A Health-Conscious Food Pattern Is Associated with Prediabetes and Gut Microbiota in the Malmö Offspring Study

Ulrika Ericson, Louise Brunkwall, Sophie Hellstrand, Peter M Nilsson, and Marju Orho-Melander

1Department of Clinical Sciences, Diabetes and Cardiovascular Disease, Genetic Epidemiology, Lund University, Malmö, Sweden; and 2Department of Internal Medicine, Skåne University Hospital, Lund University, Malmö, Sweden

ABSTRACT

Background: Diet is a determinant of gut microbiota. Both diet and gut microbiota have been linked to metabolic diseases.

Objective: We aimed to examine data-driven food patterns in relation to the prevalence of prediabetes and gut microbiota composition and food pattern–associated bacteria in relation to prediabetes.

Methods: Food patterns were extracted using principal component analysis in 1726 individuals (aged 18–71 y, 55% women, mean BMI = 25.5 kg/m²) without diabetes from the population-based Malmö Offspring Study. The gut (fecal) microbiota was analyzed by sequencing the 16S ribosomal RNA gene (V1–V3 region). Prediabetes classification was based on fasting glucose ≥6.0 mmol/L and/or glycated hemoglobin ≥42 mmol/L at baseline and/or type 2 diabetes diagnosis during follow-up (0–3.8 y). Logistic regression was used to investigate cross-sectional associations with prediabetes, and the general linear model to examine associations between food patterns and bacterial genera.

Results: Two food patterns, the Health-conscious and the Sugar and High-Fat Dairy patterns, were identified. Adherence to the Health-conscious pattern was associated with a lower prevalence of prediabetes (OR comparing highest quintile with lowest: 0.54; 95% CI: 0.32, 0.92; P-trend = 0.03) and with the abundance of several gut bacterial genera, of which the most robust findings were with a higher abundance of Roseburia and Lachnospira and with a lower abundance of Eubacterium. Roseburia was also associated with a lower prevalence of prediabetes (OR comparing highest quintile with lowest: 0.56; 95% CI: 0.35, 0.92; P-trend = 0.01) and the association between the Health-conscious pattern and prediabetes was attenuated after adjustment for abundance of Roseburia and BMI. Adherence to the Sugar and High-Fat Dairy pattern was associated with a higher prevalence of prediabetes in women (P-trend across food pattern quintiles = 0.03).

Conclusions: In this Swedish population-based study, a Health-conscious food pattern showed an inverse association with the prevalence of prediabetes. Potential underlying explanations may involve links between healthy diet and BMI, as well as gut microbiota, especially a higher abundance of Roseburia. J Nutr 2020;150:861–872.

Keywords: food intake, food patterns, epidemiology, type 2 diabetes, gut microbiota

Introduction

Type 2 diabetes (T2D) affects the quality of life of many individuals worldwide and the high and increasing prevalence is an enormous burden to society. Identification of modifiable lifestyle factors including diet, that have beneficial effects on glucose metabolism and counteract the development of T2D is of great value to public health. However, diet is a complex exposure of interacting food components, consumed in different combinations. It is therefore important to capture overall healthy food patterns that could be translated into dietary guidelines. Many data-driven food patterns referred to as “Healthy” or “Prudent” have been associated with decreased risk of T2D (1–3). Diet is also an important factor for gut microbiota composition and richness (4), which have been found to be altered in many disease states including obesity and T2D (5–7), and a recent study indicated that the gut microbiota may play a causal role in the development of T2D via the effects on SCFA production from fermentable food components (8). Studies on overall dietary patterns in relation to gut microbiota are scarce and have mainly focused on associations with the Mediterranean dietary patterns or plant-based versus animal-based diets (9–13). Although diet has been identified as 1 of
the main determinants of gut microbiota composition, large observational studies are lacking regarding foods commonly consumed in modern societies and no study has examined data-driven food patterns in relation to gut bacterial composition. Moreover, the importance of gut microbiota in health and disease in a large number of individuals, or in population-based cohorts, have only recently been studied (4, 14).

Our aim was to identify data-driven food patterns using principal component analysis in 1726 men and women from the Malmö Offspring Study (MOS) without a previous diabetes diagnosis, and to examine whether the extracted food patterns were associated with prevalence of prediabetes at baseline, independently of other lifestyle factors. In addition, we wanted to examine the identified food patterns in relation to gut microbiota composition. Finally, the food pattern-related gut bacteria were examined in relation to prediabetes.

**Methods**

**Study population and data collection**

MOS is an ongoing population-based cohort study where children and grandchildren (aged >18 y) of participants in the Malmo Diet and Cancer Study—Cardiovascular Cohort are recruited (15, 16). Participants were invited via letter and visited the research clinic on 2 occasions with about a week in between the visits. There were no exclusion criteria. At the first visit, venous blood was drawn after an overnight fast, and anthropometrics and blood pressure were measured. The study participants were instructed on how to collect the fecal samples at home, how to record their food intake during 4 d (starting the day after the first visit), and how to fill in a web-based food propensity questionnaire and a comprehensive questionnaire on other lifestyle and socioeconomic factors (before the second visit). At the second visit the fecal samples were brought to the clinic. The study protocols were approved by the Ethics Committee of Lund University (protocol number DNR 2012/594) and all participants provided written informed consent.

From the start of the study in March 2013 until the end of April 2017, 2644 individuals participated in baseline examinations (47% of the eligible participants) (Supplemental Figure 1). Of the nonparticipants, 29% answered that they were not willing to take part; 28% did not reply, 26% did not come to their appointment, 5% died or moved before they had received an invitation, 4% wanted to participate later, 5% stated lack of time due to work or other activities, 1% were ill, and 2% had other reasons. In total, 1788 participants contributed dietary data by completing a web-based food record. Out of those, we excluded 62 individuals with prevalent diabetes (according to information from national and local registries and questionnaire data at baseline), leaving 1726 (55% women) individuals for the present study. For gut microbiota analyses, 1477 individuals with gut microbiota data were included. As use of antibiotics and probiotics are known to alter the gut microbiota composition, we performed sensitivity analysis on the gut microbiota in a reduced sample (n = 851) after the exclusion of: 1) individuals reporting to have used antibiotics during the 6 mo previous to data collection (n = 169) or with missing data on use of antibiotics (n = 177), and/or 2) individuals reporting probiotic use >3 times per week (n = 77) or with missing data on the use of probiotics (n = 363).

