Serum metabolites reflecting gut microbiome alpha diversity predict type 2 diabetes

Cristina Menni , Jialing Zhu , Caroline I Le Roy , Olatz Mompeo , Kristin Young , Casey M. Rebholz , Elizabeth Selvin , Kari E. North , Robert P Mohney , Jordana T Bell , Eric Boerwinkle , Tim D Spector , Massimo Mangino , Bing Yu & Ana M Valdes

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER/REPORT

OPEN ACCESS | Check for updates

Serum metabolites reflecting gut microbiome alpha diversity predict type 2 diabetes

Cristina Menni [ID][a], Jialing Zhu[b], Caroline I Le Roy [ID][a], Olatz Mompeo[a], Kristin Young [ID][c], Casey M. Rebholz [ID][d], Elizabeth Selvin [ID][d], Kari E. North[c], Robert P Mohney [ID][e], Jordana T Bell [ID][a], Eric Boerwinkle[b,f], Tim D Spector [ID][a], Massimo Mangino [ID][a], Bing Yu [ID][b], and Ana M Valdes [ID][a,g,h]

[a]Department of Twin Research and Genetic Epidemiology, Kings College London, London, UK; [b]School of Public Health, University of Texas Health Science Center, Houston, TX, USA; [c]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA; [d]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; [e]Metabolon Inc., Morrisville, NC, USA; [f]Baylor College of Medicine, Houston, TX, USA; [g]School of Medicine, Nottingham, UK; [h]NIHR Nottingham Biomedical Research Centre, Nottingham, UK

ABSTRACT

Type 2 diabetes (T2D) is associated with reduced gut microbiome diversity, although the cause is unclear. Metabolites generated by gut microbes also appear to be causative factors in T2D. We therefore searched for serum metabolites predictive of gut microbiome diversity in 1018 females from TwinsUK with concurrent metabolomic profiling and microbiome composition. We generated a Microbial Metabolites Diversity (MMD) score of six circulating metabolites that explained over 18% of the variance in microbiome alpha diversity. Moreover, the MMD score was associated with a significantly lower odds of prevalent (OR[95%CI] = 0.22[0.07;0.70], P = .01) and incident T2D (HR [95%CI] = 0.31[0.11,0.90], P = .03). We replicated our results in 1522 individuals from the ARIC study (prevalent T2D: OR[95%CI] = 0.79[0.64,0.96], P = .02, incident T2D: HR[95%CI] = 0.87[0.79,0.95], P = .003). The MMD score mediated 28%[15%,94%] of the total effect of gut microbiome on T2D after adjusting for confounders. Metabolites predicting higher microbiome diversity included 3-phenylpropionate(hydrocinnamate), indolepropionate, cinnamoylglycine and 5-alpha-pregnan -3beta,20 alpha-diol monosulfate(2) of which indolepropionate and phenylpropionate have already been linked to lower incidence of T2D. Metabolites correlating with lower microbial diversity included glutarate and imidazole propionate, of which the latter has been implicated in insulin resistance. Our results suggest that the effect of gut microbiome diversity on T2D is largely mediated by microbial metabolites, which might be modifiable by diet.

Introduction

Type 2 diabetes (T2D) is a major public health burden with a 10-fold increase in T2D rates globally between 1980 (30 million) and 2010 (300 million).[1] Accumulating evidence indicates that bacteria in the human gut play a vital role in the pathophysiology of T2D and related traits,[2–7] with higher gut microbiome alpha diversity having a beneficial effect.

Gut microbiome composition is regulated jointly by the host genome, the colonic milieu and diet.[6] During digestion, gut microbes coproduce a wide range of metabolites such as short chain fatty acids (SCFAs),[8] lipopolysaccharide (LPS),[9] imidazole propionate,[10] among others, that are essential for regulating multiple host-microbiome pathways.[11] Previous studies in animals and humans have identified a gut microbiome signature of T2D.[6,7] Individuals with T2D have lower gut microbiome alpha diversity compared with healthy controls and they also show impaired SCFA butyrate production,[12] modifications in incretin secretion, reducing intestinal permeability leading to metabolic endotoxemia (increased LPS in blood)[13,14] that is linked to increasing metabolic inflammation and insulin resistance induced by high-fat diet.[15]

In addition, a number of studies have shown that microbial metabolites, other than SCFAs, are linked to the development of T2D,[16] suggesting that part of the benefit of having a diverse gut microbiome is related to an improvement of the metabolic cross-

Supplemental data for this article can be accessed on the publisher's website.

talk occurring between the gut microbes and their host.[17]

A recent study has found that 40 metabolites, including 13 of them of microbial origin[18] can predict 45% of the inter-individual variation in alpha diversity. However, the study did not focus on the specific relevance of bacterial metabolites to T2D, but primarily on the usefulness of human and bacterial metabolites as biomarkers predictive of diversity.

