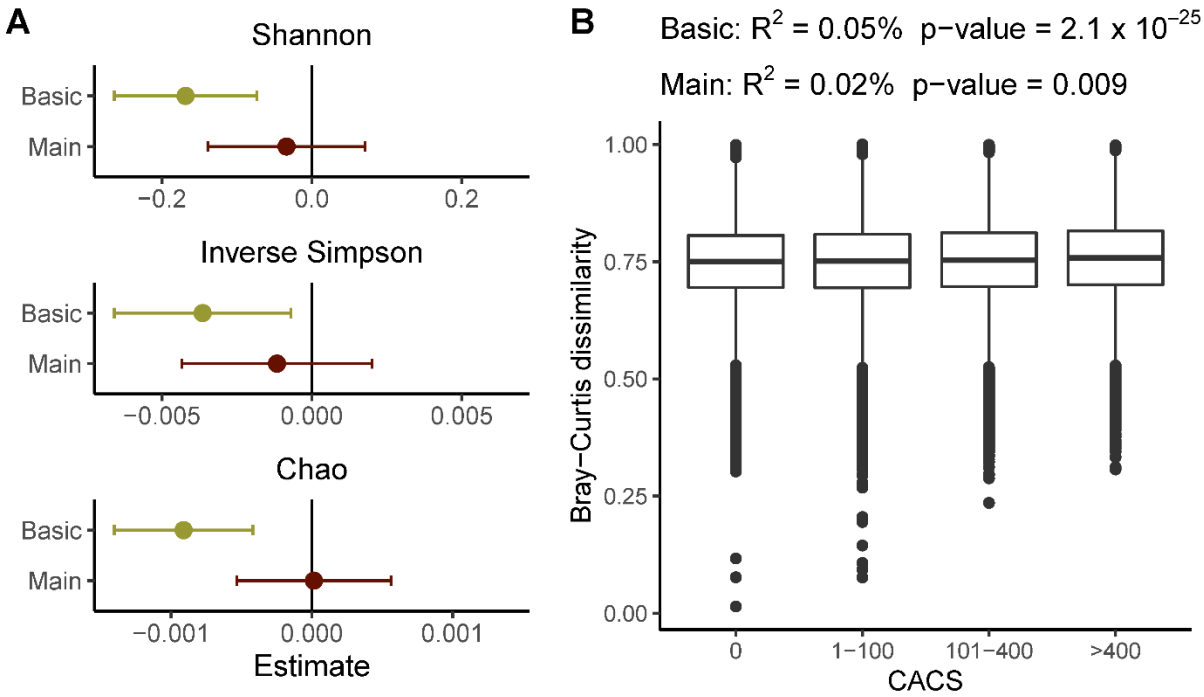


Figure 2. Association of alpha and beta diversity with CACS.

The association of alpha diversity (a) with coronary atherosclerosis was assessed by fitting linear mixed models with CACS as the outcome and the alpha diversity indices as the exposure. The basic model (n=8972) was adjusted for age, sex, country of birth, and metagenomics extraction plate as fixed effects and family relatedness as random effect. The main model (n=8118) was further adjusted for, smoking, physical activity, energy-adjusted intake of carbohydrate, protein, and fiber, systolic blood pressure (SBP), and diastolic blood pressure (DBP), total cholesterol, HDL and LDL cholesterol, triglycerides, BMI, diabetes, and self-reported medication for dyslipidemia, hypertension, and diabetes as fixed effects. For beta diversity (b), a distance-based multivariate analysis of variance of CACS as the independent variable was applied and the same fixed effect adjustments were made without the random effect adjustment. In the basic (n=8757) and the full model (n=7921) one participants for each family relatedness cluster was removed. Box plots represent the Bray-Curtis dissimilarity between participants with CACS=0 and the other CACS groups. For the CACS=0, the boxplot represents the within group dissimilarity.