**Clinical measurements**

Height (m), without shoes and hats, was measured to the nearest centimeter directly under the meter with the legs together looking straight ahead. Weight (kg) was measured in light clothing on a calibrated balance beam or digital scale. BMI (kg/m²) was calculated from measurements of weight and height. Resting blood pressure (mmHg) was measured as a mean of 2 readings in the supine position after 10 min rest by use of an automatic device (Omron). Blood samples were analyzed at the Department of Clinical Chemistry, Malmo. Whole blood glycated hemoglobin (HbA1c) concentrations were measured within 4 h using the Capillaries 3 Tera HbA1c kit (Sebia). Plasma glucose concentrations were measured directly using the HemoCue Glucose 201+ System (HemoCue AB). Total cholesterol, triglycerides, and HDL cholesterol plasma concentrations were measured within 4 h by enzymatic methods using the COBAS system (Roche Diagnostics). LDL cholesterol was calculated using the Friedewald equation.

**Dietary data**

Diet was assessed using a 4-d web-based food record, the Riksmaten2010, developed by the Swedish National Food Institute (17). The relative validity of the Riksmaten2010 method has been evaluated by comparing the reported energy intake to objectively measured total energy expenditure with the doubly labelled water technique (τ = 0.40) (17). Estimated intakes of fiber sources with the Riksmaten2010 method have been compared with objective plasma biomarkers; the Pearson correlation coefficients for fruit and vegetable intake were 0.46 and 0.20, and for whole grain intake 0.30 and 0.29, in women and men, respectively (18). In addition, repeated 4-d records have been collected for 323 individuals indicating reliable data with Pearson correlation coefficients of 0.55, 0.40, 0.56, and 0.61 for energy-adjusted intakes of carbohydrates, fat, protein, and fiber, respectively (unpublished data).

The mean daily food intakes were calculated from frequency and portion estimates from the food records. Food intakes were converted into energy and nutrient intakes using the national food database; Riksmaten vuxna 2010, version 10–05-05. Food intakes were aggregated into 43 food groups considered to represent overall dietary intake. Our aim was to cover as many parts of the overall diet as possible, but to avoid an overly detailed level on foods consumed irregularly, which may not be satisfactorily captured on a 4-d basis. This was, for example, the reason that the intake of lean and fatty fish were grouped together. Characteristics related to both dietary behaviors and nutrient content were considered when aggregating the foods. In order to minimize the effects of misreporting, energy-adjusted intakes of the food groups were calculated by regressing the intakes on nonalcohol energy intake using the residual method (19).

**Other variables**

Lifestyle variables were based on answers from a web-based self-administered questionnaire on lifestyle and socioeconomic factors. Education was based on the participants highest level of completed education defined as primary (< 9 y), secondary (9 y), upper secondary (12 y), and a university or college degree. Smoking status was defined as never-smoker, ex-smoker, irregular smoker, and regular smoker (based on the participant's own definition of regular smoking). Total physical activity level was estimated by summing 4-grade scale answers regarding physical activity at work (very light = 1, light = 2, moderately heavy = 3, and heavy/very heavy = 4) and leisure time (sedentary = 1, moderate activity = 2, moderate and regular activity = 3, and regular training = 4) into a 7-grade scale (2–8 points, e.g. 2 = 1 + 1 for very light physical activity at work and sedentary leisure time). Alcohol consumption was defined by a 7-category variable. Participants

Supported by the European Research Council ERC-CoG-2014-649021 (MO-MI), The European Foundation for the Study of Diabetes/Lilly Award 2014, the Swedish Research Council (MO-M and PMN), the Heart and Lung Foundation, the Swedish Diabetes Foundation, the Region Skåne, the Skåne University Hospital, the Novo Nordic Foundation, the Albert Påhlsson Research Foundation, the Swedish Research Council (for Strategic Research Area Exobiol, Dnr 2009-1039), the Swedish Foundation for Strategic Research (Dnr IRC15-0067), and the Swedish Research Council Linnaeus grant, Dnr 266/2006-237).

**Author disclosures:** The authors report no conflicts of interest.

Supplemental Tables 1–3 and Supplemental Figures 1 and 2 are available from the “Supplementary data” link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/jn/.

 UE and LB contributed equally to this study.

Additions to correspondence to UE (e-mail: eriksa.med.lu.se).

Abbreviations used: HbA1c, glycated hemoglobin; MOS, Malmö Offspring Study; OTU, operational taxonomical unit; T2D, type 2 diabetes.
reporting zero consumption during the 4-d record and reporting to never consume alcohol in the lifestyle questionnaire were categorized as zero-reporters. The second category was defined as no consumption during the 4-d record, but indication of alcohol consumption in the questionnaire, the other category ranges were <10 g/d, >10–20 g/d, >20–30 g/d, >30–40 g/d, and >40 g/d, according to reported consumption during the 4-d record. Use of antibiotics was based on the question “Have you used antibiotics during the previous 6 months?” Use of probiotics was defined as ≤3 times per week or >3 times per week based on information from the food propensity questionnaire.

**Definition of prediabetes**

In total, 260 cases were classified as having prediabetes. The classification was based on fasting plasma glucose ≥6.0 mmol/L and/or HbA1c ≥42 mmol/L at baseline (2013–2017) (n = 258) and/or individuals diagnosed with T2D during follow-up according to register data (n = 12) [follow-up until 31 December 2016, mean follow-up = 1.6 y (0–3.8)]. Register data were obtained from national and regional registries (3).