Here we identify serum metabolites reflective of gut microbiome function that can mediate the effect of gut microbiome composition on T2D and related metabolic traits in a large UK cohort. We replicate our results in an independent US sample.

Results

The descriptive characteristics of the study populations are presented in Table 1. We included 1018 middle-age females from TwinsUK with 16 s gut microbiome data and concurrent fasting serum metabolomics. The replication cohort consisted of 1522 male and female participants from the ARIC study of European ancestry.

We randomly divided the TwinsUK study sample into two independent sets: a discovery set of 409 individuals and a test set of 609 individuals. Six metabolites were associated with Shannon diversity after adjusting for age, BMI, family relatedness (as random effect) and multiple testing (using Bonferroni correction $P < 1.68x10^{-4} = 0.1/596$ metabolites) in the discovery set (Table S1). Of these, there were four amino acid derivatives, one steroid and one xenobiotic. We then created a microbial metabolite diversity (MMD) score in

females from TwinsUK to assess the combined effects of all metabolic traits identified, by including the six metabolites into a linear regression. We found that the MMD score explains 27% of the variance in Shannon Diversity (Figure 1) and the variance inflation factor (VIF) between the selected feature was low (VIF = 1.38) suggesting low collinearity.[19]

$$MMD\ score = -0.0476306 + 0.1746576$$
$$\times\ (3 - phenylpropionate(hydrocinnamate)))$$
$$-\ 0.0867652 \times (imidazolepropionate)$$
$$+\ 0.1461903 \times (cinnamoylglycine))$$
$$+\ 0.1674953 \times \begin{pmatrix} 5alphapregnan\_3beta, \\ 20alpha\_diolmonosulfate(2) \end{pmatrix}$$
$$-\ 0.145269 \times (glutarate(pentanedioate))$$
$$+\ 0.0683178 \times (indolepropionate)$$

We then replicated the Shannon-metabolite association in 609 independent individuals from the test set. All metabolites were significantly associated with Shannon diversity after adjusting for covariates and multiple testing in the test set (Figure 1). We also tested the predictive value of the MMD score in the test set and we found that it explained 18.68% of the variance in Shannon diversity (Beta (SE) = 1(0.09), $P = 3.32x10^{-26}$). We combined the results of the discovery and test set using Fixed Effect Meta-Analysis (Figure 1).

To further investigate the predictive value of the MMD score, we tested it for association with the other alpha diversity indexes in the overall cohort. As shown in Figure 2 and Table S1, the MMD score was significantly associated with the observed number of OTUs (Beta (SE) = 0.99(0.06), $P = 2.84x10^{-55}$; r = 0.42, R2 = 0.19), Simpson Index (0.65 (0.08), $P = 5.97x10^{-16}$; r = 0.3, R2 = 0.08) and Chao1 Index (0.46 (0.07), $P = 4.21x10^{-12}$; r = 0.2, R2 = 0.04)

We then tested the association between the MMD score and T2D and related phenotypes adjusting for age and BMI. We find that the MMD score was associated with a lower odds of both prevalentT2D (OR[95%CI] = 0.22[0.07;0.70],P = .01) and with lower risk of incident T2D (HR[95%CI] = 0.31 [0.11,0.90], P = .03). We repeated the adjusting also for baseline fasting glucose and results were consistent. The MMD score was also negatively correlated with BMI, visceral fat mass, fasting glucose and

Table 1. Descriptive characteristics of the study population.

