BMC Medicine

RESEARCH

Open Access

Circulating short-chain fatty acids and Mediterranean food patterns. A potential role for the prediction of type 2 diabetes risk: The Di@bet.es Study

Gemma Llauradó[1,2*†], Lídia Cedó[2,3†], Elisenda Climent[1], Joan Badia[4], Gemma Rojo-Martínez[2,5], Juana Flores-Le Roux[1], Oscar Yanes[2,6,7], Maria Vinaixa[2,6,7], Minerva Granado-Casas[2,8,9], Didac Mauricio[2,10], Sonia Fernández-Veledo[2,3,11] and Joan Vendrell[2,3,11,12*]

Abstract

Background Identifying nutritional patterns associated with developing type 2 diabetes (T2D) can facilitate more effective and personalized dietary interventions. Short-chain fatty acids (SCFAs), key metabolites derived from gut microbiota, are produced through the anaerobic fermentation of dietary fibers. This study aimed to evaluate whether circulating concentrations of SCFAs are associated with specific food consumption patterns and to assess their association with T2D development in at-risk subjects within a prospective cohort (The Di@bet.es Study).

Methods The Di@bet.es study is a prospective, population-based study utilizing random cluster sampling from the Spanish population aged over 18 years ($n = 5,072$). Among these participants, 4,347 were free of T2D at baseline. Follow-up losses were approximately 45%, resulting in a final re-screened sample of 2,408 subjects. A qualitative food frequency questionnaire evaluated Mediterranean diet (MedDiet) adherence and high-fiber food consumption. The risk of developing T2D was assessed using the FINDRISK. Metabolomics-driven analyses of SCFAs were conducted using gas chromatography-mass spectrometry.

Results Subjects who developed T2D after a median follow-up of seven years had higher baseline circulating concentrations of butyrate and isobutyrate. Circulating concentrations of SCFAs were associated with high-fiber food consumption at baseline. In multivariate analysis, baseline circulating concentrations of butyrate and isobutyrate were independently associated with incident T2D after adjusting for traditional clinical factors. The C-statistics for predicting T2D were 0.847 (95%CI:0.816–0.877) for butyrate and 0.843 (95%CI:0.812–0.875) for isobutyrate in adjusted models, similar to the reference model based on traditional clinical factors (0.840 [95%CI: 0.807–0.873]). Both models improved risk prediction compared to FINDRISK. Dietary patterns did not add predictive value. Sensitivity analysis excluding subjects with prediabetes at baseline confirmed these results. In addition, an association between baseline consumption

†Gemma Llauradó and Lídia Cedó contributed equally to this work.

*Correspondence:
Gemma Llauradó
gllauradoc@gmail.com
Joan Vendrell
jvortega@iispv.cat
Full list of author information is available at the end of the article

of high-fiber foods with incident T2D emerged, suggesting a different behaviour between healthy and prediabetic subjects.

Conclusions Baseline circulating concentrations of SCFAs are associated with high-fiber food consumption and independently predict the development of T2D over seven years of follow-up. However, they offer limited improvement in risk prediction compared to traditional risk factors, though they enhance risk prediction as assessed by FINDRISK. Further studies are necessary to evaluate the impact of dietary interventions on SCFA.

Keywords Short-chain fatty acids, Type 2 diabetes mellitus, Mediterranean Diet, High-fiber food consumption, Risk prediction

Background

Obesity has reached epidemic proportions, and its comorbidities, most notably type 2 diabetes mellitus (T2D), have become a major public health concern [1]. Traditional diabetes risk factors have limited utility in identifying individuals at high risk for developing T2D. Consequently, cross-sectional epidemiological studies are being conducted to identify new biomarkers and better estimate this risk. In this context, it is essential to discover novel, reliable biomarkers to improve the identification of at-risk individuals.

Metabolites are generally considered to function mainly as energy sources or building blocks in metabolic pathways. Additionally, many metabolites may function intracellularly as signaling molecules by regulating the activity of some enzymes or modulating intracellular receptors such as nuclear hormone receptors [2]. However, it is becoming increasingly evident that key metabolites act as extracellular signaling molecules similar to neurotransmitters and hormones. These "signaling metabolites" can activate specific G-protein-coupled receptors (GPCRs) and include microbiota-derived products [2, 3]. These new players may act in an autocrine, paracrine, and endocrine manner, but they also function as "metabokines" [2]. In this sense, GPCR-mediated signaling metabolites act as efficient pro- and anti-inflammatory regulators of crucial immune cells. They may drive the low-grade inflammation associated with obesity and insulin resistance [3].

Short-chain fatty acids (SCFAs) are probably the most well-known gut microbiota-derived metabolites, their major source of the anaerobic fermentation of dietary fibers in the intestine [4]. The SCFAs receptors (FFA2 and FFA) are activated by acetate, propionate, and butyrate (albeit with somewhat different affinity) [5]. Both FFA2 (GPR43) and FFA3 (GPR41) are expressed in the endocrine cells of the intestine (where their activation stimulates glucagon-like peptide-1 (GLP-1) secretion and regulates expression of PYY) and pancreatic islets [3]. Emerging evidence points to some metabolites, such as acetate, as potential regulators of glucose-induced insulin secretion by pancreatic beta-cells [6]. Furthermore, FFA2 is expressed in white adipose tissue, where it inhibits lipolysis [7], and in various immune cells, where it has pro-inflammatory effects [3]. FFA3 is also expressed in sympathetic ganglion cells where it mediates the stimulation of sympathetic activity [8]. Increased fecal SCFAs have been reported in overweight, obese, and subjects with T2D [9]. Although fecal SCFAs are commonly used to indicate microbial fermentation, they reflect the net result of colonic production and absorption and may not accurately reflect in vivo fermentation. Thus, evidence on circulating SCFA levels, as opposed to fecal SCFA levels, remains scarce.

The regulation of energy metabolism requires the well-balanced control of opposing metabolic pathways, often disturbed in obesity and T2D. Given the numerous functions of these microbiota-derived metabolites and the central role in coordinating metabolic processes and regulating low-grade inflammation response, these metabolites are highly likely to be useful as biomarkers for predicting the risk of T2D and as potential targets for drugs that modulate metabolic fluxes and, consequently, influence the outcome metabolic disorders. In addition, identifying nutritional patterns associated with the development of T2D might help to develop more efficient and personalized nutritional interventions. This study aimed to evaluate whether baseline circulating concentrations of SCFAs are related to the development of T2D in at-risk subjects (incident cases) within a prospective study (The Di@bet.es Study) and whether they are associated with specific food consumption patterns.

Methods

Study subjects

The subjects from this study were part of the Di@bet.es Study cohort who participated in the T2D incidence study [10]. The Di@bet.es study was the first nationwide, prospective, population-based study in Spain designed to determine the prevalence and incidence of T2D [10, 11]. A cluster sampling design was used to select participants, forming a representative random sample of the Spanish population census. Of eligible adults, 55.8%

underwent examination, and 9.9% were excluded by protocol design (institutionalized, severe disease, surgery within the previous month, pregnancy, or recent delivery). A final sample of 5,076 individuals aged ≥ 18 years (58.4% women) was included in the study. Of these, 4,347 were free of T2D at baseline (prevalence study) and were invited to participate in the follow-up study (incidence study). Follow-up losses were about 45%. At baseline and follow-up, people with serious illness, pregnancy, recent delivery or lactation, or surgery within the previous month were excluded. The final sample for re-screening finally comprised 2,408 subjects, 293 (12%) with prediabetes [10]. The prevalence study was undertaken between 2008 and 2010, and the subjects were evaluated a second time (incidence study) between 2016 and 2018 (median follow-up: 7.5 (7.1–7.9) years). In the present study, we included 2,900 subjects with available frozen blood samples to measure baseline circulating concentrations of SCFA (659 subjects with T2D at baseline [prevalence study] and 2,241 subjects free from T2D at baseline -subjects at-risk- who participated in the follow-up [incidence study]) (Fig. 1).

Ethical considerations

The research was carried out following the Declaration of Helsinki (2008) of the World Medical Association. The Ethics and Clinical Research Committee of the Hospital Regional Universitario de Málaga (Málaga, Spain) approved the baseline and follow-up studies. All the participants were informed about the nature of the study and provided written informed consent in the two phases, a document approved by the committee mentioned above. Although the cohorts included in this application were already established, this study was also approved by the Clinical Research Ethics Committee (CEIM) of the Hospital del Mar Medical Research Institute (IMIM).

Study procedures

The study procedures have been previously described in detail [10, 11]. At baseline and follow-up, the participants were invited to attend a single examination visit at their health center with a nurse specially trained for this project. Information was collected using an interviewer-administered structured questionnaire, followed by a physical examination, blood sampling, and an oral glucose tolerance test (OGTT). A structured questionnaire included closed questions was used to collect the following data: sex, age, educational level (none, basic, high school or college), smoking (current, ex- or never smoker), personal history of T2D (yes/no), hypertension (blood pressure equal or higher 140/90 mmHg or receiving antihypertensive treatment) and dyslipidemia (triglycerides were equal or higher 1.7 mmol/L or HDL cholesterol less than 1.03 mmol/L in men or less than 1.29 mmol/L in women or medication), family history of T2D (at least one first degree relative with T2D), medications (with a particular focus on T2D, blood pressure or dyslipidemia treatment); a physical activity (International physical activity questionnaire, IPAQ), and a semiquantitative food frequency consumption questionnaire [12]. The physical examination included measurement of body weight, height, and waist and hip