**Gut microbiota**

Fecal samples were collected without preservatives at home in sterile plastic tubes (80.9924.014 polypropylene, Sarstedt) and stored in a home freezer until they were brought to the clinic where they were stored at −80 °C. The microbial DNA was extracted using the QIAamp column Stool Kit and the V1–V3 region (300bp^2) of the 16S rRNA gene was amplified and sequenced on a HiSeq Illumina at the GATC Biotech (Constance). The fastq files were then aligned by FLASH and binned together to operational taxonomical units (OTUs) using QIIME 1 (Quantitative Insight Into Microbial Ecology) (20, 21). The sequences were matched with the reference database Greengene (v.13.8). OTUs with <0.01% counts assigned to them (out of the pool of all counts assigned to all OTUs) and OTUs occurring in <3 individuals were removed, leaving 64 bacteria characterized at genus level and belonging to 8 microbial phyla, for further analysis. All association analyses were performed using normalized absolute abundances, i.e. counts that were normalized by cumulative sum scaling (CSS) in R using the metagenomicSeq package. In addition, to descriptively illustrate percentage abundances of the bacterial genera in the study population, relative abundances were used. The Shannon diversity index was calculated using diversity within the R package vegan.

**Statistical analysis**

The SPSS statistical computer package (version 24.0; IBM Corporation) was used for all statistical analyses. All food variables were log transformed (e-log) to normalize the distribution before analysis. To handle log transformation of zero intakes, we added a very small amount (0.01 g). All food intakes were energy adjusted with the residual method.

We used principal component analysis (eigenvalues >1 and varimax rotation) to reduce 43 energy-adjusted food groups into factors representing food patterns. From the obtained scree plots (Supplemental Figure 2), we decided to retain and rotate the 2 factors with eigenvalues >2, that explained most of the variance in the data (6.8% and 5.2%, respectively). These factors were possible to interpret and translate into food patterns based on their loadings for the initial food group variables. In addition, these factors were found to be similar in men and women, indicating robust patterns. Reported characteristics of the patterns were based on food group loadings <−0.25 or >0.25. All individuals were assigned scores for each of the 2 factors that represented food patterns, corresponding to the agreement of their diet to the patterns.

We examined baseline characteristics according to prediabetes status and quintiles of the factors representing food patterns with the general linear model for continuous variables (adjusted for age and sex) and with the chi-square test for categorical variables. Food patterns in relation to prediabetes were examined using logistic regression.

The basic model included adjustments for age, sex, and total energy intake. A second multivariable model also included physical activity level, smoking, alcohol consumption, and level of education, and a third model additionally included BMI. Missing data for the potential confounding variables were treated as separate categories. We examined Spearman’s correlations between retained food patterns and microbiota genera. In addition, we examined microbiota genera across quintiles of the food patterns with the general linear model adjusted for age, sex, physical activity level, smoking, and alcohol consumption, as well as with additional adjustments for BMI and fiber intake. Finally, the gut bacterial genera found to associate with food patterns, after adjustment for lifestyle factors including BMI, were examined in relation to prediabetes with logistic regression. Tests for interactions between gender and food patterns on prediabetes and gut bacterial composition were performed [gender × quintile of food pattern (treated as continuous variables)].

In sensitivity analysis, we only included individuals who had not used antibiotics during the previous 6 mo and those who did not use probiotics >3 times per week.

All statistical tests were 2-sided. Statistical nominal significance was assumed at P < 0.05.

To correct for multiple testing, when analyzing dietary patterns in relation to 64 gut bacterial genera, the Bonferroni correction was applied and therefore statistical significance was assumed at P < 8 × 10^{-4} (0.05/64).

**Results**

We retained 2 factors from the principal component analysis. The first derived factor explained 6.8% of the variance in the data. We called it the Health-conscious food pattern, because it was characterized by (loadings <−0.25 or >0.25) high intakes of fruits and berries, nuts and seeds, legumes, other vegetables (nonlegumes), plain yogurt, fresh cheese, tea, animal replacement foods, breakfast cereals, cooked grains such as bulgur, oil-based dressings, fish and fiber-rich bread, and by low intakes of sugar-sweetened beverages, red and processed meat, white bread and fried/deep-fried potatoes (Figure 1A). The second food pattern, named the Sugar and High-Fat Dairy pattern, explained 5.2% of the variance and was characterized by high intakes of pastry and desserts, high-fat milk, cream, traditional sauces, jam and sugar, white bread, boiled potatoes, cheese, butter, eggs, processed meat, and sweets, and by low intake of food replacement products, such as weight loss powders (Figure 1B).

The factors showed strongest loadings for the same foods in women and men (data not shown), with the main exception that the Sugar and High-Fat Dairy pattern did not show a loading above 0.25 for cream in men (loading = 0.16).

**Baseline characteristics**

The study participants with high adherence to the Health-conscious food pattern were more often women and they were characterized by higher age and level of education, compared with those with low adherence to that food pattern (Table 1). In addition, those adhering to the Health-conscious pattern reported less sedentary leisure time, less heavy work, less smoking, and lower energy intake, and they had a lower BMI, blood pressure, fasting glucose, triglycerides, and higher HDL cholesterol. Their diets contained more polyunsaturated fat and fiber, but less sucrose, and they reported somewhat higher alcohol consumption. Those adhering to the Sugar and High-Fat Dairy pattern were also more often women and of higher age, level of education, and energy intake, compared with those with low adherence to the Sugar and High-Fat Dairy pattern.
In addition, they had a more sedentary leisure time and their diet contained more saturated fat and sucrose. Finally, they reported much higher alcohol consumption during the 4-d food record.

Individuals identified as having prediabetes were older and had a higher BMI, waist circumference, and blood pressure compared with those without prediabetes (Supplemental Table 1). They were also found to have a more sedentary leisure time,
lower LDL cholesterol and HDL cholesterol concentrations, and they reported lower alcohol consumption during the 4-d food record period.

The most abundant bacterial genus in MOS was *Bacteroides*, followed by an unclassified genus in the family Ruminococcaceae, another in the family Rikenellaceae, and 1 in the order Clostridiales, as well as *Faecalibacterium*, all with relative abundances of over 5% in the feces samples collected at baseline (Figure 2).