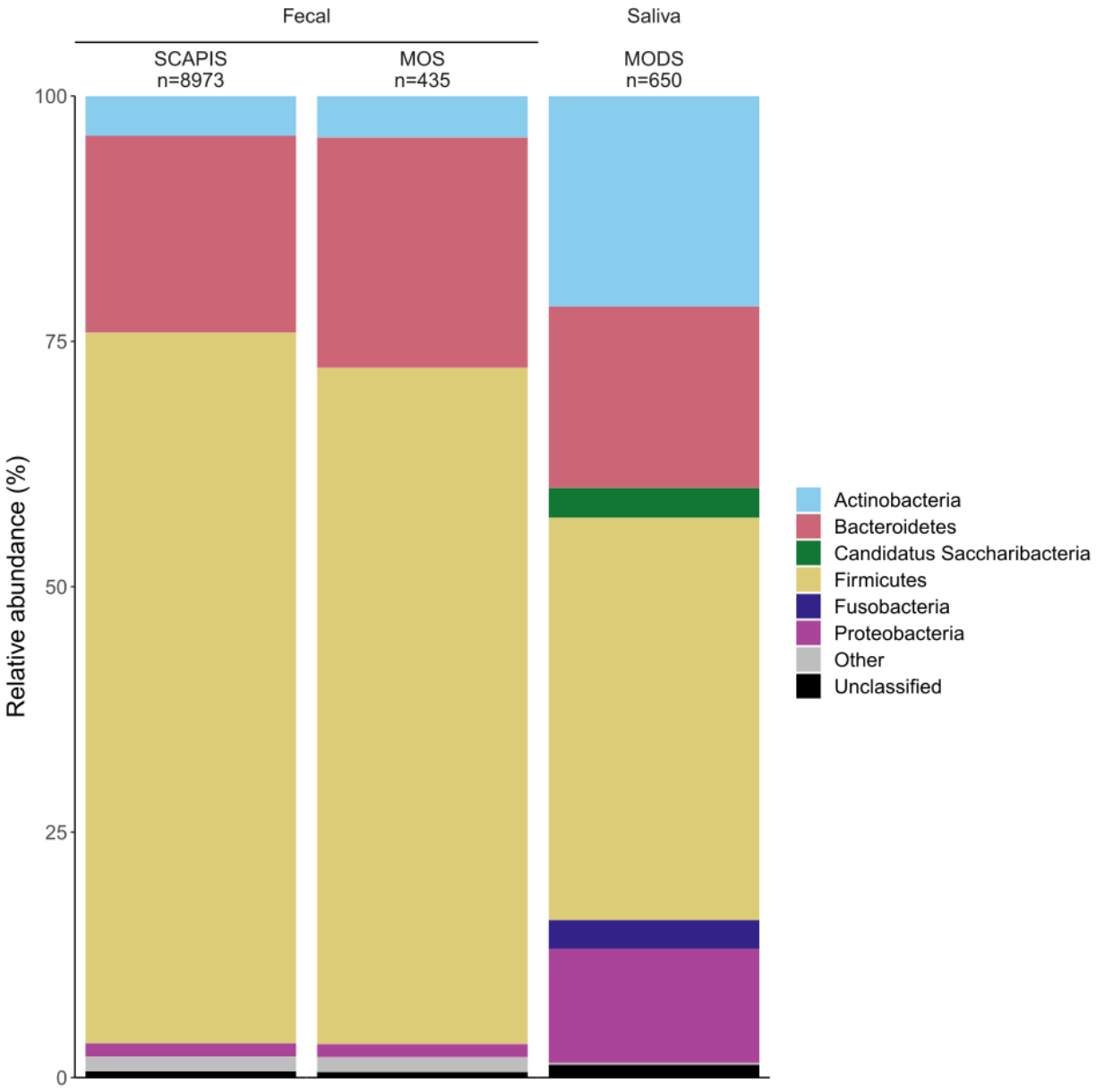


Figure 3. Phylum-level taxonomical profiling of the fecal and saliva samples across the three data sets.