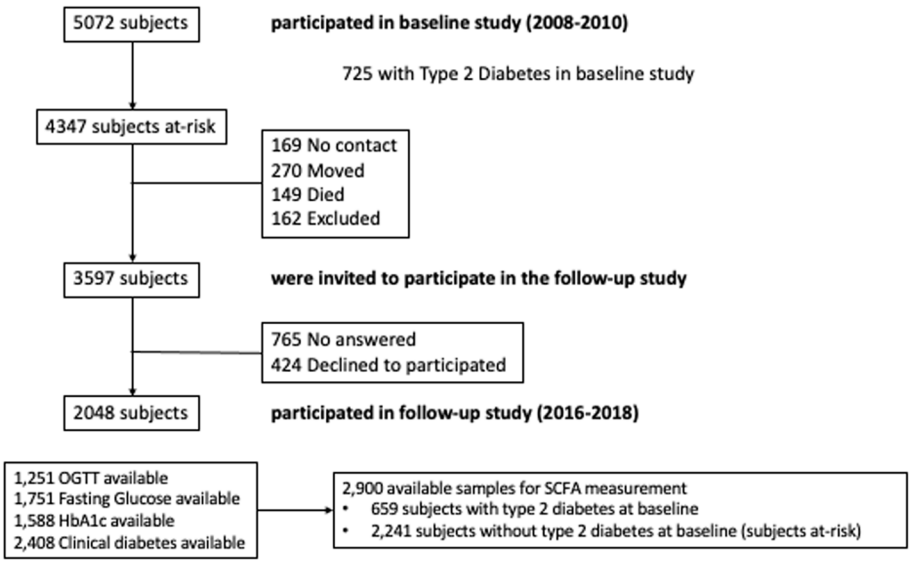

Fig. 1 Participation flow chart. Flow-chart of the Di@bet.es Study including the number of participants for both prevalence study (baseline) and incidence of T2D study. SCFA: schort-chain fatty acids; OGTT: oral glucose tolerance test

circumferences performed using standardized methods [13]. Besides, body mass index (BMI) and waist-to-hip ratio (WHR) were calculated.

Clinical and laboratory analyses

Both at baseline and follow-up, fasting and 2-h glucose were quantified before and after the administration of 75 g OGTT for individuals with fasting capillary glucose < 7.8 mmol/L (measured by OneTouch® system, Lifescan, Johnson & Johnson, S.A., Madrid, Spain). Biochemical measurements were performed with standard procedures. Glucose was determined by the hexokinase enzymatic method and HbA1c (only at follow-up) by high-performance liquid chromatography (analyzer ADAMS A1C HA-8180V, ARKRAY, Minneapolis, MN, US). The Homeostatic Model Assessment for Insulin Resistance (HOMA-IR) was calculated as insulin (IU)* glucose (mmol/L)/22.5. Blood was withdrawn in fasting conditions and frozen at -80 ºC for SCFA analyses. The samples were deposited in the Biobank of the Hospital Regional Universitario de Málaga-IBIMA, which belongs to the Andalusian Public Health System Biobank, and the biorepository CIBERDEM managed by IDIBAPS Biobank (Barcelona, Spain).

Food Frequency Questionnaire and Mediterranean Diet Score

A qualitative food frequency questionnaire (FFQ), previously used in a population-based study in southern Spain [12], was administered face-to-face by a trained dietician. Dieticians were aware of the participants' previous diabetic status. Annual frequency consumption of 50 food items was prespecified in 11 different categories as follows: never/seldom, 1 and 2–3 times/month, 1, 2–3 and 4–6 times/week, and 1, 2, 3, 4 and > 4 times/day. For this study, food items in the FFQ were assigned to food groups as follows: cereals (including non-refined cereals, bread, pasta, rice, and couscous), fruit (whole fruit and natural juices), vegetables (salads and raw or cooked vegetables), legumes (including nuts), potatoes, fish (fresh or canned and shellfish); meat and meat products (including beef, pork, and poultry), dairy products (including cheese, milk, yogurt, custard and ice-cream), olive oil (for either dressing, deep-frying or cooking) and wine (red and white/cava wines). Mediterranean diet (MedDiet) and consumption of High-Fiber Food were scored using the data provided by the FFQ at baseline. Adherence to the MedDiet was evaluated following an adaptation of the Panagiotakos score in the ATTICA study [14] and as previously described [15]. Based on a priori recommendations [14, 16], we appropriately scored the weekly frequency of the consumption of food groups characteristic of the MedDiet (nonrefined cereals, vegetables, fruit, legumes, potatoes, and fish) from 5 to 0. Conversely, a reverse rating (0–5) was gradually assigned from higher to lower consumption frequencies for food groups not typical of this diet pattern, particularly meat and meat derivatives and dairy products. For wine, we assigned a score of 5 to consumption seven times per week and 0 to consumption < 2–3 times per month or to clearly unhealthy wine consumption of > 28 or 21 times per week for males and females, respectively. Finally, since olive oil is an essential component of the MedDiet, we assigned a score of 5 to those individuals who always chose olive oil (among olive, sunflower, or other vegetable oils) as the sole source of oil for dressing, frying, and cooking, and a score of 0 to those who did not always choose it. The MedDiet Score was thus calculated for each participant as the sum of scores for each food group ($n = 9$) and had a potential range of 0–45. Higher values of this diet score indicate greater adherence to the MedDiet [14, 17]. In addition, as the major source of SCFA comes from the anaerobic fermentation of dietary fibers in the intestine, we created a priori high-fiber food group including nonrefined cereals, fruits (whole fruit and natural juices), vegetables (salads and raw or cooked vegetables), legumes (including nuts) and potatoes, and using the same rating used in the MedDiet score. The High-Fiber Food score was thus calculated for each participant as the sum of scores for each food group ($n = 5$) and had a potential range of 0–25.

Finnish Diabetes Risk Score

The baseline clinical characteristics of subjects in the present study were used to estimate their risk of developing T2D according to the Finnish Diabetes Risk Score (FINDRISK) [18], previously validated for the Spanish population [19].

Targeted metabolomics analysis

A metabolomics-driven analysis of SCFA (acetate, propionate, butyrate, and isobutyrate) was carried out in baseline samples obtained from the Di@bet.es cohort to identify subjects who developed T2D after seven years of follow-up. A targeted analysis by gas chromatography-mass spectrometry (GC–MS) was carried out to determine the selected metabolites in plasma samples [20]. The study was performed with the Centre for Omic Sciences (COS) equipment, the Joint Unit of the Universitat Rovira i Virgili and Eurecat. The experimental procedure of the analysis was, firstly, the extraction of the metabolites of the matrix by liquid–liquid extraction, followed by the analysis of the extracts using gas chromatography separation before determining the metabolites by

mass spectrometry. The intra- and interassay coefficient of variance was 2.5% and 6.3% for acetate, 7.4% and 8.2% for propionate, 11% and 15.7% for butyrate and 1.2% and 5.8% for isobutyrate, respectively.

Definition of outcomes at follow-up

As previously mentioned, the incidence of T2D was recorded after a median follow-up of 7.5 (7.1–7.9) years. Incident T2D was defined as fasting serum glucose equal to or higher than 7 mmol/L or 2-h post-load serum glucose equal to or higher than 11 mmol/L or HbA1c equal or higher than 6.5% or use of glucose-lowering medication at the follow-up examination [21]. Prediabetes was defined as impaired fasting glucose (IFG; fasting glucose between 6.1 and 6.9 mmol/L), impaired glucose tolerance (IGT; glucose levels between 7.8 and 11 mmol/L after 2-h glucose overload), or both (IFG/IGT).

Statistical analysis

All data were tested for normality using the Shapiro–Wilk test. Data are presented as percentages, means (standard deviation (SD)) for normally distributed quantitative variables, or medians (interquartile ranges) for non-normally distributed quantitative variables. Non-normally distributed quantitative variables were analyzed using a non-parametrical test. To analyze the differences for categorical variables, we used the X^2 test. Statistical significance between the two groups was tested using an unpaired T-test or Mann–Whitney U test. As needed, one-way analysis of variance (ANOVA) or the Kruskal–Wallis test was used to compare groups (3 or more) of normally and non-normally distributed quantitative variables. The Bonferroni procedure (parametric) and the Dunn's test (non-parametric) were used for post hoc analyses for multiple comparisons. Logistic regression analyses were performed to identify the potential role of baseline circulating concentrations of SCFA and the other factors independently related to the risk of incident T2D. All associated variables in the univariate analyses ($0.67 < \text{odds ratio [OR]} > 1.67$ and $p < 0.2$) and those variables known or likely to be associated with remission (based on previous literature) were included in the logistic regression models as potential independent variables. Receiver operating characteristic (ROC) curves, in which sensitivity is plotted as a function of 1-specificity, were developed to assess the predictive value of baseline circulating concentrations of SCFA for the prediction of incident T2D and to compare it with the prediction of traditional risk factors and the FINDRISK. Subsequently, the equality between the different ROC curve areas obtained was tested. In addition, to validate model performance, the Net Reclassification Improvement (NRI) and the Integrated Discrimination Improvement (IDI) were also assessed. Finally, sensitivity analyses were performed, excluding subjects with prediabetes at baseline. Two-tailed p-values < 0.05 will be considered statistically significant. The calculations and figures were made using STATA v.18.0 for Mac (StataCorp LP, College Station, TX, USA) and GraphPad Prism software v 6.0 for Mac (GraphPad Software Inc., San Diego, CA, USA).