**Food patterns and prediabetes**

High adherence to the Health-conscious food pattern was associated with a lower prevalence of prediabetes, both in the basic model and after adjustment for several lifestyle factors ($P$ for trend across quintiles of the food pattern = 0.03) (Table 2). However, after additional adjustment for BMI the association did not remain significant ($P$-trend = 0.19). We did not observe

**TABLE 1** Baseline characteristics across quintiles of food patterns in 1726 individuals from the Malmö Offspring Study without prevalent diabetes

| Baseline characteristics | Quintile of the Health-conscious food pattern | Quintile of the Sugar and High-Fat Dairy pattern |
|--------------------------|---------------------------------------------|---------------------------------------------|
|                         | 1 ($n = 345$) | 2 ($n = 345$) | 3 ($n = 346$) | 4 ($n = 345$) | 5 ($n = 345$) | 1 ($n = 345$) | 2 ($n = 345$) | 3 ($n = 346$) | 4 ($n = 345$) | 5 ($n = 345$) | $P$-trend $^3$ | $P$-trend $^3$ |
| Age, y                   | 33.8 ± 0.73 | 38.5 ± 0.72 | 41.8 ± 0.72 | 42.2 ± 0.73 | 43.1 ± 0.74 | 0.001 | 0.001 |
| BMI, kg/m$^2$            | 26.9 ± 0.23 | 25.8 ± 0.23 | 25.8 ± 0.23 | 25.0 ± 0.23 | 23.9 ± 0.23 | 0.001 | 0.001 |
| Systolic BP, mmHg        | 115 ± 0.69 | 115 ± 0.67 | 115 ± 0.66 | 115 ± 0.67 | 114 ± 0.68 | 0.001 | 0.001 |
| Diastolic BP, mm Hg      | 71 ± 0.43 | 72 ± 0.43 | 72 ± 0.43 | 70 ± 0.44 | 69 ± 0.44 | 0.001 | 0.001 |
| P-Glucose, mmol/L        | 5.8 ± 0.11 | 5.8 ± 0.11 | 5.8 ± 0.11 | 6.2 ± 0.11 | 6.6 ± 0.11 | 0.001 | 0.001 |
| Energy intake, MJ/d      | 80.0 ± 0.14 | 82.0 ± 0.13 | 86.0 ± 0.13 | 87.0 ± 0.14 | 92.0 ± 0.14 | 0.001 | 0.001 |
| Age, y                   | 32.5 | 43.5 | 55.2 | 65.2 | 75.4 | 0.001 |
| Gender, female, %        | 32.5 | 43.5 | 55.2 | 65.2 | 75.4 | 0.001 |
| Smokers, current, %      | 21.8 | 14.6 | 14.2 | 9.8 | 6.3 | 0.001 |
| Higher education, %      | 23.4 | 34.5 | 38.9 | 45.4 | 56.1 | 0.001 |
| LPA, sedentary/low %     | 55.1 | 49.5 | 48.6 | 38.7 | 30.4 | 0.001 |
| Work AL, very light/ligt | 51.1 | 56.8 | 60.0 | 66.3 | 69.8 | 0.001 |

**FIGURE 2** Mean relative abundance (%) of bacterial genera in feces samples collected in 1477 individuals at baseline of the Malmö Offspring Study. Other refers to the 50 identified genera not specified in the figure (Supplemental Table 2).

Food patterns, prediabetes, and gut microbiota 865

(Continued)
any interaction between the Health-conscious food pattern and gender ($P = 0.72$).

The Sugar and High-Fat Dairy pattern did not indicate any overall association with the prevalence of prediabetes. However, we observed a statistical interaction with gender ($P_{interaction} = 0.03$); no significant association was seen in men ($P_{trend} = 0.40$), but women adhering to the Sugar and High-Fat Dairy pattern were more likely to have prediabetes (OR comparing the highest quintile with the lowest: 2.16; 95% CI: 1.02, 4.54; $P_{trend} = 0.001^5$). The association remained significant after adjustment for BMI ($P_{trend} = 0.04$).

Food patterns and gut microbiota

Adherence to the Health-conscious food pattern was significantly correlated to the abundance of 19 bacterial genera after the Bonferroni correction (Supplemental Table 2). After adjustment for lifestyle factors, a high adherence to the pattern was associated with a higher abundance of 4 bacterial genera and lower abundance of 6 genera. After additional adjustment for BMI, the highest abundance of *Lachnospira* and *Roseburia* genus in the RF39 order (Figure 3A), and the lower abundance of *Blautia*, *Anaerotruncus*, and *Eubacterium* (Figure 3B), remained significant. After additional adjustment for fiber intake, only the higher abundances of *Lachnospira* and *Roseburia*, and the lower abundance of *Eubacterium* remained significant. In a sensitivity analysis, restricted to individuals reporting no use of antibiotics during the previous 6 mo and consumption of probiotics $\geq 3$ times per week, the higher abundances of *Lachnospira* ($\beta = 0.22$; $P_{trend}$ across quintiles of the food pattern = $1 \times 10^{-4}$) and *Roseburia* ($\beta = 0.25$; $P_{trend} = 1 \times 10^{-4}$) with a higher adherence to the Health-conscious food pattern, and the lower abundance of *Eubacterium,* remained similar as in the whole study sample. However, the lower abundance of *Eubacterium* was only nominally significant ($\beta = -0.33$, $P_{trend} = 0.002$). In addition, higher adherence to the Health-conscious food pattern was associated with a significantly lower abundance of *Anaerotruncus* ($\beta = -0.25$; $P_{trend} = 3 \times 10^{-4}$) in the sensitivity analysis.

Adherence to the Sugar and High-Fat Dairy pattern correlated to a genus within the family Christensenellaceae and a genus within the order SFA98, but it did not significantly associate with any bacterial genera after adjustment for potential confounders, although a few nominally significant associations were observed (Supplemental Table 3).

We did not observe any significant association between the food patterns and $\alpha$-diversity (Shannon index across the food pattern quintiles, $P_{trend} \geq 0.16$).

Gender was not found to significantly modify the association between the food patterns and gut bacteria (data not shown).

Food pattern related gut bacteria and prediabetes

The abundance of *Roseburia* was inversely associated with the prevalence of prediabetes (OR per quintile of gut bacterial abundance: 0.86; 95% CI: 0.76–0.96; $P_{trend} = 0.01$) (Figure 4). The association remained significant after adjustment for adherence to the Health-conscious food pattern. The other 5 gut bacterial genera that associated with the Health-conscious food pattern were not associated with the prevalence of prediabetes, after adjustment for lifestyle factors including BMI (Table 4).