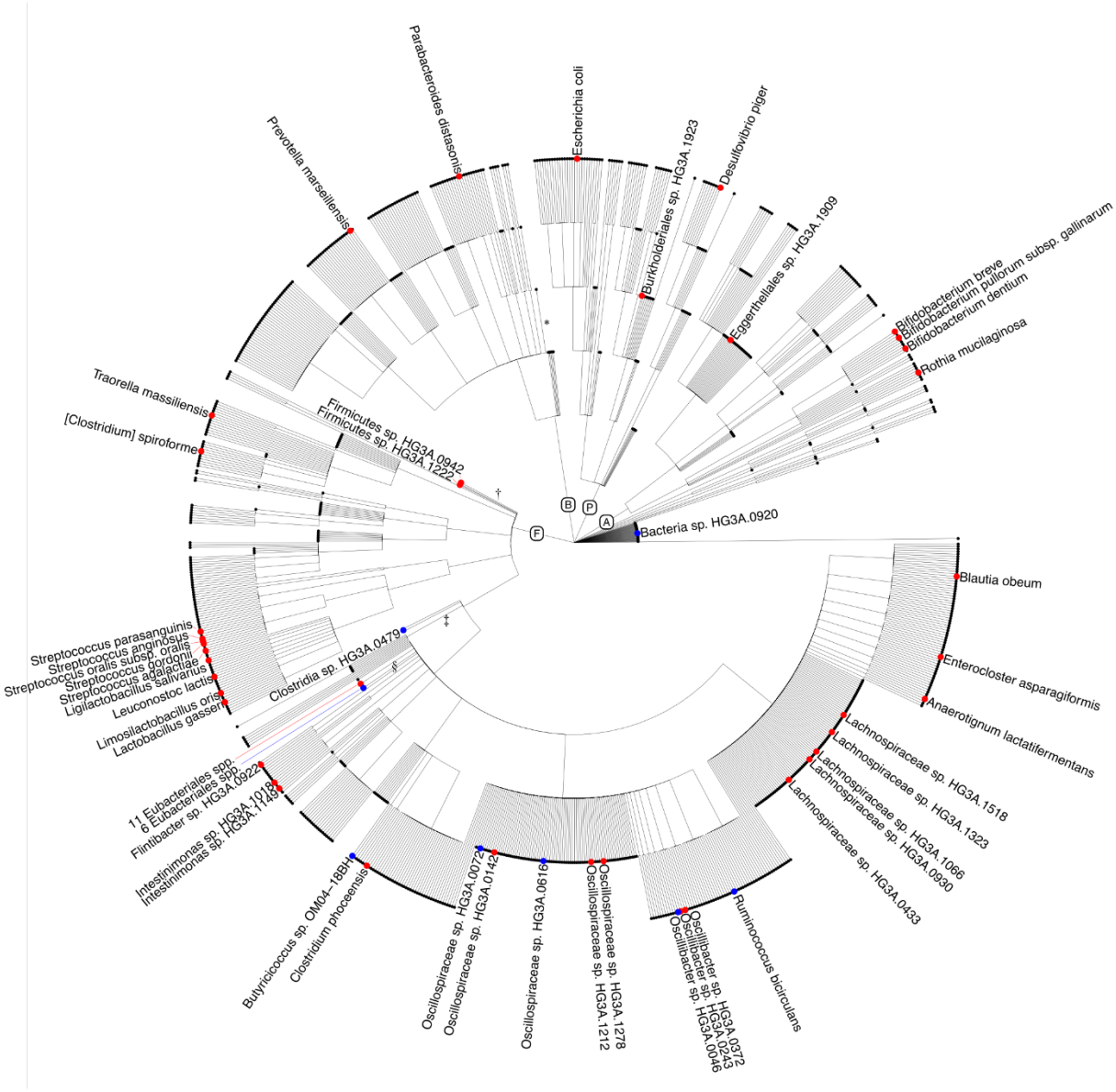