Results

The baseline clinical and metabolic characteristics of subjects included in the study are shown in Tables 1 and 2 for the baseline and the follow-up study, respectively. The mean age was about 50 years, and 61% of the subjects were females. Obesity was found in 26% of the participants, 12% had prediabetes, and 50% presented a first-degree family history of T2D. As mentioned, we included 2,900 subjects (659 subjects with T2D at baseline [prevalence study] and 2,241 subjects free from T2D at baseline -subjects at-risk- who participated in the follow-up [incidence study]). Of the 2,241 participants included in the follow-up study who were at risk at baseline, 149 developed T2D (6,6%) during or at the end of the 7.5 (7.1–7.9)

Table 1 Baseline clinical and metabolic characteristics of the subjects included in the study

	BASELINE		
	No T2D	Prevalent T2D	*p*
n	2241	659	-
Age (years)	47.8 (14.8)	64.7 (12.0)	< 0.001
Female (n, %)	1357 (61%)	309 (47%)	< 0.001
Current smokers (n, %)	580/2235 (26%)	112/658 (17%)	< 0.001
Educational level (n, %)			< 0.001
No studies	168/2234 (8%)	200/658 (30%)	
Elementary	1065/2234 (48%)	334/658 (51%)	
Secondary	608/2234 (27%)	90/658 (14%)	
University	393/2234 (18%)	34/658 (5%)	
Family history of T2D	1121 (50%)	491 (75%)	< 0.001
MedDiet Score	17 (14–21)	17 (14–21)	NS
High-fiber Food Score	8 (7–10)	8 (7–10)	NS
Hypertension	834/2233 (37%)	538/656 (82%)	< 0.001
Dyslipidaemia (n, %)	1011 (45%)	491 (75%)	< 0.001
Obesity (n, %)	575/2226 (26%)	354/653 (54%)	< 0.001
BMI (kg/m^2)	27.5 (4.7)	30.9 (5.3)	< 0.001
Prediabetes (n, %)	267 (12%)	-	-
Glucose at 0′ (mmol/L)	5.1 (0.7)	8.0 (3.0)	< 0.001
Glucose at 120′ (mmol/L)	5.7 (1.7)	12.8 (4.1)	< 0.001
HOMA-IR	2.0 (1.4)	6.2 (17.6)	< 0.001

All values are given as percentage, mean (SD), or median (interquartile range).
Abbreviations: *T2D* Type 2 diabetes, *MedDiet* Mediterranean Diet, *BMI* Body mass index, *HOMA-IR* Homeostatic Model Assessment for Insulin Resistance, *NS* Not significant.

Table 2 Baseline clinical and metabolic characteristics of the subjects included in the follow-up study

	FOLLOW-UP		
	No T2D	**Incident T2D**	***p***
n	2092	149	-
Age (years)	47.2 (14.8)	56.0 (12.1)	< 0.001
Female (n, %)	1279 (61%)	78 (52%)	< 0.001
Current smokers (n, %)	542/2087 (26%)	38/148 (26%)	NS
Educational level (n, %)			< 0.001
No studies	146/2086 (7%)	22/148 (15%)	
Elementary	975/2086 (47%)	90/148 (61%)	
Secondary	585/2086 (28%)	23/148 (16%)	
University	380/2086 (18%)	13/148 (9%)	
Family history of T2D	1028 (49%)	93 (62%)	< 0.001
MedDiet Score	17 (14–21)	17 (15–20)	NS
High-fiber Food Score	8 (7–10)	8 (7–9)	NS
Hypertension	743/2085 (36%)	91/148 (62%)	< 0.001
Dyslipidaemia (n, %)	916 (44%)	95 (64%)	< 0.001
Obesity (n, %)	493/2079 (24%)	82/147 (56%)	< 0.001
BMI (kg/m^2)	27.3 (4.6)	31.1 (5.0)	< 0.001
Prediabetes (n, %)	187 (9%)	80 (54%)	< 0.001
Glucose at 0′ (mmol/L)	5.1 (0.6)	5.7 (0.7)	< 0.001
Glucose at 120′ (mmol/L)	5.5 (0.6)	7.5 (1.8)	< 0.001
HOMA-IR	1.9 (1.3)	3.1 (1.8)	< 0.001

All values are given as percentage, mean (SD), or median (interquartile range).
Abbreviations: *T2D* Type 2 diabetes, *MedDiet* Mediterranean Diet, *BMI* body mass index, *HOMA-IR* Homeostatic Model Assessment for Insulin Resistance, *NS* Not significant.

follow-up period. At baseline, subjects who developed T2D were older and with a higher proportion of males. In addition, they had a higher prevalence of first-degree family history of T2D, arterial hypertension, dyslipidemia, and obesity. As expected, those subjects with prediabetes presented a higher probability of developing T2D compared to those subjects with normal glucose metabolism at inclusion (Table 2).

Association between baseline circulating concentrations of SCFA and baseline food consumption patterns

To further examine the relationship between baseline circulating concentrations of SCFA and baseline food consumption patterns and to explore non-linear associations, we categorized the sample according to tertiles of the MedDiet and High-Fiber Food scores. In general terms, the baseline circulating concentrations of SCFA were associated with the consumption of high-fiber foods at baseline. As shown in Fig. 2, the baseline circulating concentrations of acetate [27.3 (20.5–36.2) μmol/L vs. 28.0 (20.6–37.6) μmol/L vs. 31.0 (22.8–40.4) μmol/L for the 1st, 2nd and 3rd tertiles respectively; $p < 0.001$] (Fig. 2, panel A), propionate [3.3 (1.6–4.1) μmol/L vs. 3.4 (2.2–4.2) μmol/L vs. 3.6 (2.9–4.4) μmol/L; $p < 0.001$] (Fig. 2, panel B), butyrate [11.0 (1.0–14.0) μmol/L vs. 11.5 (1.5–13.9) μmol/L vs. 11.9 (8.9–14.5) μmol/L; $p < 0.001$] (Fig. 2, panel C), and isobutyrate [0.88 (0.43–1.09) μmol/L vs. 0.91 (0.58–1.12) μmol/L vs. 0.95 (0.76–1.15) μmol/L; $p < 0.001$] (Fig. 2, panel D) increased in parallel with the increase in the baseline High-Fiber Food score tertiles. Additionally, baseline circulating concentrations of acetate increased in parallel with the increase in the baseline MedDiet Score tertiles [28.0 (20.7–37.6) μmol/L vs. 28.0 (20.7–37.4) μmol/L vs. 29.7 (21.5–39.2) μmol/L for the 1st, 2nd and 3rd tertiles respectively; $p = 0.037$] (Fig. 3, panel A). However, no differences were found between tertiles for the rest of the baseline circulating concentrations of SCFA (Fig. 3, panels B, C and D).

Association between circulating concentrations of SCFA and T2D at baseline

At baseline, the circulating concentrations of propionate [3.4 (2.1–4.1) μmol/L in no T2D vs. 3.6 (2.4–4.5) μmol/L; $p < 0.001$] (Fig. 4, panel B) and isobutyrate [0.90 (0.59–1.10) μmol/L vs. 0.94 (0.59–1.19) μmol/L; $p = 0.001$] (Fig. 4, panel D) were higher in those subjects with pre-existing T2D, with no differences for circulating concentrations of acetate [28.6 (21.0–38.1) μmol/L vs. 27.7 (20.8–36.9) μmol/L; not significant (NS)] (Fig. 4, panel A) or butyrate [11.4 (1.6–14.2) μmol/L vs. 11.3 (1.4–14.0) μmol/L; NS] (Fig. 4, panel C). In the same line, the concentrations of these two metabolites increased in parallel with the carbohydrate metabolism status if we compared subjects with normal glucose with those with prediabetes and those with T2D [propionate: 3.4 (2.2–4.1) μmol/L vs. 3.6 (2.0–4.3) μmol/L vs. 3.6 (2.4–4.5) μmol/L; $p < 0.001$; isobutyrate: 0.90 (0.59–1.09) μmol/L vs. 0.93 (0.64–1.17) μmol/L vs. 0.94 (0.59–1.19) μmol/L; $p = 0.001$] (Fig. 5, panels B and D respectively).

Association between baseline circulating concentrations of SCFA and T2D at 7-year follow-up

We further explored the potential association between baseline circulating concentrations of SCFA and incident T2D. Up to 2,241 subjects of the 2,408 subjects at-risk from the baseline study [prevalence study] who participated in the follow-up study [incidence study] had available samples for the metabolomics analysis (93.1% of the Di@bet.es follow-up study). Subjects who developed T2D (incident cases) after a median follow-up of 7 years had higher baseline circulating concentrations of butyrate [11.4 (1.6–14.1) μmol/L vs. 12.3 (6.9–15.2) μmol/L; $p = 0.040$] (Fig. 6, panel C) and isobutyrate [0.90 (0.58–1.09) μmol/L vs. 0.96 (0.72–1.23) μmol/L, $p = 0.004$]

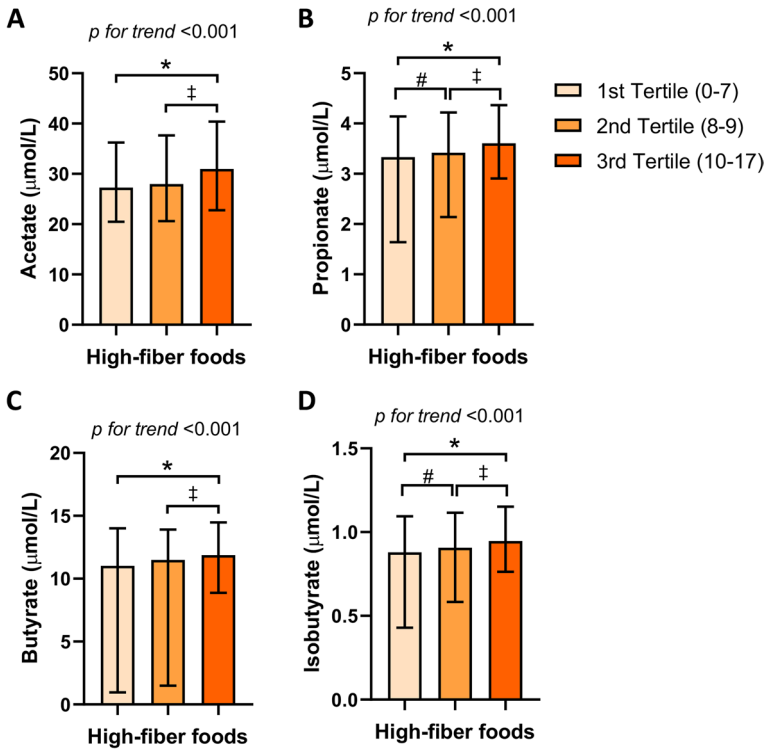

Fig. 2 Circulating concentrations of SCFA and High-Fiber Food Score. Circulating concentrations of acetate (panel A), propionate (panel B), butyrate (panel C), and isobutyrate (panel D) at baseline according to tertiles of the High-Fiber Food Score. The baseline circulating concentrations of acetate, propionate, butyrate, and isobutyrate increased in parallel with the increase of baseline High-Fiber Food score tertiles. Data are presented as median ± interquartile ranges. Kruskal–Wallis test was used to compare the three groups, and the Dunn's test was used for post hoc analyses for multiple comparisons. # $p < 0.05$ for 2n tertile compared with 1st tertile; * $p < 0.05$ for 3rd tertile compared with 1st tertile; and ‡ $p < 0.05$ for 3rd tertile compared with 2nd tertile

(Fig. 6, panel D) with no differences for baseline circulating concentrations of acetate [28.6 (21.0–38.4) μmol/L vs. 29.1 (20.6–36.7) μmol/L; NS] or propionate [3.4 (2.1–4.1) μmol/L vs. 3.6 (2.6–4.2) μmol/L; NS) (Fig. 6, panels A and B, respectively).