Adjusting for gut bacteria associating with both diet and prediabetes

In line with the result of the whole study cohort (Figure 5A), adherence to the Health-conscious food pattern tended to associate with a lower prevalence of prediabetes in the sample only including those with data on gut microbiota ($n = 1477$) (OR per quintile: 0.88; 95% CI: 0.79–1.00; $P_{trend} = 0.056$ after adjustment for potential confounders) (Figure 5B). The association was also attenuated after adjustment for BMI (OR: 0.93; 95% CI: 0.82–1.06; $P_{trend} = 0.30$) (Figure 5C). When additionally adjusting for the abundance of *Roseburia*, the association between the Health-conscious food pattern and prevalence of prediabetes was further attenuated (OR: 0.96; 95% CI: 0.84–1.10; $P_{trend} = 0.56$) (Figure 5D). The association was also slightly attenuated when adjusting for *Roseburia* but not BMI (OR: 0.91; 95% CI: 0.80–1.04; $P_{trend} = 0.15$).

Discussion

In this large observational study with data on diet and microbiota, 2 food patterns were extracted from principal component analysis. Participants adhering to the first food pattern, characterized by Health-conscious food choices, such...
as fiber-rich plant foods, were less likely to have prediabetes. Several gut bacterial genera correlated with adherence to this Health-conscious food pattern. A higher abundance of *Roseburia* was found to associate with both a higher adherence to the Health-conscious food pattern and with a lower prevalence of prediabetes. The association between the Health-conscious food pattern and prevalence of prediabetes was attenuated after adjustment for BMI and abundance of gut bacterial genera, the association with *Roseburia* is of special interest, as the food pattern was so strongly associated with *Roseburia* and as *Roseburia* per se was found to associate with a lower prevalence of prediabetes, independently of adherence to the Health-conscious food pattern.

As *Roseburia* is known to be a butyrate-producing genera, dependent on fermentable carbohydrates in the diet (25), it is biologically plausible that *Roseburia* could affect diabetes development, as butyrate and other SCFAs act as signal substances with beneficial effects on glucose metabolism (20, 26). Moreover, a recent study indicated that the butyrate-producing activity of the gut may causally affect the glucoses—stimulated insulin response and that *Roseburia* was among the bacteria showing the strongest correlation to such activity (8). Since different types of dietary fibers are utilized as substrates in bacterial SCFA production, it is possible that potential effects

### TABLE 3

| Sugar and high-fat dairy pattern | Quintile of food pattern | 1 OR | 2 OR (95% CI) | 3 OR (95% CI) | 4 OR (95% CI) | 5 OR (95% CI) | P-trend |
|---------------------------------|--------------------------|------|--------------|--------------|--------------|--------------|---------|
| Women                           |                          |      |              |              |              |              |         |
| Cases/controls                  | 18/136                   | 26/162| 27/159       | 27/169       | 39/175       |              |         |
| Basic model                     | 0.12 ± 0.07              | 1.00 | 1.21 (0.62, 2.33) | 1.21 (0.62, 2.34) | 1.28 (0.66, 1.50) | 1.42 (0.90, 2.31) | 0.10    |
| Multivariable model             | 0.18 ± 0.08              | 1.00 | 1.33 (0.62, 2.82) | 1.28 (0.60, 2.71) | 1.61 (0.76, 3.42) | 2.16 (1.02, 4.54) | 0.03    |
| Multivariable model with BMI    | 0.17 ± 0.08              | 1.00 | 1.42 (0.66, 2.83) | 1.32 (0.61, 2.83) | 1.61 (0.74, 3.48) | 2.21 (1.04, 4.72) | 0.04    |
| Men                             |                          |      |              |              |              |              |         |
| Cases/controls                  | 35/156                   | 24/133| 25/135       | 18/131       | 21/110       |              |         |
| Basic model                     | −0.09 ± 0.08             | 1.00 | 0.77 (0.43, 1.38) | 0.77 (0.42, 1.39) | 0.56 (0.29, 1.08) | 0.80 (0.42, 1.55) | 0.27    |
| Multivariable model             | −0.08 ± 0.09             | 1.00 | 0.71 (0.35, 1.42) | 0.72 (0.36, 1.45) | 0.59 (0.28, 1.24) | 0.78 (0.36, 1.65) | 0.40    |
| Multivariable model with BMI    | −0.06 ± 0.09             | 1.00 | 0.73 (0.37, 1.47) | 0.76 (0.38, 1.56) | 0.63 (0.30, 1.33) | 0.82 (0.38, 1.78) | 0.52    |

1* indicates mean difference per intake quintile ± SEs.
2 Adjusted for age, sex, and total energy intake.
3 Adjusted for age, sex, total energy intake, level of education, smoking, alcohol intake, and level of physical activity.
4 Adjusted for age, sex, total energy intake, level of education, smoking, alcohol intake, level of physical activity, and BMI.
FIGURE 3 High adherence to the Health-conscious food pattern was significantly associated with a higher abundance of \textit{Lachnospira}, \textit{Roseburia}, and a genus in the RF39 order (panel A), and with a lower abundance of \textit{Blautia}, \textit{Anaerotruncus}, and \textit{Eubacterium} (panel B) after adjustment for lifestyle factors including BMI (P values for trend across quintiles of the food pattern < $8 \times 10^{-4}$) in 1477 individuals from the Malmö Offspring Study. Data points show mean normalized absolute abundances (with 95% CIs) of the bacterial genera in quintiles (Q) of the food pattern. Q1, \( n = 286 \); Q2, \( n = 292 \); Q3, \( n = 306 \); Q4, \( n = 293 \); Q5, \( n = 300 \).