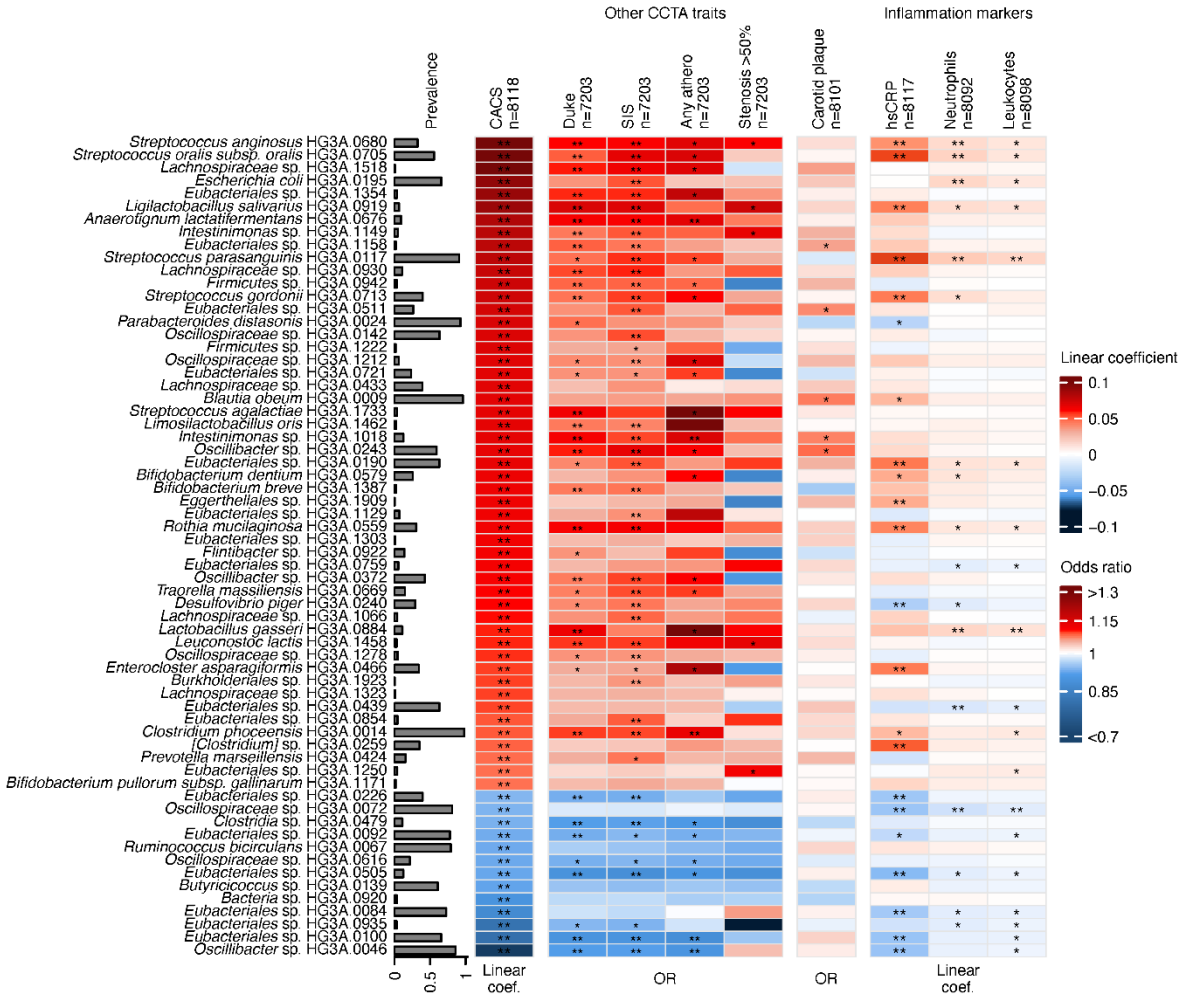