|	TwinsUK N (%)	ARIC N (%)
N	1018	1522
Females, %	1018 (100%)	827 (54.3%)
T2D prevalent cases,	47 (4.72%)	119 (7.8%)
T2D incident cases	23	467 (34.7%)
	Mean (SD)	Mean (SD)
Age, yrs	65.19 (7.74)	54.56 (5.74)
BMI, kg/m²	26.19 (4.83)	27.25 (5.01)
Homa2-IR	0.93 (0.67)	
Glucose, mmlol/l	4.75 (0.63)	5.84 (1.54)
VFM, (g)	606.77 (290.07)	
Indices of microbiome diversity		
Shannon Diversity	5.14 (0.7)	NA
Observed number of OTUs	89.38 (113.52)	NA
Simpson diversity	0.92 (0.06)	NA
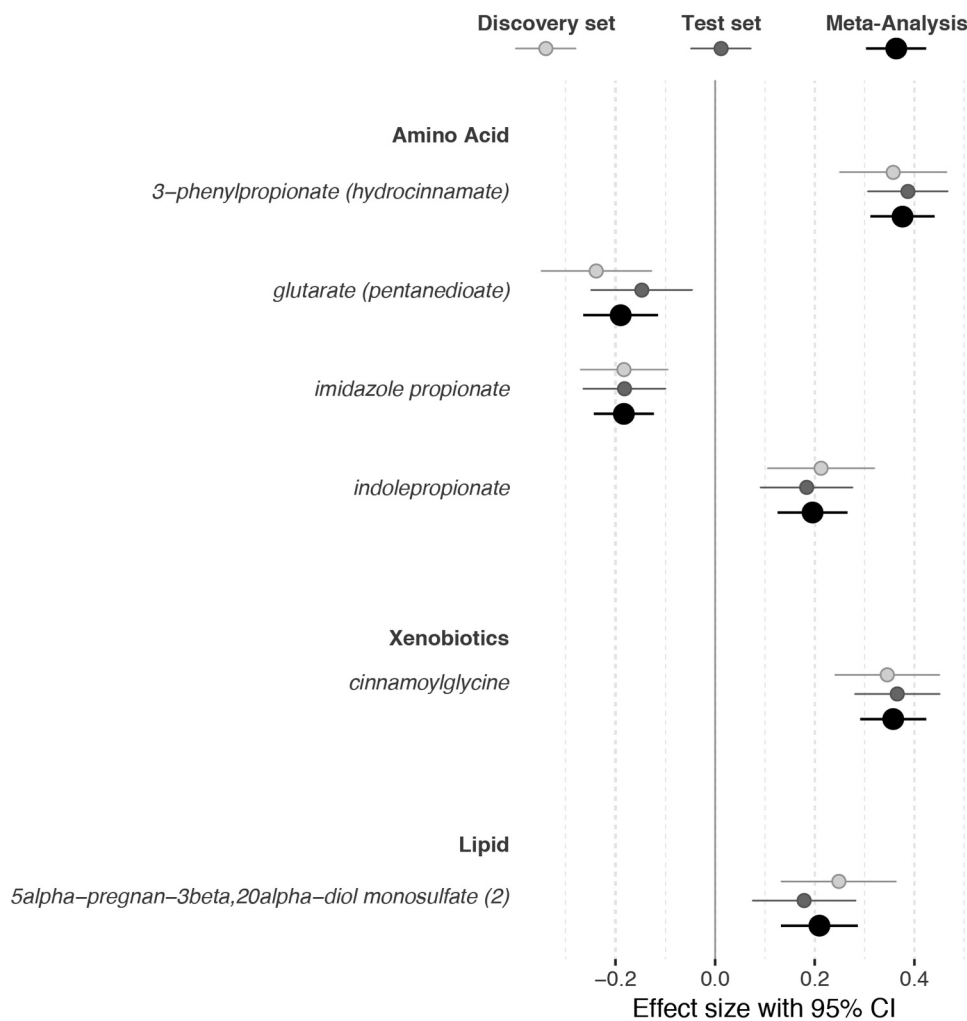
Figure 1. Metabolites trait significantly associated with gut microbiome alpha diversity (Shannon Index) in the discovery and test set of the TwinsUK cohort and in the overall cohort (using inverse variance fixed effect meta-analysis). Analyses adjusted by age, sex, body mass index, family relatedness, and multiple testing. CI indicates confidence interval.

HOMA2-IR, as shown in Figure 2 and in Table S1. We find that the associations between the MMD score and clinical variables were actually stronger than those with actual microbiome diversity (Table S1). To confirm the relevance of the metabolites identified we replicated the metabolite score association with both prevalent and incident T2D in 1522 individuals of European Ancestry from the ARIC cohort. As shown in Figure 3, we find that the MMD score was significantly associated with a lower risk of prevalent and incident T2D both in ARIC (*Prevalent T2D*: OR[95%CI] = 0.79[0.64,0.96], P = .02; *Incident T2D*: HR[95%CI] = 0.87[0.79,0.95], P = .003) and in the meta-analysis which pooled estimates from TwinsUK and ARIC (*Prevalent T2D*: OR[95%CI] = 0.75[0.62,0.92], Fixed Effect P = .006, Han & Eskin[20] Random-Effect P = .006;

Incident T2D: HR[95%CI] = 0.86[0.78,0.95], P = .002, Han & Eskin Random-Effect P = .004).

We further replicated the cross-sectional association of the MMD score with BMI (Beta (SE) = −0.59 (0.13), P = 5.03×10^{-6})). Considering the gender difference in microbiota and metabolism, we rerun the analysis in ARIC stratifying by gender and we find that the MMD score performance is not always better in females than males though the directionality is always consistent (Table S2).