FIGURE 4 The prevalence of prediabetes was found to be lower in individuals with a higher abundance of \textit{Roseburia} (OR per quintile of gut bacterial abundance: 0.86; 95% CI: 0.76–0.96; \( P_{\text{trend}} = 0.01 \)) in the Malmö Offspring Study (\( n = 1477 \)). Data points show ORs (with 95% CIs) of prediabetes in quintiles (Q) of bacterial abundance (normalized absolute abundance of \textit{Roseburia}). Q1, \( n = 295 \); Q2, \( n = 296 \); Q3, \( n = 295 \); Q4, \( n = 296 \); Q5, \( n = 295 \).
the abundance of *Eubacterium* was lower at higher adherence to the pattern and was not found to associate with prediabetes. Different bacteria could compete for dietary substrates for butyrate production (33) and this may explain the contrasting associations. The potentially pathogenic *Anaerotruncus* (34) has, in accordance with our results showing a lower abundance of this genera at higher adherence to the Health-conscious food pattern could also have contributed to our observations. The association between dietary patterns and microbiota. Our consideration of studies with data on gut microbiota, detailed dietary data, and that only few observational studies have reported associations between dietary patterns and microbiota. Our intention was to study overall food patterns instead of single foods or nutrients. This approach has several advantages: foods or nutrients, and foods are consumed together and may interact. For example, it is possible that foods with probiotic components, such as some fermented dairy products, may also be easier to detect compared with those of single foods or nutrients, and foods are consumed together and may interact. For example, it is possible that foods with probiotic components, such as some fermented dairy products, may have greater effects if accompanied by foods with prebiotic qualities such as fiber-rich foods (37, 38). In addition, it cannot be excluded that a pattern indicating Health-conscious food choices might be a better marker of long-term fiber intake than assessed fiber intake per se reported during 4 specific days. On the other hand, fiber consumption close to feces sampling is also

### Table 4

| Gut bacterial genus | β | 1 OR | 2 OR (95% CI) | 3 OR (95% CI) | 4 OR (95% CI) | 5 OR (95% CI) | P-trend |
|---------------------|---|------|---------------|---------------|---------------|---------------|---------|
| *Roseburia*         |   |      |               |               |               |               |         |
| Cases/controls      | 58/237 | 46/250 | 45/250 | 36/260 | 44/251 |         |
| Basic model         |      | −0.11 ± 0.05 | 1.00 | 0.65 (0.42, 1.00) | 0.68 (0.44, 1.06) | 0.50 (0.31, 0.79) | 0.68 (0.44, 1.06) | 0.04 |
| Multivariable model |      | −0.16 ± 0.06 | 1.00 | 0.59 (0.37, 0.93) | 0.54 (0.33, 0.87) | 0.43 (0.26, 0.71) | 0.56 (0.35, 0.92) | 0.01 |
| Multivariable model with BMI | | | | | | | |

| Blautia              |   |      |               |               |               |               |         |
| Cases/controls      | 52/243 | 41/255 | 46/249 | 37/259 | 53/242 |         |
| Basic model         |      | 0.03 ± 0.05 | 1.00 | 0.82 (0.52, 1.31) | 1.02 (0.65, 1.59) | 0.77 (0.48, 1.23) | 1.20 (0.78, 1.85) | 0.53 |
| Multivariable model |      | 0.06 ± 0.06 | 1.00 | 0.80 (0.49, 1.32) | 0.95 (0.58, 1.54) | 0.82 (0.49, 1.35) | 1.30 (0.62, 2.07) | 0.32 |
| Multivariable model with BMI | | | | | | | |

| Lachnospira         |   |      |               |               |               |               |         |
| Cases/controls      | 50/245 | 49/247 | 40/255 | 40/256 | 50/245 |         |
| Basic model         |      | −0.03 ± 0.05 | 1.00 | 0.97 (0.63, 1.51) | 0.74 (0.47, 1.18) | 0.76 (0.48, 1.28) | 0.98 (0.63, 1.51) | 0.58 |
| Multivariable model |      | 0.001 ± 0.06 | 1.00 | 1.22 (0.75, 1.98) | 0.92 (0.55, 1.54) | 1.01 (0.61, 1.67) | 1.10 (0.68, 1.79) | 0.99 |
| Multivariable model with BMI | | | | | | | |

| Anaerotruncus       |   |      |               |               |               |               |         |
| Cases/controls      | 63/315 | 34/179 | 35/296 | 47/249 | 46/249 |         |
| Basic model         |      | 0.02 ± 0.05 | 1.00 | 0.99 (0.63, 1.58) | 0.86 (0.55, 1.33) | 1.12 (0.74, 1.71) | 1.04 (0.68, 1.59) | 0.71 |
| Multivariable model |      | −0.002 ± 0.06 | 1.00 | 0.96 (0.58, 1.58) | 0.85 (0.53, 1.35) | 0.96 (0.59, 1.54) | 1.01 (0.64, 1.59) | 0.97 |
| Multivariable model with BMI | | | | | | | |

| Eubacterium         |   |      |               |               |               |               |         |
| Cases/controls      | 40/255 | 47/249 | 41/254 | 53/243 | 48/247 |         |
| Basic model         |      | +0.03 ± 0.05 | 1.00 | 1.28 (0.81, 2.04) | 1.03 (0.64, 1.66) | 1.44 (0.92, 2.27) | 1.11 (0.70, 1.77) | 0.54 |
| Multivariable model |      | +0.08 ± 0.06 | 1.00 | 1.15 (0.69, 1.93) | 1.04 (0.61, 1.75) | 1.44 (0.88, 2.33) | 1.32 (0.82, 2.33) | 0.17 |
| Multivariable model with BMI | | | | | | | |

| Genus within the order RF39 |   |      |               |               |               |               |         |
| Cases/controls            | 70/375 | 41/216 | 45/214 | 33/225 | 40/218 |         |
| Basic model               |      | −0.05 ± 0.05 | 1.00 | 0.98 (0.67, 1.58) | 0.76 (0.69, 1.59) | 0.91 (0.63, 1.46) | 0.45 (0.62, 1.16) | 0.33 |
| Multivariable model       |      | −0.05 ± 0.06 | 1.00 | 0.92 (0.57, 1.47) | 1.04 (0.65, 1.63) | 0.65 (0.39, 1.06) | 0.91 (0.57, 1.45) | 0.33 |
| Multivariable model with BMI | | | | | | | |

1. β indicates mean difference per quintile of normalized absolute abundance ± SEs.
2. Calculated with the general linear model. Adjusted for age and sex (continuous) when appropriate.
3. Adjusted for age, sex, smoking, alcohol intake, and physical activity.
4. Adjusted for age, sex, smoking, alcohol intake, physical activity, and BMI.

Favorable effects of antioxidants in fruits and vegetables characterizing the Health-conscious food pattern could also have contributed to the observed lower prevalence of prediabetes, as antioxidants may affect glucose metabolism (36).