Baseline circulating concentrations of SCFA as a predictive biomarker of incident T2D

We developed logistic regression models to evaluate the independent factors associated with incident T2D at 7 years follow-up. In the univariate analyses, age (OR 1.04 [95%CI: 1.03–1.05]; $p < 0.001$), male (OR 1.43 [95%CI: 1.03–1.99]; $p < 0.001$), family history of T2D (OR 1.43 [95%CI: 1.18–1.74]; $p < 0.001$), obesity (OR 4.06 [95%CI: 2.89–5.71]; $p < 0.001$), arterial hypertension (OR 2.88 [95%CI: 2.04–4.06]; $p < 0.001$), dyslipidemia (OR 2.26 [95%CI: 1.60–3.19]; $p < 0.001$), prediabetes (OR 11.81 [95%CI: 8.28–16.85]; $p < 0.001$), glucose concentrations at 0' (OR 6.20 [95%CI: 4.52–8.50]; $p < 0.001$) and 120' (OR 1.89 [95%CI: 1.67–2.12]; $p < 0.001$) post-OGTT, insulin resistance (OR 1.42 [95%CI: 1.29–1.56]; $p < 0.001$), assessed by HOMA-IR, and butyrate (OR 1.93 [95%CI: 1.24–3.01]; $p < 0.001$) at baseline were positively associated with incident T2D. On the contrary, the educational level at baseline (OR 0.26 [95%CI: 0.14–0.48]; $p < 0.001$ and OR 0.23 [95%CI: 0.11–0.47]; $p < 0.001$ for secondary and university studies, respectively) was inversely associated with incident T2D. No association was found between baseline dietary patterns (adherence to MedDiet or consumption of high-fiber foods) in the univariate analyses. In the multivariate analysis, baseline circulating concentrations of butyrate (OR 1.03 [95%CI: 1.01–1.07]; $p = 0.037$) and isobutyrate (OR 1.9 [95%CI: 1.2–3.1]; $p = 0.009$) were independently associated with development of T2D after adjusting for age, sex, family history of T2D, prediabetes, obesity, hypertension and dyslipidemia at baseline (Table 3).

ROC curve analyses were performed to explore the ability of baseline circulating concentrations of SCFA to improve the prediction of the incidence of T2D at 7-year follow-up beyond traditional risk factors. The C-statistics for predicting incident T2D was 0.840 [95%CI: 0.807–0.873] for the model including conventional risk factors (model 1: age, sex, family history

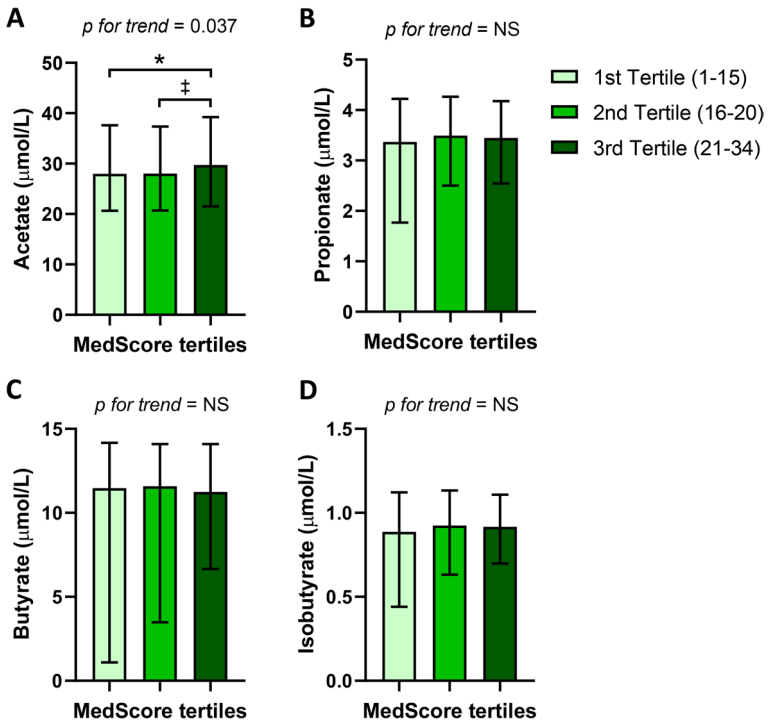

Fig. 3 Circulating concentrations of SCFA and MedDiet adherence. Circulating concentrations of acetate (panel A), propionate (panel B), butyrate (panel C), and isobutyrate (panel D) at baseline according to tertiles of the Mediterranean diet score. Baseline circulating concentrations of acetate increased in parallel with the increase of baseline Mediterranean diet score tertiles. Data are presented as median ± interquartile ranges. Kruskal–Wallis test was used to compare the three groups, and the Dunn's test was used for post hoc analyses for multiple comparisons. # $p < 0.05$ for 2n tertile compared with 1st tertile; *$p < 0.05$ for 3rd tertile compared with 1st tertile; and ‡$p < 0.05$ for 3rd tertile compared with 2nd tertile. MedDiet: Mediterranean diet (MedDiet); NS: non-significant

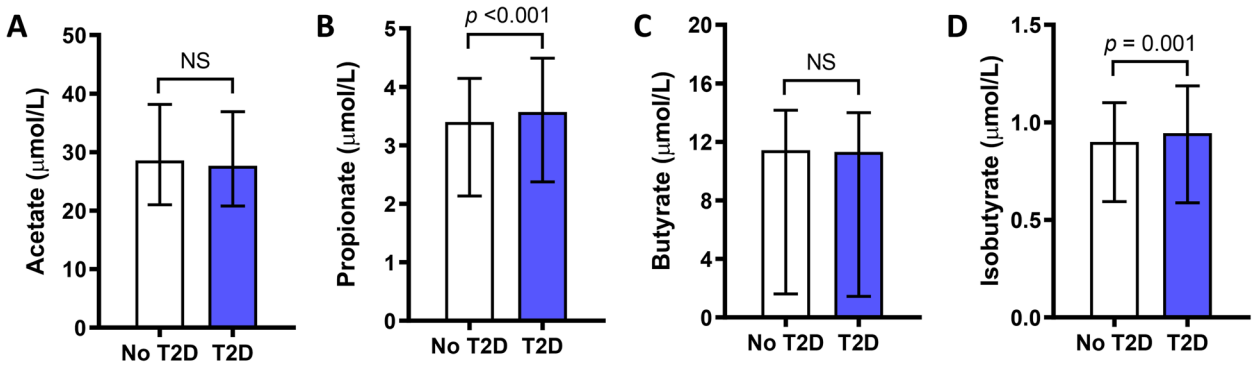

Fig. 4 Circulating concentrations of SCFA and type 2 diabetes at baseline. Circulating concentrations of acetate (panel A), propionate (panel B), butyrate (panel C), and isobutyrate (panel D) for subjects without T2D (n = 2,241) vs. subjects with T2D (n = 659) at baseline. Baseline circulating concentrations of propionate and isobutyrate were higher in those subjects with pre-existing T2D, with no differences for circulating concentrations of acetate or butyrate. Data are presented as median ± interquartile ranges. Mann–Whitney test was used to compare the two groups. NS: non-significant; T2D: type 2 diabetes

of T2D, prediabetes, obesity, hypertension, and dyslipidemia at baseline), 0.847 (95%CI: 0.816–0.877) for butyrate (model 1 + baseline butyrate) and 0.843 (95%CI: 0.812–0.875) for isobutyrate (model 1 + baseline isobutyrate) (Fig. 7). Thus, both baseline circulating concentrations of butyrate (0.847 vs. 0.840; $p = 0.149$) and isobutyrate (0.843 vs. 0.840; $p = 0.390$), when added to traditional risk factors, offered a good ability to predict T2D, although similar to the reference model based only on baseline clinical factors. Finally, to assess the