To test our hypothesis that the association between diversity and T2D is due to the metabolites produced by the microbes we then retested for an association between T2D with the Shannon diversity index after adjusting for metabolites. We found there was no significant association between T2D and alpha diversity when adjusting for the MMD
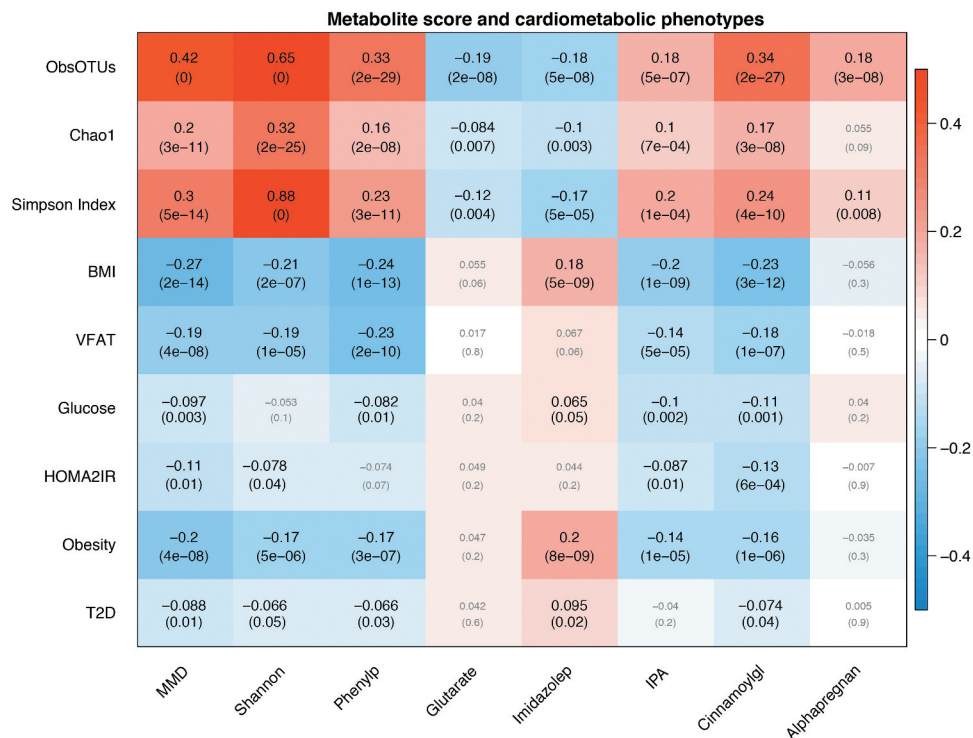
Figure 2. Microbial traits (MMD score, alpha diversity, microbial metabolites), T2D and related traits. Each cell of the matrix contains the correlation coefficient between one microbial trait and a metabolic phenotype score and the corresponding *P* value. The table is color coded by correlation according to the table legend (red for positive and blue for negative correlations). Analyses are adjusted for age and BMI. ObsOTUs = number of observed OTUs, BMI = body mass index, VFAT = visceral fat mass, HOMA2IR = insulin resistance, T2D = type 2 diabetes, MMD = Microbial Metabolites Diversity score, Shannon = Shannon Alpha Diversity Index, Phenylp = 3-phenylpropionate (hydrocinnamate), Imidazolep = imidazole propionate, IPA = indolepropionate, Cinnamoylgl = cinnamoylglycine, Alphapregan = 5alpha-pregnan-3beta,20alpha-diol monosulfate (2).
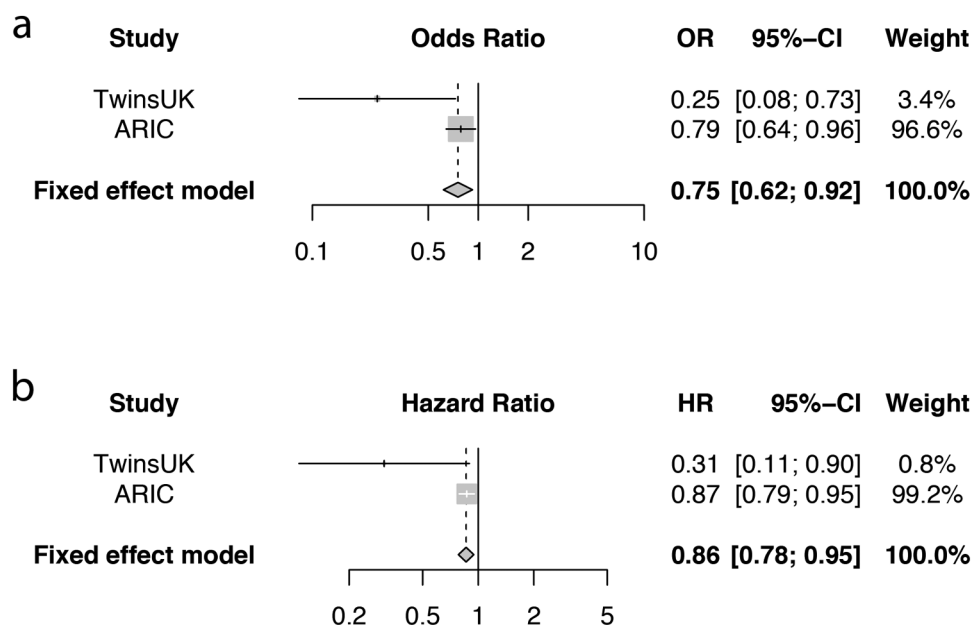