The strengths of this study are the large size, when considering studies with data on gut microbiota, detailed dietary data, and that only few observational studies have reported associations between dietary patterns and microbiota. Our intention was to study overall food patterns instead of single foods or nutrients. This approach has several advantages; food intakes are correlated and it may be challenging to disentangle their individual importance, cumulative effects of several foods may also be easier to detect compared with those of single foods or nutrients, and foods are consumed together and may interact. For example, it is possible that foods with probiotic components, such as some fermented dairy products, may have greater effects if accompanied by foods with prebiotic qualities such as fiber-rich foods (37, 38). In addition, it cannot be excluded that a pattern indicating Health-conscious food choices might be a better marker of long-term fiber intake than assessed fiber intake per se reported during 4 specific days. On the other hand, fiber consumption close to feces sampling is also
of importance, as the changes in fiber intake can rapidly affect the gut bacterial composition (39).

A limitation of our study is that we could not examine incident T2D, as the participants were followed for only a few years resulting in very few incident cases. Instead, we could examine the prevalence of prediabetes, defined according to baseline concentrations of fasting glucose and HbA1c or the development of T2D during the short follow-up. However, the observed association between the Health-conscious food pattern and the butyrate-producing genera *Roseburia* is biologically plausible, considering that the pattern was characterized by fiber-rich foods that provide substrates for SCFA production, which in turn affect glucose metabolism. We also need to acknowledge that measurement error is a major problem in studies with self-reported diet and 1 consequence might have been unsatisfactory adjustment for fiber intake. Moreover, we cannot completely exclude the risk of overadjustment, when including fiber intake in the same model as the Health-conscious food pattern. Another drawback is that both diet and microbiota were only measured at baseline in the whole study sample. Nevertheless, data from a subsample of our study population with repeat measurements of diet indicate acceptable agreement between the measurements; with Pearson correlation coefficients of 0.6 and 0.7 for energy-adjusted intakes of fiber and fat, respectively (40). Unfortunately, we do not have any repeat measurements of microbiota. Another concern is that our results may not be generalizable to other populations; different dietary patterns may, for example, occur in other study settings. In addition, dietary patterns could represent overall lifestyle, and despite adjustment for several confounders, we cannot completely rule out residual confounding. On the other hand, the fact that the Health-conscious pattern was characterized by foods that per se have been found to be associated with cardiometabolic disease in other studies (41–48) makes our findings more credible. Finally, a lack of power may be an important issue in the sensitivity analyses, as almost half the study sample was removed.

To conclude, our findings suggest that the association between the Health-conscious food pattern and lower prevalence of prediabetes may partly be explained by links between healthy diet, BMI-status, and abundance of *Roseburia*, as the included fiber-rich foods may constitute substrates for gut bacterial SCFA production and thereby affect glucose metabolism. The association between adherence to the Sugar and High-Fat Dairy pattern and higher prevalence of prediabetes may be explained by more direct links to glucose metabolism. Future studies are warranted to replicate our findings in other populations and to evaluate their importance in experimental settings.

**Acknowledgments**

We thank data manager Anders Dahlin for extensive quality control of MOS data, and Johan Hultman for working with setting up the microbiota pipeline. The authors’ contributions were as follows—UE, LB, and MO-M: designed the research; UE, LB, and SH: contributed to data collection; LB: performed the microbiome analysis from raw DNA sequencing data; UE: performed the statistical analysis and wrote the paper; UE, LB, SH, PMN, and MO-M: contributed to the interpretation of results and revision of the manuscript; UE: had primary responsibility for final content; and all authors read and approved the final manuscript.

**References**

1. Maghsoudi Z, Ghiassvand R, Salehi-Abargouei A. Empirically derived dietary patterns and incident type 2 diabetes mellitus: a systematic review and meta-analysis on prospective observational studies. Public Health Nutr 2016;19:230–41.

2. McEvoy CT, Cardwell CR, Woodside JV, Young IS, Hunter SJ, McKinley MC. A posteriori dietary patterns are related to risk of type 2 diabetes: findings from a systematic review and meta-analysis. J Acad Nutr Diet 2014;114:1759–73 e1754.

3. Ericson U, Brunkwall L, Alves Dias J, Drake I, Hellstrand S, Gullberg B, Sonesstedt E, Nilsson PM, Wirfalt E, Orlo-Melaner M. Food patterns in relation to weight change and incidence of type 2 diabetes, coronary events and stroke in the Malmo Diet and Cancer cohort. Eur J Nutr 2018;58(5):1801–14.

4. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, Kurilshikov A, Bonder MJ, Valles-Colomer M, Vandeputte D, et al. Population-level analysis of gut microbiome variation. Science 2016;352:560–4.

5. Ley RE, Turnbaugh PJ, Klein S, Gordon JL. Microbial ecology: human gut microbes associated with obesity. Nature 2006;444:1022–3.
6. Qin J, Li Y, Cai Z, Li S, Zhu J, Zheng F, Liang S, Zhang W, Guan Y, Shen D, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature 2012;490:55–60.

7. Ottosson F, Brunkwall L, Ericsson U, Nilsson PM, Almgren P, Fernandez C, Melander O, Orho-Melander M. Connection between BMI-related plasma metabolite profile and gut microbiota. J Clin Endocrinol Metab 2018;103:1491–501.

8. Sanna S, van Zuydam NR, Mahajan A, Kurilshikov A, Vich Vila A, Vosa U, Muigaj Z, Mascele AAM, Jonkers D, Oosting M, et al. Causal relationships among the gut microbiota, short-chain fatty acids and metabolic diseases. Nat Genet 2019;51(4):600–5.

9. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Weir TL. Linking dietary fiber intake to health outcomes: a systematic review and meta-analysis. Am J Clin Nutr 2018;107:965–83.

10. De Filippis F, Pellegrini N, Vannini L, Jeffery RB, La Storia A, Laghi L, Serrazanetti DL, Di Cagno R, Ferrocino I, Lazzi C, et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut 2015;65(11):1812–21.

11. Garcia-Mantrana I, Selma-Royo M, Alcantara C, Collado MC. Shifts on gut microbiota associated to Mediterranean diet adherence and specific dietary intakes on general adult population. Front Microbiol 2018;9:890.

