

Sensitivity analysis for factors that may affect the species-coronary artery calcium score (CACS) relationship. To ensure the statistical robustness of our findings, partial Spearman correlation were used as an alternative robust method for the identified species in the main primary hypothesis analysis. Partial Spearman correlation models revealed that 62 species remained significantly associated with CACS in the basic model and 31 in the full model (**Supplementary Tables 6 and 7**). The 31 species that did not replicate in the full model at 5% false discovery rate in the Spearman correlation could be explained due to a reduced statistical power, or due to highly influential observations in the linear regression, or because the true association was non-linear. Therefore, we investigated whether this discrepancy was due to influential observations by calculating unscaled dfbeta values, which measures the difference in each parameter estimated with and without the influential observation. If the most influential observations had high dfbeta values compared to the regression coefficient and the values were in the same direction as the regression coefficient, the finding from the linear regression was deemed unreliable. None of the 31 species was indeed the case and after removing the observation with a higher Cook's distance the model still was significant. Next, for all the 31 associations we tested whether they can be explained by non-linear associations. To carry out this verification the residuals from the linear regression were plotted against the species (the exposure in the model) to notice whether there is a trend. None of them were neither explained by non-linear associations. Hence, the findings were considered robustly associated with CACS. We then tested if our findings could be explained due to the addition of covariates. The addition of body mass index and proton-pump inhibitors-use as co-variates into the full model reduced the number of CACS associated species to 57, but the effect estimates remained similar across the species (**Extended Data Fig. 4 and Supplementary Table 7**). Further, inclusion of established cardiovascular risk factors (body mass index, diabetes, systolic blood pressure, diastolic blood pressure, total cholesterol, cholesterol in low-density lipoproteins, and cholesterol in high-density lipoproteins) in the full model revealed that 51 species were associated with CACS independently of these factors (**Extended Data Fig.4 and Supplementary Table 7**). We reanalyzed the findings without adjusting for Shannon diversity index and the results were similar (**Extended Data Fig.4, Supplementary Table 6 and 7**). We then excluded specific groups of individuals who received antibiotic treatment or self-reported

gastrointestinal disease. After excluding 1,797 participants who had received antibiotic treatment during the year preceding the baseline visit, 65 species were associated with CACS in the basic model and 52 species in the full model (**Extended Data Fig. 4 Supplementary Table 6 and 7**). The effect estimates remained very similar indicating that the loss of significance was due to the smaller sample size and not confounding. The exclusion of 96 participants with Crohn's disease or ulcerative colitis generally did not impact the analysis results, except for the association between *Eubacteriales* sp. (HG3A.0373) and CACS which was no longer associated at 5% false discovery rate in the full model (**Extended Data Fig. 4, Supplementary Table 6 and 7**).