Figure 3. MMD score and risk of (a) prevalent and (b) incident T2D in the TwinsUK (TUK) and ARIC cohorts and results from fixed effect meta-analyses.

score (OR[95%CI] = 0.81[0.47, 1.38], P = .437), while the MMD score is significantly associated to a lower risk of T2D when adjusting for Shannon diversity (OR[95%CI] = 0.26[0.07,0.93], P = .038). We therefore conducted formal mediation analysis to determine the indirect effect of the MMD score on the effect between Shannon diversity (high versus low defined as above or below the median, respectively) and T2D. The mediation model, which intends to evaluate the strength of the indirect effects, found that the direct relationship between microbiome factors and T2D was not statistically significant (path coefficient = −.026[−0.06; 0.002]). For indirect effect, the effect of the MMD was statistically significant (path coefficient = −0.24 [−0.87; −0.39]). The VAF score for the MMD was 0.28[0.15; 0.94].

Discussion

In this study, we identified a panel of circulating metabolites reflective of gut microbiome alpha diversity metrics, which was associated with T2D, obesity, adiposity, and other metabolic traits. We found the associations between microbiome diversity and T2D related metabolic traits were stronger with the metabolite surrogate measures than with the actual diversity measures. Our mediation analysis suggests that 28% of the effect the gut microbiome alpha diversity has on T2D is mediated by gut microbiota-associated metabolites. Some of these metabolites are produced by gut bacteria and absorbed through the gut barrier to then reach the bloodstream (3-phenylpropionate, IPA, cinnamoylglycine and imidazole propionate), suggesting that the associations between microbiome diversity and metabolic outcomes are simply reflecting the effect of some of these metabolites. A more diverse microbiome contains a higher number of metabolic pathways and hence is likely to be producing a relatively greater number of metabolites that contribute to lower inflammation and improved metabolic health, such as IPA, cinnamic acid and 3-phenylpropionate. Four of these compounds are the result of microbial amino acid metabolism pathways, one of them is the result of steroid microbial metabolism and one is a phenolic conjugate of a short chain fatty acid.

Importantly, all of the metabolites that we have identified as being strongly predictive of microbiome diversity can be found in human stool samples (see, e.g.,[21]).

A much larger number of metabolites have been shown to be predictive of alpha diversity in a recent study with a smaller sample size[18] using a LASSO approach. However, less than 25% of those metabolites are of microbial origin, which for our purposes, i.e., to identify metabolites that mediate the effect of the gut microbiome on T2D development, risks confounding causes and effects by including metabolites of purely human origin. In this study, we have focused on identifying metabolites of microbial origin that are individually associated and replicated after adjustment for multiple tests making these much more likely to be implicated in the causal link between gut microbiome diversity and T2D.

Compounds associated with lower microbiome alpha diversity

Glutarate is produced by the human body as part of the metabolism of tryptophan and lysine. It is produced via lysine degradation in microorganisms[22] and the microbiome is now known to be an important source of this compound which has been linked to cancer and organic acidurias.[22] Importantly, glutarate has already been reported to correlate with lower microbiome richness linked to severe obesity.[23]

Imidazole propionate is a microbially produced histidine-derived metabolite.[10] This compound is present at higher concentrations in the portal and peripheral blood of subjects with type 2 diabetes.[10] Imidazole propionate can worsen glucose tolerance when administered to mice and impairs insulin signaling at the level of insulin receptor substrate through the activation of p38γ MAPK, which promotes p62 phosphorylation and, subsequently, activation of mTORC1.[10] The association with obesity, T2D and glucose that we found is therefore not surprising and in fact confirms the results seen in the Dutch cohorts studied for this compound.[10]

Compounds associated with higher microbiome alpha diversity:

Indolepropionic acid (IPA) is a deamination product of tryptophan that can only be produced by

gut microbes.[24] It regulates gastrointestinal barrier function via its interaction with the pregnane X receptor (PXR), which is predictive of the onset of type 2 diabetes,[16] has been shown to be strongly correlated with higher gut microbiome diversity and its levels in serum are determined primarily by gut microbiome composition.[25]

Among the compounds that intestinal bacteria can metabolize there are sterols, bile acids, and steroid hormones.[26] The metabolites generated by gut bacteria action on host steroids, when reabsorbed in the intestine back into the circulation, produce biological effects.[26] The formation of sulfated steroids with a 3 alpha-hydroxy-5 alpha is normally derived from progesterone and account for at least 50% of the metabolism of progesterone in late pregnancy.[27] Compounds in this class have been implicated in the severity of premenstrual symptoms.[28] This compound however is also present in human feces[21] and importantly it has been, along with IPA, reported to be decreased in the plasma of patients with primary dilated cardiomyopathy.[29] We found that higher serum levels of 5-alpha-pregnan-3beta,20 alpha-diol monosulfate (2) were positively correlated with microbiome diversity and in the context of the metabolite score, with lower BMI, VFAT and risk of T2D. The negative association with primary dilated cardiomyopathy is particularly relevant as that study included both men and women with an average age of 47 y, suggesting that the microbial origin of the metabolite may be the same for both genders. Indeed, Wilmanski and colleagues[18] reported that serum concentrations of 5-alpha-pregnan-3beta,20 alpha-diol monosulfate (2) were slightly higher in men than in women and that this steroid is associated with microbiome diversity in both males and females.

3-phenylpropionic acid is an organic acid that is abundantly found in the products of lactic acid bacterial fermentation[30] and in honey.[31] It is an antimicrobial compound with broad-spectrum activity against a wide range of Gram-positive bacteria, some Gram-negative bacteria and fungi.[32] 3-phenylpropionic acid along with phenylacetic acid is one of the main phenolic metabolites present in human feces. Its levels in feces have been shown to be inversely correlated with fecal levels of propionate and positively correlated with fecal levels of acetate. It appears therefore to be directly related to SCFA metabolism.[33] Given its correlation with dietary fiber, it is not surprising that it is associated positively with microbiome diversity. The mechanism by which it influences cardiometabolic traits is unknown but we hypothesize that its antimicrobial properties may result in lower production of lipopolysaccharide and that its antioxidant activities may contribute to decreased risk of obesity and insulin resistance.[34]

Cinnamic acid is a naturally occurring bioactive compound synthesized in plants by the shikimate pathway, where phenylalanine and tyrosine are two precursor molecules. It is found in plant-based foods such as fruits, vegetables, and whole grains and several potential benefits of cinnamic acid and its derivatives have been demonstrated in *in vitro* and preclinical studies with regards to T2D, although there is no clinical evidence.[35] Cinnamoylglycine is a glycine conjugate of cinnamic acid and it is known to be produced by gut microbes because it is abundant in the serum of conventional mice but present in minimal concentrations in the serum of germfree mice.[24] Cinnomoylglycine has substantially greater urinary clearance than creatinine[36] and hence accumulates in plasma under conditions of diminished renal function. In our data, we found that the association between cinnomoylglycine, diversity and clinical traits was not affected when adjusting for creatinine. Although its functional effects in humans are unknown, urinary excretion levels of cinnamoylglycine have been proposed as markers of colonization resistance against *Clostridium difficile*, i.e., as a marker of a healthy gut microbiome that can inhibit the growth of pathogenic microorganisms.[37]

The current study has several strengths. The study benefits from a big sample size, the use of internal subsets for discovery and retesting, a large independent replication cohort and a large number of metabolites. Both TwinsUK and ARIC have extensive phenotypic data available that allowed us to accurately classify subjects according to T2D status. Finally, the robustness of our result is highlighted by the fact that we confirm many previous findings, and the predictive value of our metabolic score is successfully validated in an independent sample.

Our study has some limitations. Because TwinsUK is a sample representative of the general UK population the number of prevalent and incident cases of T2D in a sample of ~1000 individuals is relatively modest (n = 47 and 23, respectively) hence the large confidence intervals. Moreover, the definitions for T2D are different in the two cohorts. In the TwinsUK subjects, glucose concentration was only measured in the fasting state, potentially underestimating T2D incidence. This and the lack of men in the analysis are study limitations that are at least in part overcome by the inclusion of the ARIC cohort, which includes men and a much larger number of individuals with incident and prevalent T2D. Also, we were not able to test for the association between the MMD score and Shannon diversity and the mediation analysis results in ARIC as data on 16 s gut microbiome data was not available.