Figure 4. Cladogram of species investigated in the study.

Circles represent metagenomic species annotated to the lowest taxonomic level. Dots indicate the species positively (red) and negatively (blue) associated with CACS in the main 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 §651 species with lowest annotation to the order *Eubacteriales*.



629

630 **Figure 5. Associations between CACS-related gut species and alternate measurements of**
631 **atherosclerosis and markers of inflammation and infection.** Heatmap showing the
632 associations of the 64 coronary artery calcium score (CACS)-associated species (False
633 Discovery Rate (FDR) q-value<0.05) with CACS, other coronary computed tomography
634 angiography atherosclerosis traits, carotid plaque, and inflammation biomarkers. Associations
635 with CACS and with inflammatory markers hsCRP (high-sensitivity C-reactive protein),
636 neutrophil and leukocyte counts are presented as linear regression coefficients after adjustment
637 for the main model covariates (age, sex, country of birth, metagenomics extraction plate,
638 smoking, physical activity, energy-adjusted carbohydrate, protein and fiber intake, systolic and

639 diastolic blood pressure, high-density lipoprotein (HDL), low-density lipoprotein (LDL), and
640 total cholesterol, triglycerides, body mass index (BMI), diabetes, and self-reported medication
641 for dyslipidemia, hypertension, and diabetes as fixed effects and first-degree family relatedness
642 as a random effect). Associations with atherosclerosis traits (i.e., modified Duke index,
643 segment involvement score (SIS), any atherosclerosis, and ≥50% stenosis) and presence of
644 carotid plaques are presented as odds ratio (OR) after adjustment for main model covariates.
645 Associations with FDR q-value <0.05 are marked with two asterisks (**) and those with P<0.05
646 are marked with one asterisk (*).

REFERENCES

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton AZ, Benjamin EJ, Benziger CP, et al. 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. doi: https://doi.org/10.1016/j.jacc.2020.11.010
2. Jonsson AL, Bäckhed F. Role of gut microbiota in atherosclerosis. *Nat Rev Cardiol*. 2017;14:79-87. doi: 10.1038/nrcardio.2016.183
3. Chakaroun RM, Olsson LM, Backhed F. The potential of tailoring the gut microbiome to prevent and treat cardiometabolic disease. *Nat Rev Cardiol*. 2023;20:217-235. doi: 10.1038/s41569-022-00771-0
4. Ott SJ, Mokhtari NEE, Musfeldt M, Hellmig S, Freitag S, Rehman A, Kühbacher T, Nikolaus S, Namsolleck P, Blaut M, et al. Detection of diverse bacterial signatures in atherosclerotic lesions of patients with coronary heart disease. *Circulation*. 2006;113:929-937. doi: 10.1161/CIRCULATIONAHA.105.579979
5. Fromentin S, Forslund SK, Chechi K, Aron-Wisnewsky J, Chakaroun R, Nielsen T, Tremaroli V, Ji B, Prifti E, Myridakis A, et al. Microbiome and metabolome features of the cardiometabolic disease spectrum. *Nature Medicine 2022*. 2022;37:1-12. doi: 10.1038/s41591-022-01688-4
6. Jie Z, Xia H, Zhong SL, Feng Q, Li S, Liang S, Zhong H, Liu Z, Gao Y, Zhao H, et al. The gut microbiome in atherosclerotic cardiovascular disease. *Nature Communications*. 2017;8:845. doi: 10.1038/s41467-017-00900-1
7. Liu H, Chen X, Hu X, Niu H, Tian R, Wang H, Pang H, Jiang L, Qiu B, Chen X, et al. Alterations in the gut microbiome and metabolism with coronary artery disease severity. *Microbiome*. 2019;7:68. doi: 10.1186/s40168-019-0683-9
8. Schmidt TS, Hayward MR, Coelho LP, Li SS, Costea PI, Voigt AY, Wirbel J, Maistrenko OM, Alves RJ, Bergsten E, et al. Extensive transmission of microbes along the gastrointestinal tract. *Elife*. 2019;8. doi: 10.7554/eLife.42693
9. Koren O, Spor A, Felin J, Fåk F, Stombaugh J, Tremaroli V, Behre CJ, Knight R, Fagerberg B, Ley RE, et al. 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. doi: 10.1073/pnas.1011383107
10. Jönsson D, Orho-Melander M, Demmer RT, Engström G, Melander O, Klinge B, Nilsson PM. Periodontal disease is associated with carotid plaque area: the Malmö Offspring Dental Study (MODS). *Journal of Internal Medicine*. 2020;287:301-309. doi: 10.1111/JOIM.12998
11. Ryden L, Buhlin K, Ekstrand E, de Faire U, Gustafsson A, Holmer J, Kjellstrom B, Lindahl B, Norhammar A, Nygren A, et al. Periodontitis Increases the Risk of a First Myocardial Infarction: A Report From the PAROKRANK Study. *Circulation*. 2016;133:576-583. doi: 10.1161/CIRCULATIONAHA.115.020324
12. Tonetti MS, D'Aiuto F, Nibali L, Donald A, Storry C, Parkar M, Suvan J, Hingorani AD, Vallance P, Deanfield J. Treatment of periodontitis and endothelial function. *N Engl J Med*. 2007;356:911-920. doi: 10.1056/NEJMoa063186
13. Bergström G, Berglund G, Blomberg A, Brandberg J, Engström G, Engvall J, Eriksson M, Faire Ud, Flinck A, Hansson MG, et al. The Swedish CArdioPulmonary BioImage Study: Objectives and design. *Journal of Internal Medicine*. 2015;278:645-659. doi: 10.1111/joim.12384
14. Brunkwall L, Jönsson D, Ericson U, Hellstrand S, Kennbäck C, Östling G, Jujic A, Melander O, Engström G, Nilsson J, et al. The Malmö Offspring Study (MOS): design, methods and first results. *European Journal of Epidemiology*. 2021;36:103-116. doi: 10.1007/s10654-020-00695-4
15. Bergström G, Persson M, Adiels M, Björnson E, Bonander C, Ahlström H, Alfredsson J, Angerås O, Berglund G, Blomberg A, et al. Prevalence of Subclinical Coronary Artery