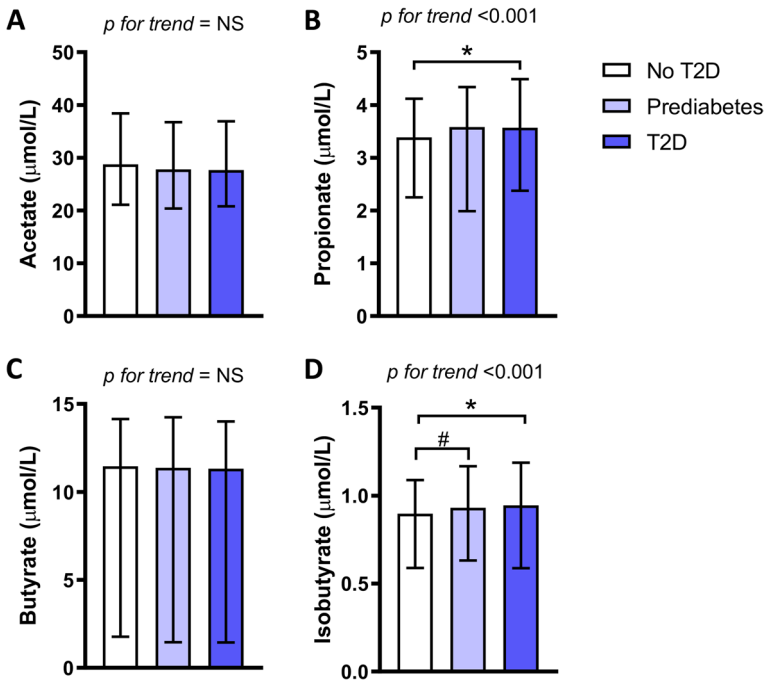

Fig. 5 Circulating concentrations of SCFA and prediabetes and type 2 diabetes at baseline. Circulating concentrations of acetate (panel A), propionate (panel B), butyrate (panel C), and isobutyrate (panel D) for subjects without T2D (n = 1,974) and subjects with prediabetes (n = 267) or T2D (n = 659) at baseline. Concentrations of propionate and isobutyrate increased in parallel with the carbohydrate metabolism status. Data are presented as median ± interquartile ranges. Kruskal–Wallis test was used to compare the three groups, and and the Dunn's test was used for post hoc analyses for multiple comparisons. $^{\#}p < 0.05$ for prediabetes compared with no T2D; $^{*}p < 0.05$ for T2D compared with no T2D; and $^{\ddagger}p < 0.05$ for T2D compared with prediabetes. NS: non-significant; T2D: type 2 diabetes

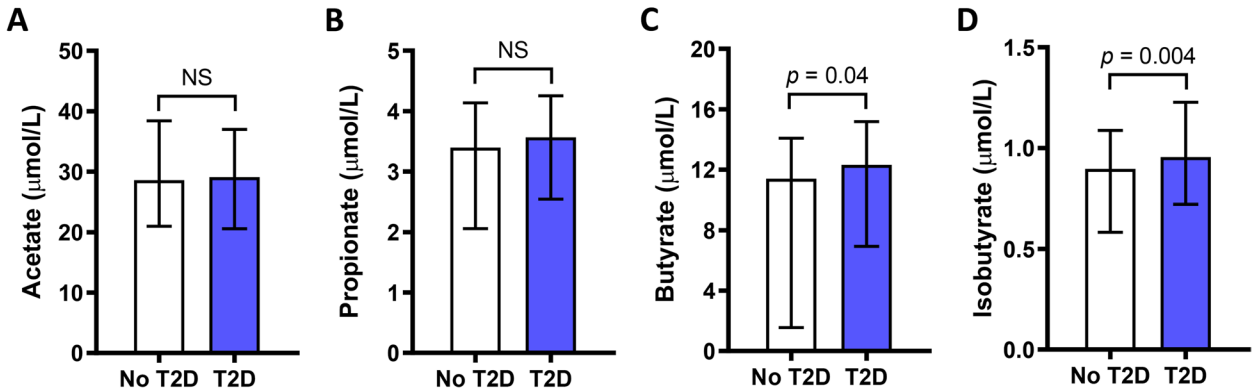

Fig. 6 Circulating concentrations of SCFA and type 2 diabetes at 7-year follow-up. Circulating concentrations of acetate (panel A), propionate (panel B), butyrate (panel C), and isobutyrate (panel D) for subjects without T2D (n = 2,092) vs. subjects with incident T2D (n = 149) after 7 years of follow-up. Subjects who developed T2D after a median follow-up of 7 years had higher baseline circulating concentrations of butyrate and isobutyrate with no differences for baseline circulating concentrations of acetate or propionate. Data are presented as median ± interquartile ranges. Mann–Whitney test was used to compare the two groups. NS: non-significant; T2D: type 2 diabetes

accuracy of baseline circulating concentrations of SCFA as a biomarker to predict incident T2D, we compared the accuracy of the existing risk score (FINDRISK) for predicting the risk of T2D with the models, including baseline circulating concentrations of butyrate and isobutyrate, respectively. Accordingly, the FINDRISK score was calculated for each patient, and the ROC curve was developed. The C-statistics for the prediction of T2D was 0.785 [95%CI: 0.746–0.823] for the FINDRISK (Fig. 7). Thus, both models, including baseline

Table 3 Association between baseline circulating concentrations of SCFA and incident T2D

	Unadjusted		Adjusted	
Incident T2D	**OR (CI 95%)**	***p***	**OR (CI 95%)**	***p***
Acetate	1.00 (0.98–1.01)	NS	-	-
Propionate	1.09 (0.97–1.22)	NS	-	-
Butyrate	1.93 (1.24–3.01)	0.004	1.03 (1.00–1.07)	0.037
Isobutyrate	1.02 (0.99–1.05)	NS	1.02 (1.18–3.12)	0.009

Abbreviations: T2D type 2 diabetes, *OR* Odds Ratio, *CI* Confidence interval, *NS* Non-significant.

circulating concentrations of butyrate (0.847 vs. 0.785; $p<0.001$) and isobutyrate (0.843 vs. 0.785; $p<0.001$), improved the risk prediction compared to FINDRISK. Baseline dietary patterns, including adherence to MedDiet or consumption of high-fiber foods, did not add any prediction value.

In addition, to validate the model performance, the NRI and IDI were also calculated. In the first comparison (model 1 vs. model 1 + baseline butyrate), there was a non-significant improvement in predicting the risk of T2D, with a positive 1.51% NRI and an IDI of 0.34%. The reclassification of T2D cases showed a net gain of 2 correct classifications out of 142 cases (1.4% improvement), and the reclassification of non-T2D cases similarly showed a net gain of 2 correct classifications out of 2042 non-T2D cases (0.1% improvement). In the second comparison (model 1 vs. model 1 + baseline isobutyrate), there was a slight, non-significant decline in the prediction of the risk of T2D, with an NRI of -0.70% and an IDI of 0.06%. Here, the reclassification of T2D cases resulted in 1 fewer correct classification out of 142 cases (0.7% decline), while the reclassification of non-T2D cases showed no net change. In the third comparison (FINDRISK vs. model 1 + baseline butyrate), the model demonstrated a statistically significant improvement in the prediction of the risk of T2D, with a NRI of 12.6% and an IDI of 5.19%. This improvement was driven by a net gain of 9 correct reclassifications in T2D cases out of 142 (6.3% improvement) and a net gain of 128 additional correct reclassifications out of 2042 non-T2D cases (a 6.3% improvement). In the fourth comparison (FINDRISK vs. model 1 + baseline isobutyrate), the model also showed a statistically significant improvement in the prediction of the risk of T2D, with an 8.65% NRI and a 4.97% IDI. This improvement consisted in a net reclassification gain of 4 detected out of 142 cases (2.8% improvement on T2D reclassification) and a net reclassification gain of 119 out of 2042 non-T2D cases (5.8% improvement on non-T2D reclassification).

Analysis of Sensitivity

Sensitivity analyses were performed to avoid potential reverse causality, excluding subjects with prediabetes ($n=267$, 12%) at baseline. From the 1,974 subjects with normal glucose metabolism at inclusion, $n=69$ (3.5%) developed T2D during the study follow-up. As

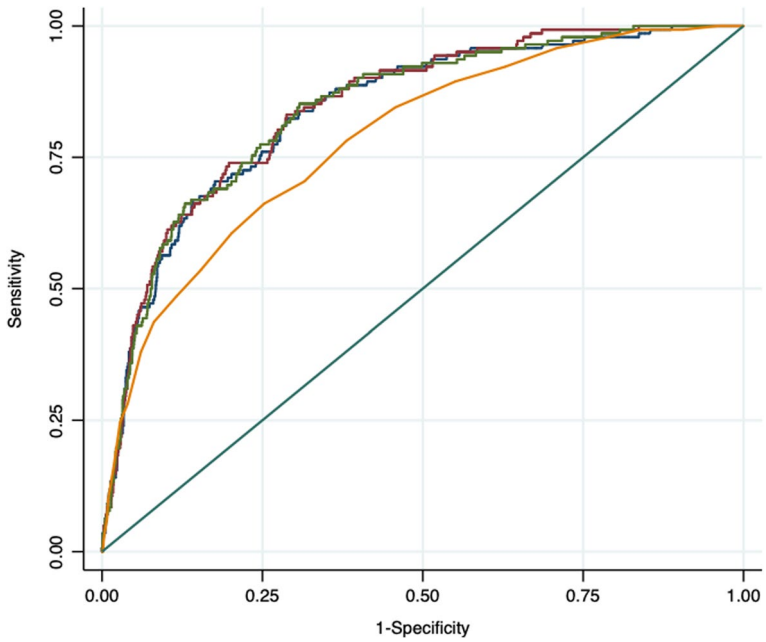

Fig. 7 ROC curves for prediction of incident type 2 diabetes. The ROC curves for conventional risk factors (blue: model 1), butyrate (red: model 1 + butyrate), isobutyrate (green: model 1 + isobutyrate), and FINDRISK (yellow) in predicting incident type 2 diabetes after a 7-year follow-up

previously, after a multivariate adjustment, the association between baseline circulating concentrations of butyrate (OR 1.07 [95%CI: 1.02–1.12]; $p=0.003$) and isobutyrate (OR 3.00 [95%CI: 1.54–5.85]; $p=0.001$) with incident T2D remained unchanged or strengthened, as compared with the analyses including subjects with prediabetes. The C-statistics for predicting incident T2D was 0.776 [95%CI: 0.725–0.827] for the model including conventional risk factors (model 1: age, sex, family history of T2D, obesity, hypertension, and dyslipidemia at baseline), 0.789 [95%CI: 0.740–0.839] for butyrate (model 1 + baseline butyrate) and 0.791 [95%CI: 0.743–0.840] for isobutyrate (model 1 + baseline isobutyrate) and 0.676 [95%CI: 0.619–0.734] for the FINDRISK. Thus, both baseline circulating concentrations of butyrate (0.775 vs. 0.792; $p=0.238$) and isobutyrate (0.784 vs. 0.771; $p=0.306$), when added to baseline traditional risk factors, offered a good ability to predict T2D, although similar to the reference model and improved significantly the risk prediction compared to FINDRISK (butyrate: 0.775 vs. 0.676, $p<0.001$; and isobutyrate: 0.784 vs. 0.676, $p<0.001$; respectively).