In conclusion, in this study, we find that a set of six microbially derived or microbially modified metabolites in serum was predictive of gut microbiome diversity and of T2D risk in two independent populations, and we report that the effect of gut microbiome diversity on risk of T2D is in part mediated by these metabolites. Clinical studies should investigate the pathways identified and determine whether modifying these compounds through diet could help reduce or monitor T2D disease.

Subject and methods

Study subjects

We included a random subset of 1018 females from TwinsUK representative of the general middle age female UK population[38] with concurrent fasting serum metabolomic profiling (592 metabolites) and fecal 16 S sequencing (measuring microbiome composition) along with diabetic and metabolic information. T2D prevalent cases were defined as individuals with fasting glucose≥7 mmol/l or with a physician's letter confirming diagnosis or use of medication for diabetes. Visceral fat (VF), BMI, obesity and Homa2-IR were measured as previously described.[25] Incident T2D cases were defined with the same criteria but having developed T2D at follow-up (mean (SD) = 8.16 (1.3) y).

The study was approved by NRES Committee London–Westminster, and all twins provided informed written consent.

The replication cohort consisted of 1522 European American (T2D incident cases 467, T2D prevalent cases 119) from the Atherosclerosis Risk in Communities (ARIC) study with fasting metabolomic profiles.[39] In the ARIC study, prevalent diabetes at baseline was defined as fasting glucose ≥7.0 mmol/l, non-fasting glucose ≥11.1 mmol/l, self-reported diagnosis of diabetes by a physician or use of medication for diabetes within the previous 2 weeks. The incidence of diabetes was ascertained from visit 1 (1987–1989) until December 31, 2015. Incident diabetes was defined as elevated glucose (fasting glucose ≥7.0 mmol/l or non-fasting glucose ≥11.1 mmol/l) at any of three triennial follow-up visits and visit 5 (2010–2013), self-report of a diabetes diagnosis by a physician at a study visit or annual follow-up telephone interview, or self-report of diabetes medication use during a study visit or annual follow-up telephone interview.[40] The incidence of diabetes was ascertained from visit 1 until December 31, 2015.

Note, the accumulative T2D incidence per year is 0.4%[0.16%; 1.02%] in the subset of TwinsUK individuals included in this study, 1.1%[0.62%; 1.96%] in the participants from ARIC included here. This compares to 0.67% [0.53%;085%] in US individuals according to the National Diabetes Statistics Report.[41]

Metabolomics profiling

Metabolomics profiling was conducted using ultra-high-performance liquid chromatography-tandem mass-spectrometry by the metabolomics provider Metabolon Inc. (Morrisville, USA) on fasting serum samples from participants in the TwinsUK study and ARIC, as described previously,[39,42] for details see Supplementary Text. The metabolomic dataset measured by Metabolon includes 592 known metabolites containing the following broad categories – amino-acids, peptides, carbohydrates, energy intermediates, lipids, nucleotides, cofactors and vitamins, and xenobiotics. These include metabolites of established microbial origin.[43] We inverse normalized circulating metabolite levels as the metabolite concentrations were not normally distributed. We imputed missing values using the minimum run-day measures.

Microbiome analysis

Gut microbiome composition was determined by 16 S rRNA gene sequencing carried out as previously described.[44] Briefly, the V4 region of the 16 S rRNA gene was amplified and sequenced on Illumina MiSeq. Reads were then summarized to operational taxonomic units (OTUs) Quality control was carried out on a per sample basis, discarding paired-ends with an overlap of less than 200nt and removing chimeric sequences using de novo chimera detection in USEARCH.[45] De novo OTU clustering was then carried across all reads using Sumaclust within QIIME 1.9.0, grouping reads with a 97% identity threshold.[46,47] OTU counts were converted to log transformed relative abundances, with zero counts handled by the addition of an arbitrary value (10–6). In order to calculate alpha diversity, the complete OTU count table was rarefied to 10,000 sequences per sample 50 times. Alpha diversity metrics were calculated for each sample in each of the rarefied tables and final diversity measures taken as the mean score across all 50. Alpha diversities were quantified as observed OTU counts and Shannon and Simpson diversity indices, commonly used to characterize species diversity in a community. They both account for abundance and evenness of the species present. Alpha diversity indexes were standardized to have mean 0 and SD 1.

Statistical analysis

Statistical analysis was carried out using Stata version 12 and R version 3.6.0.