Atherosclerosis in the General Population. *Circulation*. 2021;144:916-929. doi: 10.1161/CIRCULATIONAHA.121.055340

16. Ohnesorge B, Flohr T, Fischbach R, Kopp AF, Knez A, Schroder S, Schopf UJ, Crispin A, Klotz E, Reiser MF, et al. Reproducibility of coronary calcium quantification in repeat examinations with retrospectively ECG-gated multisection spiral CT. *Eur Radiol*. 2002;12:1532-1540. doi: 10.1007/s00330-002-1394-2

17. Touboul PJ, Hennerici MG, Meairs S, Adams H, Amarenco P, Bornstein N, Csiba L, Desvarieux M, Ebrahim S, Hernandez RH, et al. 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. doi: 10.1159/000343145

18. Nielsen HB, Almeida M, Juncker AS, Rasmussen S, Li J, Sunagawa S, Plichta DR, Gautier L, Pedersen AG, Chatelier EL, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature Biotechnology*. 2014;32:822-828. doi: 10.1038/nbt.2939

19. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic acids research*. 2016;44:D733-D745. doi: 10.1093/NAR/GKV1189

20. Vieira-Silva S, Falony G, Darzi Y, Lima-Mendez G, Garcia Yunta R, Okuda S, Vandeputte D, Valles-Colomer M, Hildebrand F, Chaffron S, et al. Species–function relationships shape ecological properties of the human gut microbiome. *Nature Microbiology*. 2016;1:16088. doi: 10.1038/nmicrobiol.2016.88

21. Dekkers KF, Sayols-Baixeras S, Baldanzi G, Nowak C, Hammar U, Nguyen D, Varotsis G, Brunkwall L, Nielsen N, Eklund AC, et al. An online atlas of human plasma metabolite signatures of gut microbiome composition. *Nature Communications*. 2022;13:5370. doi: 10.1038/s41467-022-33050-0

22. Benjamini Y, 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. doi: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

23. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. *bioRxiv*. 2021:060012. doi: 10.1101/060012

24. Caselli E, Fabbri C, D'Accolti M, Soffritti I, Bassi C, Mazzacane S, Franchi M. Defining the oral microbiome by whole-genome sequencing and resistome analysis: the complexity of the healthy picture. *BMC Microbiol*. 2020;20:120. doi: 10.1186/s12866-020-01801-y

25. Liu S, Zhao W, Liu X, Cheng L. Metagenomic analysis of the gut microbiome in atherosclerosis patients identify cross-cohort microbial signatures and potential therapeutic target. *FASEB Journal*. 2020;34:14166-14181. doi: 10.1096/fj.202000622R

26. Doern CD, 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. doi: 10.1128/JCM.01563-10

27. Lockhart PB, Brennan MT, Sasser HC, Fox PC, Paster BJ, Bahrani-Mougeot FK. Bacteremia associated with toothbrushing and dental extraction. *Circulation*. 2008;117:3118-3125. doi: 10.1161/CIRCULATIONAHA.107.758524

28. Viehman JA, Smith BJ, Nanjappa S, Li SK, Perez CO, Fernandes CR. Chapter 3 - Microbiology of endocarditis. In: Kilic A, ed. *Infective Endocarditis*. Academic Press; 2022:43-59.