In addition, when the sensitivity analyses were performed, an association between baseline consumption of high-fiber foods (OR 0.886 [0.792–0.990], $p=0.033$) and incident T2D emerged in the univariate analyses. Thus, we added the baseline consumption of high-fiber foods to the previously developed models. When baseline consumption of high-fiber foods was added to the model, it improved the prediction of T2D risk, although the differences were not statistically significant to the previous models. The C-statistics for predicting incident T2D was 0.805 [95%CI: 0.751–0.856] for the model including traditional risk factors (model 1: age, sex, family history of T2D, obesity, hypertension dyslipidemia + consumption of high-fiber foods at baseline), 0.812 [95%CI: 0.761–0.862] for butyrate (model 1 + consumption of high-fiber foods at baseline + baseline butyrate) and 0.812 [95%CI: 0.758–0.866] for isobutyrate (model 1 + consumption of high-fiber foods at baseline + baseline isobutyrate).

Discussion

This study provides new evidence on the association between SCFAs and the risk of developing T2D. Elevated baseline circulating concentrations of butyrate and isobutyrate were identified as independent predictors of T2D incidence over a median follow-up period of seven years. These associations remained significant after adjustment for traditional risk factors, including age, sex, family history of T2D, prediabetes, obesity, hypertension, and dyslipidemia. The findings of this study are consistent with the established role of SCFAs as key metabolites produced by the gut microbiota through the anaerobic fermentation of dietary fibers. Previous research has underscored the critical role of SCFAs in maintaining gut health and influencing systemic metabolic processes [4]. This study builds upon existing knowledge by identifying an association between SCFAs and the risk of T2D, emphasizing their potential utility as biomarkers for disease prediction.

Our findings further demonstrate a positive association between the consumption of high-fiber foods at baseline and baseline circulating concentrations of SCFA, which agrees with previous studies that identify dietary fibers as the primary substrates for SCFA production [4]. In this study, higher intake of high-fiber foods at baseline was associated with higher baseline circulating concentrations of acetate, propionate, butyrate, and isobutyrate, emphasizing the critical role of dietary patterns in modulating SCFA levels. Moreover, adherence to the MedDiet has been linked to elevated SCFA concentrations and inversely associated with markers of intestinal permeability, such as lipopolysaccharide-binding protein and zonulin, suggesting that SCFAs may serve as a mechanistic link between diet and the maintenance of intestinal barrier integrity [22]. Notably, while baseline circulating concentrations of acetate increased with the increase of the adherence to the MedDiet at baseline, other SCFAs did not exhibit a similar pattern, indicating that specific dietary fibers may differentially influence SCFA production [22]. This differential effect may be attributed to the distinct fermentation properties of individual fibers, their interactions with specific gut microbiota compositions, and the varying affinities of SCFAs for their receptors, such as FFA2 and FFA3, as suggested by Ulven et al. [5]. These findings highlight the complexity of the interplay between dietary fibers, SCFA production, and gut microbiota.

Observational epidemiological and interventional clinical studies indicate that both fiber consumption [23–26] and the adherence to a MedDiet [27–29] are causally related to a higher risk of T2D. Nevertheless, although there is an association between the dietary patterns and the circulating concentrations of SCFA, we did not find any association between the consumption of high-fiber foods and the adherence to MedDiet with T2D prediction, which is probably a paradoxical result. However, we found an association between baseline consumption of high-fiber foods and incident T2D when performing the sensitivity analyses (excluding subjects with prediabetes). In this sense, some authors have suggested that the effect of SCFA on the substrate and energy metabolism may differ between metabolic phenotypes [30–32]. Canfora et al. demonstrated that supplementation of a fiber mixture increased distal colonic bacterial fermentation

in lean individuals, which improved metabolic health parameters. However, in subjects with overweight/obese and prediabetes, the supplementation with a fiber mixture increased circulating acetate concentrations. Still, it did not improve metabolic health parameters [33]. Therefore, the observed differences in behaviour between healthy people and those with prediabetes may be consistent with the findings presented in this study.

The potential of butyrate and isobutyrate as predictive biomarkers for T2D is particularly compelling. Both baseline circulating concentrations of SCFAs demonstrated strong predictive capabilities for T2D, with C-statistics comparable to traditional clinical risk factors. However, their inclusion in a conventional risk model resulted in only marginal improvements in predictive accuracy, which were not statistically significant. Despite this, both baseline circulating concentrations of SCFAs significantly enhanced risk prediction compared to the FINDRISK Score, emphasizing their potential clinical utility. These observations align with those of Husted et al. [2], who highlighted the role of metabolites as signaling molecules capable of modulating metabolic pathways and influencing disease outcomes. Integrating SCFAs into existing predictive models could offer a more refined assessment of T2D risk, particularly for individuals with borderline or ambiguous clinical profiles. Such an approach underscores the relevance of these metabolites as complementary biomarkers, providing additional insights beyond those offered by traditional risk factors [2].

Although the bulk of evidence suggests that increased SCFA production benefits the host by exerting antiobesity and antidiabetic effects [34, 35], some in vitro and in vivo studies have indicated that overproduction or accumulation of SCFAs in the bowel may also lead to obesity or T2D, owing to increased energy accumulation, higher capacity to harvest energy through SCFA production or increased de novo lipogenesis [9, 36–38]. In our study, baseline circulating concentrations of SCFA were higher in individuals with pre-existing T2D and those who developed T2D during the follow-up period. This observation aligns with prior studies [9, 36], which reported elevated fecal SCFA levels in overweight, obese, and subjects with T2D. However, focusing on circulating concentrations of SCFA offers a more precise representation of in vivo fermentation processes, as fecal SCFA levels may not fully capture the systemic metabolic state. Elevated baseline circulating concentrations of butyrate and isobutyrate, in particular, may signify an adaptive response to the metabolic dysregulation associated with T2D, highlighting the intricate interactions between gut-derived metabolites and host metabolism. These findings underscore the need for further research into the underlying mechanisms of this relationship.

Additionally, the results are consistent with Zhao et al. findings, who observed a significant increase in SCFA and zonula occludens-1 concentrations (used for assessing tight junctions integrity and regulation of intestinal permeability) in subjects with T2D [39]. The authors hypothesised that the intestinal barrier destruction, driven by hyperglycemia, increased the passive reabsorption of SCFAs, which resulted in the rising circulating concentrations of SCFA in subjects with T2D. Thus, the study suggested that excessive SCFA absorption may occur via a leaky gut caused by gut microbiota dysbiosis, leading to higher levels of circulating SCFAs in subjects with T2D and that the leaky gut might be caused by the disordered gut microbiota [39]. These observations further emphasize the role of SCFAs in metabolic dysregulation and gut barrier integrity, providing important insights into the pathophysiology of T2D.

The association of SCFAs with the development of T2D underscores the pivotal role of gut microbiota and their metabolic products in maintaining metabolic health. The independent predictive value of butyrate and isobutyrate warrants further investigation into the mechanistic pathways through which these SCFAs influence glucose metabolism and insulin sensitivity. The primary fermentation products resulting from digestion of dietary fibre in the intestine are SCFA, which are widely recognized for their role as key energy sources in metabolic pathways. Normal colonic epithelia obtain 60–70% of their energy supply from SCFA, particularly butyrate. The liver primarily absorbs propionate, serving as a precursor for gluconeogenesis, liponeogenesis and protein synthesis. Acetate enters systemic circulation to be metabolised by peripheral tissues and acts as a substrate for cholesterol synthesis. An excess production of SCFA from dietary compounds, potentially caused by alterations of gut microbiota, may escape digestion in the small intestine, therby providing the host with an additional energy source [9].

However, the potential of SCFAs to act as extracellular signalling molecules also offers a plausible explanation for their involvement in modulating metabolic and inflammatory responses, which are key contributors to T2D pathogenesis [3]. Notably, SCFAs, particularly butyrate, are known to activate GPCRs such as FFA2 and FFA3, which are expressed in several tissues, including human colonic tissues, as well as pancreatic islets, adipose tissue, skeletal muscle, liver, and immune cells. These receptors mediate various metabolic and immune processes, suggesting that SCFAs may influence T2D development through multiple mechanisms [3, 6, 7]. Observational and interventional studies have shown that SCFAs affect a range of host systems and energy homeostasis by activating GPCRs, preserving intestinal integrity and barrier

function, influencing immune regulation and inflammatory responses (by the reduction of the secretion of pro-inflammatory cytokines and chemokines), and modulating the enteric nervous system and brain-gut axis. Furthermore, SCFAs may play crucial roles in maintaining glucose homeostasis and insulin sensitivity by: 1) in pancreatic islets, by regulating β-cell insulin secretion; 2) in adipose tissue, by enhancing lipid buffering capacity, increasing the oxidation of fatty acids in brown adipose tissue, facilitating the browning and mitochondrial function of white adipose tissue, diminishing the size of adipocytes and by the inhibition of endogenous lipolysis within adipocytes; 3) in the liver, by decreasing glycolysis and gluconeogenesis, and increasing glucose uptake and glycogen synthesis and fatty acids oxidation and improving mitochondrial function; and 4) in the skeletal muscle by improving glucose uptake and by the inhibition of ectopic lipid storage in skeletal muscle [40, 41]. Additionally, the role of SCFAs in modulating incretin hormones, essential to maintain glucose metabolism, offers further insights into the connection between gut microbiota-derived metabolites and T2D. SCFAs can also stimulate the local release of satiety hormones such as GLP-1 and peptide YY (PYY) from enteroendocrine L-cells via the activation of GPR41 or GPR43 [40, 41], while in circulation, they epigenetically regulate the expression of adipokines such as leptin, adiponectin, and resistin. Recent evidence indicates that circulating SCFAs are associated with GLP-1 concentrations, whole-body lipolysis, and peripheral insulin sensitivity in humans [42]. This emphasizes the therapeutic potential of targeting gut microbiota and their metabolites in T2D prevention and management.