12. Shelif AM, Melby CL, Carbonero F, Weir TL. Linking dietary patterns with gut microbial composition and function. Gut Microbes 2017;8:113–29.

13. Mitsou EK, Kakali A, Antonopoulou S, Mountzouris KC, Yannakoulia M, Panagiotakos DB, Kyriacou A. Adherence to the Mediterranean diet is associated with the gut microbiota pattern and gastrointestinal characteristics in an adult population. Br J Nutr 2017;117:1645–55.

14. Brunkwall L, Orho-Melander M. The gut microbiome as a target for prevention and treatment of hyperglycaemia in type 2 diabetes: from current human evidence to future possibilities. Diabetologia 2017;60:943–51.

15. Manjer J, Carlsson S, Elmståhl S, Gullberg B, Janzon L, Lindstrom M, Mattisson I, Berglund G. The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. Eur J Cancer Prev 2001;10:489–99.

16. Wirfält E, Hedblad B, Gullberg B, Mattisson I, André C, Rosander U, Janzon L, Berglund G. Food patterns and components of the metabolic syndrome in men and women: a cross-sectional study within the Malmo Diet and Cancer cohort. Am J Epidemiol 2001;154:1150–9.

17. Nybacka S, Bertes Forslund H, Wirfält E, Larsson I, Ericson U, Warenjo Lemming E, Bergstrom G, Hedblad B, Winkvist A, Lindroos AK. Comparison of a web-based food record tool and a food-frequency questionnaire and objective validation using the doubly labelled water technique in a Swedish middle-aged population. J Nutr Sci 2016;5:e39.

18. Nybacka S, Lindroos AK, Wirfalt E, Leandersson P, Landberg R, Ericson U, Larsson I, Warenjo Lemming E, Bergstrom G, Hedblad B, et al. Carotenoids and alkylresorcinols as objective biomarkers of diet quality when assessing the validity of a web-based food record tool and a food frequency questionnaire in a middle-aged population. BMC Nutrition 2016;2:53.

19. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. Am J Epidemiol 1986;124:17–27.

20. Gao Z, Yin J, Zhang J, Ward RE, Martin RJ, Leefltre M, Cefalu WT, Ye J. Butyrate improves insulin sensitivity and increases energy expenditure in mice. Diabetes 2009;58:1509–17.

21. So D, Whelan K, Rossi M, Morrison M, Holtmann G, Kelly JT, Shanahan ER, Staudacher HM, Campbell KL. Dietary fiber intervention in mice. Diabetes 2009;58:1509–17.

22. Bakker BM. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. J Lipid Res 2013;54:2325–40.

23. Feng ZM, Li TJ, Wu L, Xiao DF, Blachier F, Yin YL. Monosodium L-glutamate and dietary fat differently modify the composition of the intestinal microbiota in growing pigs. Obes Facts 2015;8:87–100.

24. Kok CR, Hutkins R. Yogurt and other fermented foods as sources of health-promoting bacteria. Nutr Rev 2018;76:4–15.

25. Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, Nielsen J, Backhed F. Gut microbiota in European women with normal, impaired and diabetic glucose control. Nature 2013;498:99–103.

26. Zhu Q, Gao R, Zhang Y, Pan D, Zhu Y, Zhang X, Yang R, Jiang R, Xu Y, Qin H. Dysbiosis signatures of gut microbiota in coronary artery disease. Physiol Genomics 2018;50(10):893–903.

27. Bang SJ, Kim G, Lim MY, Song EJ, Jung DH, Kum JS, Nam YD, Park CS, Seo DH. The influence of in vitro pectin fermentation on the human fecal microbiome. AMB Express 2018;8:98.

28. Van’t Veer P, van Leer EM, Rietdijk A, Kok FJ, Schouten EG, Hermus AR, Bewtra M, Knights D, Walters WA, Knight R, et al. Linking long-term dietary patterns with gut microbial composition and function. Gut Microbes 2018;9:2344.

29. Aksnes EM, Hokaeys M, Bishal C, Lambert K. Dietary antioxidants: do they have a role to play in the ongoing fight against abnormal glucose metabolism? Nutrition 2012;28:715–21.

30. Van’t Veer P, van Leer EM, Rietdijk A, Kok FJ, Schouten EG, Hermus RJ, Sturmans F. Combination of dietary factors in relation to breast-cancer occurrence. Int J Cancer 1991;47:649–53.

31. Rafter J, Bennett M, Caderni G, Clune Y, Hughes R, Karlsson PC, Klimander A, Örjord M, O’ Sullivan GC, Poull Zobel B, et al. Dietary symbiotics reduce cancer risk factors in polycystic and colon cancer patients. Am J Clin Nutr 2007;85:488–96.

32. Wu GD, Chen J, Hoffmann B, Raffat CJ, Kugizaki M, Liu S. Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease, and weight gain. J Nutr 2012;142:1304–13.

33. Michi R, Wallace SK, Mozaffarian D. Red and processed meat consumption and risk of incident coronary heart disease, stroke, and diabetes mellitus: a systematic review and meta-analysis. Circulation 2010;121:2271–83.

34. Patel PS, Forouhi NG, Kuijsten A, Schulze MB, van Woudenbergh GJ, Ardanaz E, Amiano P, Arriola L, Balkau B, Barricarte A, et al. The
prospective association between total and type of fish intake and type 2 diabetes in 8 European countries: EPIC-InterAct Study. Am J Clin Nutr 2012;95:1445–53.

46. Aune D, Norat T, Romundstad P, Vatten LJ. Whole grain and refined grain consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. Eur J Epidemiol 2013;28:845–58.

47. Wang M, Yu M, Fang L, Hu RY. Association between sugar-sweetened beverages and type 2 diabetes: a meta-analysis. J Diabetes Investig 2015;6:360–6.

48. Sluijs I, Forouhi NG, Beulens JW, van der Schouw YT, Agnoli C, Arriola L, Balkau B, Barricarte A, Boeing H, Bueno-de-Mesquita HB, et al. The amount and type of dairy product intake and incident type 2 diabetes: results from the EPIC-InterAct Study. Am J Clin Nutr 2012;96:382–90.