29. Lanter BB, 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. doi: 10.1128/IAI.00510-15/SUPPL_FILE/ZII999091420SO1.PDF

30. Nagata E, Toledo AD, 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. doi: 10.1111/J.2041-1014.2010.00597.X

31. Hashizume-Takizawa T, Yamaguchi Y, Kobayashi R, Shinozaki-Kuwahara N, Saito M, 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. doi: 10.1016/j.bbrc.2019.10.057

32. Kesavalu L, Lucas AR, Verma RK, Liu L, Dai E, Sampson E, Progulske-Fox A. Increased atherogenesis during Streptococcus mutans infection in ApoE-null mice. *Journal of Dental Research*. 2012;91:255-260. doi: 10.1177/0022034511435101

33. Sethi NJ, Safi S, Korang SK, Hróbjartsson A. Antibiotics for secondary prevention of coronary heart disease. *Cochrane Database Syst Rev*. 2019.

34. Jespersen CM. Randomised placebo controlled multicentre trial to assess short term clarithromycin for patients with stable coronary heart disease: CLARICOR trial. *British Medical Journal*. 2006;332:22-24. doi: 10.1136/bmj.38666.653600.55

35. Sharma D, Misba L, Khan AU. Antibiotics versus biofilm: an emerging battleground in microbial communities. *Antimicrobial Resistance & Infection Control*. 2019;8:76. doi: 10.1186/s13756-019-0533-3

36. Xue H, Chen X, Yu C, Deng Y, Zhang Y, Chen S, Chen X, Chen K, Yang Y, Ling W. Gut Microbially Produced Indole-3-Propionic Acid Inhibits Atherosclerosis by Promoting Reverse Cholesterol Transport and Its Deficiency Is Causally Related to Atherosclerotic Cardiovascular Disease. *Circ Res*. 2022;131:404-420. doi: 10.1161/CIRCRESAHA.122.321253

37. Koh A, Molinaro A, Stahlman M, Khan MT, Schmidt C, Manneras-Holm L, Wu H, Carreras A, Jeong H, Olofsson LE, et al. Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell*. 2018;175:947-961 e917. doi: 10.1016/j.cell.2018.09.055

38. Bjorne HH, Petersson J, Phillipson M, Weitzberg E, Holm L, Lundberg JO. Nitrite in saliva increases gastric mucosal blood flow and mucus thickness. *J Clin Invest*. 2004;113:106-114. doi: 10.1172/JCI19019

39. Cosby K, Partovi KS, Crawford JH, Patel RP, Reiter CD, Martyr S, Yang BK, Waclawiw MA, Zalos G, Xu X, et al. Nitrite reduction to nitric oxide by deoxyhemoglobin vasodilates the human circulation. *Nat Med*. 2003;9:1498-1505. doi: 10.1038/nm954

40. Shiva S, Huang Z, Grubina R, Sun J, Ringwood LA, MacArthur PH, Xu X, Murphy E, Darley-Usmar VM, Gladwin MT. Deoxymyoglobin is a nitrite reductase that generates nitric oxide and regulates mitochondrial respiration. *Circ Res*. 2007;100:654-661. doi: 10.1161/01.RES.0000260171.52224.6b

41. Christensen SE, Moller E, Bonn SE, Ploner A, Wright A, Sjolander A, Balter O, Lissner L, 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. doi: 10.2196/jmir.2458

42. Silness J, Loe H. Periodontal Disease in Pregnancy. Ii. Correlation between Oral Hygiene and Periodontal Condtion. *Acta Odontol Scand*. 1964;22:121-135. doi: 10.3109/00016356408993968

43. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods 2012 9:4*. 2012;9:357-359. doi: 10.1038/nmeth.1923

44. Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P, et al. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell*. 2019;176:649-662.e620. doi: 10.1016/j.cell.2019.01.001

45. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, 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. doi: 10.1016/J.YMETH.2016.02.020

46. Steinegger M, Söding J. Clustering huge protein sequence sets in linear time. *Nature communications*. 2018;9:2542. doi: 10.1038/S41467-018-04964-5

47. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, et al. 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. doi: 10.1093/NAR/GKY1085

48. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, 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. doi: 10.1093/MOLBEV/MSAB293