Recent evidence indicates that SCFA may exert divergent effects depending on the metabolic context. It has been hypothesized that intestinally produced SCFAs are processed differently in insulin-resistant states, potentially influencing glucose and lipid metabolism [43]. Tang et al. demonstrated that, under diabetic conditions, elevated SCFA concentrations interact with FFA2 and FFA3, impairing glucose-stimulated insulin secretion. Furthermore, they suggested that antagonists targeting FFA2 and FFA3 may enhance insulin secretion in people with T2D [44]. These findings highlight the bidirectional relationship between gut microbiota-derived SCFAs and T2D, underscoring the complexity of this emerging field. Despite these advancements, numerous questions remain unanswered, necessitating further research to elucidate the role of SCFAs in glucose regulation in T2D.

Although this study offers valuable insights, certain limitations should be acknowledged. As previously explained, the Di@bet.es study was the first nationwide, prospective, population-based study in Spain designed to determine the prevalence and incidence of T2D [10, 11]. Subjects were assessed at baseline (prevalence study) and a second time for the follow-up study (incidence study) but were not followed periodically between these two time points. Thus, as patients were evaluated only at the start and end of the study (after 7 years), the exact time of our endpoint (new cases of T2D) was unknown. The absence of gut microbiota analysis, resulting from the unavailability of fecal samples, restricts the ability to directly associate specific bacterial populations with SCFA production and the risk of T2D. In addition, we could not differentiate dietary vs. host-derived SCFA sources. Thus, the absence of gut microbiota data and the inability to differentiate dietary vs. host-derived SCFA sources limits our capacity to explore deeply the mechanistic insights of the role of SFCA in the development of T2D. Furthermore, the high attrition rate observed in the follow-up cohort (45%) raises concerns about potential selection bias. Nevertheless, we found few differences between people who participated in the follow-up and those who did not, as previously published [10], and thus, the possible participation bias would be minimal. While using an untargeted metabolomics approach provides a comprehensive overview, it is inherently less sensitive and specific than targeted analyses, which may impact the precision and reproducibility of the results. Finaly, the Di@bet.es Study administered a qualitative FFQ, which has been previously used in a population-based study in southern Spain [12, 45]. We are aware that ad hoc validation remains ideal. Nevertheless, the use of FFQs in the Di@bet.es study without specific ad hoc validation for the Spanish population is supported by robust scientific evidence and methodological strategies to address applicability and recall bias [46–48].

Conclusions

In conclusion, the findings of this study underscore a significant association between baseline circulating concentrations of SCFAs and the risk of developing T2D, emphasizing the critical role of dietary fiber intake in modulating these metabolites. Although baseline circulating concentrations of SCFAs alone may not dramatically improve risk prediction beyond traditional clinical factors, they provide valuable insights into the intricate metabolic interplay between diet, gut microbiota, and host health. Future research should incorporate gut microbiota profiling alongside targeted SCFA analyses to further unravel these complex relationships and facilitate the development of personalized dietary interventions for T2D prevention. These findings contribute to the expanding body of evidence on the pivotal role of gut microbiota-derived metabolites in metabolic health and disease, offering a foundation for innovative approaches to T2D risk assessment and management.

Abbreviations

T2D	Type 2 diabetes mellitus
SCFA	Short chain fatty acid
GPCR	G-protein-coupled receptor
GLP-1	Glucagon-like peptide-1
OGTT	Oral glucose tolerance test
BMI	Body mass index
WHR	Waist-to-hip ratio
HOMA-IR	Homeostatic Model Assessment for Insulin Resistance
FFQ	Food frequency questionnaire
MedDiet	Mediterranean diet
FINDRISK	Finnish Diabetes Risk Score

Acknowledgements
We wish to acknowledge the participating health centers' primary care managers and personnel, and all the nurses and dietitians who made Di@bet.es study fieldwork possible. We especially thank all the people who voluntarily participated in the study. We are indebted to the IDIBAPS Biobank, integrated into the Spanish National Biobank Network, for the sample and data procurement. Finally, we thank Antoni del Pino and Marc Riu from the Metabolomics facility of the Centre for Omic Sciences (COS) Joint Unit of the Universitat Rovira i Virgili-Eurecat, for their contribution to gas chromatography-mass spectrometry analysis.

Authors' contributions
GL, LC, and JV were responsible for drafting and writing the manuscript. Data analysis was conducted by GL, EC, JB, JF-LR, and MG-C, while OY and MV carried out the sample analyses. GL, GR-M, DM, SF-V, and JV conceptualized, designed, and performed the research study. All authors read and approved the final manuscript.

Authors' Twitter handles
GL: @LlauradoGemma; EC: @elisendacb; JF-LR: @94066Flores; OY: @yaneslab; MV: @MarionaVinaixa; DM: @DidacMauricio; SF-V: @S_FdezVeledo; JV: @: jvo2020.

Funding
The authors acknowledge the financial support provided through the project PI20/00312 (GL), funded by the Instituto de Salud Carlos III (ISCIII) and co-funded by the European Union, as well as through the project "XII Ayudas SED a Proyectos de Investigación Clínica en Diabetes dirigidos por jóvenes investigadores," funded by the Sociedad Española de Diabetes (GL). Research conducted in the authors' laboratory is supported by grants from the ISCIII (PI20/00338 and PI23/01133 to JV; PI18/0328 and PI21/01163 to DM) and the Spanish Ministry of Science and Innovation MCIN/AEI/https://doi.org/10.13039/501100011033(RTI2018-093919-B-100 and PID2021-122480OB-584–100 to SFV, and PID2020-119030RJ-I00 to LC), all co-financed by the European Regional Development Fund (ERDF). Additional financial support was provided by the CIBER -Consorcio Centro de Investigación Biomédica en Red (CB07708/0012), ISCIII, Ministerio de Ciencia e Innovación, and by the "La Caixa" Foundation (ID 100010434) under grant agreement LCF/PR/HR20/52400013 (to SFV). SFV and JV also acknowledge support from the Agency for Management of University Research Grants of the Generalitat de Catalunya (2021 SGR 01409, 2021 SGR 0089).

Data availability
The datasets supporting the conclusions of this article are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
The Ethics and Clinical Research Committee of the Hospital Regional Universitario de Málaga (Málaga, Spain) approved the cross-sectional and follow-up studies. This study was also approved by the Clinical Research Ethics Committee (CEIM) of the IMIM (CEIm 2019/8422/I).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Endocrinology and Nutrition, Hospital del Mar, Hospital del Mar Medical Research Institute (IMIM), C/ Pg. Marítim 23-25, 08003 Barcelona, Spain. [2]Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas, Instituto de Salud Carlos III (ISCIII), Madrid, Spain. [3]Department of Endocrinology and Nutrition, Research Unit, Institut d'Investigació Sanitària Pere Virgili (IISPV) – Hospital Universitari de Tarragona Joan XXIII, Tarragona, Spain. [4]Statistics and Bioinformatics Platform, Institute of Health Reserach Pere Virgili (IISPV), Reus, Spain. [5]Regional University Hospital of Malaga, Instituto de Investigación Biomédica de Málaga y Plataforma en Nanomedicina (IBIMA-Plataforma Bionand), Málaga, Spain. [6]Department of Electronic Engineering, Metabolomics Interdisciplinary Laboratory, Universitat Rovira I Virgili (URV), Institut d'Investigació Sanitària Pere Virgili (IISPV), Tarragona, Spain. [7]Metabolomics Platform, Institut d'Investigació Sanitària Pere Virgili (IISPV), Universitat Rovira I Virgili (URV), Reus, Spain. [8]GESEC Group, Department of Nursing and Physiotherapy, Faculty of Nursing and Physiotherapy, University of Lleida, Lleida, Spain. [9]Healthcare Research Group (GRECS), Institute of Biomedical Research in Lleida (IRBLleida), Lleida, Spain. [10]Department of Endocrinology and Nutrition, IR Sant Pau, Hospital de La Santa Creu I Sant Pau, Barcelona, Spain. [11]Universitat Rovira I Virgili (URV), Reus, Spain. [12]Research Unit, Institut d'Investigació Sanitària Pere Virgili (IISPV) – Hospital Universitari de Tarragona Joan XXIII, C/ Dr. Mallafrè Guasch, 4, 43005 Tarragona, Spain.

Received: 24 December 2024 Accepted: 2 June 2025
Published online: 06 June 2025

References

1. Apovian CM. The clinical and economic consequences of obesity. Am J Manag Care. 2013;19(10 Suppl):s219-228.
2. Husted AS, Trauelsen M, Rudenko O, Hjorth SA, Schwartz TW. GPCR-Mediated Signaling of Metabolites. Cell Metab. 2017;25:777–96.
3. Blad CC, Tang C, Offermanns S. G protein-coupled receptors for energy metabolites as new therapeutic targets. Nat Rev Drug Discov. 2012;11:603–19.
4. Topping DL, Clifton PM. Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. Physiol Rev. 2001;81:1031–64.
5. Ulven T. Short-chain free fatty acid receptors FFA2/GPR43 and FFA3/GPR41 as new potential therapeutic targets. Front Endocrinol (Lausanne). 2012;3:111.
6. Perry RJ, Peng L, Barry NA, Cline GW, Zhang D, Cardone RL, et al. Acetate mediates a microbiome-brain-β-cell axis to promote metabolic syndrome. Nature. 2016;534:213–7.
7. Hong Y-H, Nishimura Y, Hishikawa D, Tsuzuki H, Miyahara H, Gotoh C, et al. Acetate and propionate short chain fatty acids stimulate adipogenesis via GPCR43. Endocrinology. 2005;146:5092–9.
8. Kimura I, Inoue D, Maeda T, Hara T, Ichimura A, Miyauchi S, et al. Short-chain fatty acids and ketones directly regulate sympathetic nervous system via G protein-coupled receptor 41 (GPR41). Proc Natl Acad Sci U S A. 2011;108:8030–5.
9. Schwiertz A, Taras D, Schäfer K, Beijer S, Bos NA, Donus C, et al. Microbiota and SCFA in lean and overweight healthy subjects. Obesity (Silver Spring). 2010;18:190–5.
10. Rojo-Martínez G, Valdés S, Soriguer F, Vendrell J, Urrutia I, Pérez V, et al. Incidence of diabetes mellitus in Spain as results of the nation-wide cohort di@bet.es study. Sci Rep. 2020;10:2765.
11. Soriguer F, Goday A, Bosch-Comas A, Bordiú E, Calle-Pascual A, Carmena R, et al. Prevalence of diabetes mellitus and impaired glucose regulation in Spain: the Di@bet.es Study. Diabetologia. 2012;55:88–93.
12. Soriguer F, Rojo-Martínez G, Almaraz MC, Esteva I, Ruiz de Adana MS, Morcillo S, et al. Incidence of type 2 diabetes in southern Spain (Pizarra Study). Eur J Clin Invest. 2008;38:126–33.
13. Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser. 1995;854:1–452.

14. Panagiotakos DB, Pitsavos C, Arvaniti F, Stefanadis C. Adherence to the Mediterranean food pattern predicts the prevalence of hypertension, hypercholesterolemia, diabetes and obesity, among healthy adults; the accuracy of the MedDietScore. Prev Med. 2007;44:335–40.
15. Ortega E, Franch J, Castell C, Goday A, Ribas-Barba L, Soriguer F, et al. Mediterranean diet adherence in individuals with prediabetes and unknown diabetes: the Di@bet.es Study. Ann Nutr Metab. 2013;62:339–46.
16. Panagiotakos DB, Pitsavos C, Stefanadis C. Dietary patterns: a Mediterranean diet score and its relation to clinical and biological markers of cardiovascular disease risk. Nutr Metab Cardiovasc Dis. 2006;16:559–68.
17. Tangney CC, Kwasny MJ, Li H, Wilson RS, Evans DA, Morris MC. Adherence to a Mediterranean-type dietary pattern and cognitive decline in a community population. Am J Clin Nutr. 2011;93:601–7.
18. Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care. 2003;26:725–31.
19. Soriguer F, Valdés S, Tapia MJ, Esteva I, Ruiz de Adana MS, Almaraz MC, et al. [Validation of the FINDRISC (FINnish Diabetes RIsk SCore) for prediction of the risk of type 2 diabetes in a population of southern Spain. Pizarra Study]. Med Clin (Barc). 2012;138:371–6.
20. Lotti C, Rubert J, Fava F, Tuohy K, Mattivi F, Vrhovsek U. Development of a fast and cost-effective gas chromatography-mass spectrometry method for the quantification of short-chain and medium-chain fatty acids in human biofluids. Anal Bioanal Chem. 2017;409:5555–67.
21. American Diabetes Association. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2019. Diabetes Care. 2019;42 Suppl 1:S13–28.
22. Seethaler B, Nguyen NK, Basrai M, Kiechle M, Walter J, Delzenne NM, et al. Short-chain fatty acids are key mediators of the favorable effects of the Mediterranean diet on intestinal barrier integrity: data from the randomized controlled LIBRE trial. Am J Clin Nutr. 2022;116:928–42.
23. de Munter JSL, Hu FB, Spiegelman D, Franz M, van Dam RM. Whole grain, bran, and germ intake and risk of type 2 diabetes: a prospective cohort study and systematic review. PLoS Med. 2007;4:e261.
24. Della Pepa G, Vetrani C, Vitale M, Riccardi G. Wholegrain Intake and Risk of Type 2 Diabetes: Evidence from Epidemiological and Intervention Studies. Nutrients. 2018;10:1288.
25. Partula V, Deschasaux M, Druesne-Pecollo N, Latino-Martel P, Desmetz E, Chazelas E, et al. Associations between consumption of dietary fibers and the risk of cardiovascular diseases, cancers, type 2 diabetes, and mortality in the prospective NutriNet-Santé cohort. Am J Clin Nutr. 2020;112:195–207.
26. Cho SS, Qi L, Fahey GC, Klurfeld DM. Consumption of cereal fiber, mixtures of whole grains and bran, and whole grains and risk reduction in type 2 diabetes, obesity, and cardiovascular disease. Am J Clin Nutr. 2013;98:594–619.
27. Martínez-González MA, de la Fuente-Arrillaga C, Nunez-Cordoba JM, Basterra-Gortari FJ, Beunza JJ, Vazquez Z, et al. Adherence to Mediterranean diet and risk of developing diabetes: prospective cohort study. BMJ. 2008;336:1348–51.
28. InterAct Consortium, Romaguera D, Guevara M, Norat T, Langenberg C, Forouhi NG, et al. Mediterranean diet and type 2 diabetes risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) study: the InterAct project. Diabetes Care. 2011;34:1913–8.
29. Salas-Salvadó J, Bulló M, Babio N, Martínez-González MÁ, Ibarrola-Jurado N, Basora J, et al. Reduction in the incidence of type 2 diabetes with the Mediterranean diet: results of the PREDIMED-Reus nutrition intervention randomized trial. Diabetes Care. 2011;34:14–9.
30. Lim J, Henry CJ, Haldar S. Vinegar as a functional ingredient to improve postprandial glycemic control-human intervention findings and molecular mechanisms. Mol Nutr Food Res. 2016;60:1837–49.
31. Canfora EE, van der Beek CM, Hermes GDA, Goossens GH, Jocken JWE, Holst JJ, et al. Supplementation of Diet With Galacto-oligosaccharides Increases Bifidobacteria, but Not Insulin Sensitivity, in Obese Prediabetic Individuals. Gastroenterology. 2017;153:87-97.e3.
32. Fernandes J, Vogt J, Wolever TMS. Intravenous acetate elicits a greater free fatty acid rebound in normal than hyperinsulinaemic humans. Eur J Clin Nutr. 2012;66:1029–34.
33. Canfora EE, Hermes GDA, Müller M, Bastings J, Vaughan EE, van Den Berg MA, et al. Fiber mixture-specific effect on distal colonic fermentation and metabolic health in lean but not in prediabetic men. Gut Microbes. 2022;14:2009297.
34. Canfora EE, Jocken JW, Blaak EE. Short-chain fatty acids in control of body weight and insulin sensitivity. Nat Rev Endocrinol. 2015;11:577–91.
35. Canfora EE, Meex RCR, Venema K, Blaak EE. Gut microbial metabolites in obesity, NAFLD and T2DM. Nat Rev Endocrinol. 2019;15:261–73.
36. Sanna S, van Zuydam NR, Mahajan A, Kurilshikov A, Vich Vila A, Võsa U, et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. Nat Genet. 2019;51:600–5.
37. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444:1027–31.
38. Goffredo M, Mass K, Parks EJ, Wagner DA, McClure EA, Graf J, et al. Role of Gut Microbiota and Short Chain Fatty Acids in Modulating Energy Harvest and Fat Partitioning in Youth. J Clin Endocrinol Metab. 2016;101:4367–76.
39. Zhao L, Lou H, Peng Y, Chen S, Fan L, Li X. Elevated levels of circulating short-chain fatty acids and bile acids in type 2 diabetes are linked to gut barrier disruption and disordered gut microbiota. Diabetes Res Clin Pract. 2020;169:108418.
40. Salamone D, Rivellese AA, Vetrani C. The relationship between gut microbiota, short-chain fatty acids and type 2 diabetes mellitus: the possible role of dietary fibre. Acta Diabetol. 2021;58:1131–8.
41. Xie C, Qi C, Zhang J, Wang W, Meng X, Aikepaer A, et al. When short-chain fatty acids meet type 2 diabetes mellitus: Revealing mechanisms, envisioning therapies. Biochem Pharmacol. 2025;233:116791.
42. Müller M, Hernández MAG, Goossens GH, Reijnders D, Holst JJ, Jocken JWE, et al. Circulating but not faecal short-chain fatty acids are related to insulin sensitivity, lipolysis and GLP-1 concentrations in humans. Sci Rep. 2019;9:12515.
43. Bouter K, Bakker GJ, Levin E, Hartstra AV, Kootte RS, Udayappan SD, et al. Differential metabolic effects of oral butyrate treatment in lean versus metabolic syndrome subjects. Clin Transl Gastroenterol. 2018;9:155.
44. Tang C, Ahmed K, Gille A, Lu S, Gröne H-J, Tunaru S, et al. Loss of FFA2 and FFA3 increases insulin secretion and improves glucose tolerance in type 2 diabetes. Nat Med. 2015;21:173–7.
45. Soriguer F, Almaraz MC, García-Almeida JM, Cardona I, Linares F, Morcillo S, et al. Intake and home use of olive oil or mixed oils in relation to healthy lifestyles in a Mediterranean population. Findings from the prospective Pizarra study. Br J Nutr. 2010;103:114–22.
46. Zazpe I, Santiago S, de la O V, Romanos-Nanclares A, Rico-Campà A, Álvarez-Zallo N, et al. Validity and reproducibility of a semi-quantitative food frequency questionnaire in Spanish preschoolers - The SENDO project. Nutr Hosp. 2020;37:672–84.
47. Larroya A, Tamayo M, Cenit MC, Sanz Y. Validity and Reproducibility of a Spanish EPIC Food Frequency Questionnaire in Children and Adolescents. Nutrients. 2024;16:3809.
48. Schröder H, Covas MI, Marrugat J, Vila J, Pena A, Alcántara M, et al. Use of a three-day estimated food record, a 72-hour recall and a food-frequency questionnaire for dietary assessment in a Mediterranean Spanish population. Clin Nutr. 2001;20:429–37.

Publisher's Note