In the TwinsUK cohort, we randomly divided study subjects into two independent subsets: a discovery set including 409 individuals (1/3 of the sample) and a test set of 609 individuals (2/3) of the sample. In the discovery set, association analyses between gut microbiome composition and metabolites were performed using random intercept linear regression adjusting for age, body mass index, family relatedness and multiple testing using Bonferroni correction ($P < 1.68 \times 10^{-4} = 0.1/596$ metabolites). We linearly combined the metabolites significantly associated with Shannon Diversity to create a Microbial Metabolites Diversity (MMD) score. We validated the score in 609 females from TwinsUK. Results from the Discovery and Test set were combined using inverse variance fixed effect meta-analysis as implemented in the R package meta. We further tested the predictive value of the MMD score by testing its association with T2D and related traits (obesity, BMI, visceral fat, fasting glucose, insulin resistance) and with other measures of diversity in the overall cohort. We replicated the results in 1522 male and female participants of European ancestry from ARIC. In the ARIC study, we applied linear regression to test the association between the MMD score and BMI adjusting for age, sex and center. We next applied logistic and Cox regressions to test the association between the MMD score and prevalent and incident T2D adjusting age, sex, center and BMI. Prevalent T2D cases were excluded from the incident T2D analysis. We did not observe violation on the proportional hazards assumption for MMD score.

Results were combined using inverse variance fixed effect meta-analysis as implemented in the R package meta and, since the sex distributions of the two cohorts are different, also using Han & Eskin[20] random-effect meta-analysis as implemented in the software metasoft.

Finally, we employed mediation analysis as implemented in the STATA package MedEff to test the mediation effects of microbial metabolites (indirect effect) on the total effect of microbiome factors (Shannon Index) on T2D adjusting for age and BMI. We constructed a mediation model to quantify both the direct effect of microbiome factors on T2D and the indirect (mediated) effects mentioned above. The model goodness of fit was assessed by standardized path coefficient and effect size (f^2)[48–50] yield using the lower and upper bound of the 95% confidence interval. The variance accounted for (VAF) score, which represents the ratio of indirect-to-total effect and determines the proportion of the variance explained by the mediation process, was further used to determine the significance of mediation effect.[51]

Disclosure statement

ORCID

Cristina Menni http://orcid.org/0000-0001-9790-0571
Caroline I Le Roy http://orcid.org/0000-0002-0341-751X
Kristin Young http://orcid.org/0000-0003-0070-6145
Casey M. Rebholz http://orcid.org/0000-0002-5442-8745
Elizabeth Selvin http://orcid.org/0000-0001-6923-7151
Robert P Mohney http://orcid.org/0000-0002-9315-9541
Jordana T Bell http://orcid.org/0000-0002-3858-5986
Tim D Spector http://orcid.org/0000-0002-9795-0365
Massimo Mangino http://orcid.org/0000-0002-2167-7470
Bing Yu http://orcid.org/0000-0003-4818-1077
Ana M Valdes http://orcid.org/0000-0003-1141-4471

References

1. Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, Singh GM, Gutierrez HR, Lu Y, Bahalim AN, et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. Lancet. 2011;377(9765):557–567. doi:10.1016/S0140-6736(10)62037-5.

2. Allin KH, Tremaroli V, Caesar R, Jensen BAH, Damgaard MTF, Bahl MI, Licht TR, Hansen TH, Nielsen T, Dantoft TM, et al. Aberrant intestinal microbiota in individuals with prediabetes. Diabetologia. 2018;61(4):810–820. doi:10.1007/s00125-018-4550-1.

3. Menni C, Jackson MA, Pallister T, Steves C, Spector T, Valdes A. Gut microbiome diversity and high fibre intake are related to lower long term weight gain. Int J Obes (Lond). 2017. doi:10.1038/ijo.2017.66.

4. Jackson MA, Verdi S, Maxan ME, Shin CM, Zierer J, Bowyer RCE, Martin T, Williams FMK, Menni C, Bell JT, et al. Gut microbiota associations with common diseases and prescription medications in a population-based cohort. Nat Commun. 2018;9(1):2655. doi:10.1038/s41467-018-05184-7.

5. Forslund K, Hildebrand F, Nielsen T, Falony G, Le Chatelier E, Sunagawa S, Prifti E, Vieira-Silva S, Gudmundsdottir V, Pedersen HK, et al. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. Nature. 2015;528(7581):262–266. doi:10.1038/nature15766.

6. Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, Nielsen J, Backhed F. Gut metagenome in european women with normal, impaired and diabetic glucose control. Nature. 2013;498(7452):99–103. doi:10.1038/nature12198.

7. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang 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(7418):55–60. doi:10.1038/nature11450.

8. van Deventer CA, Lindeque JZ, van Rensburg PJ, Malan L, van der Westhuizen FH, Louw R. Use of metabolomics to elucidate the metabolic perturbation associated with hypertension in a black south african male cohort: the sabpa study. J Am Soc Hypertens JASH. 2015;9(2):104–114. doi:10.1016/j.jash.2014.11.007.

9. Hersoug LG, Moller P, Loft S. Gut microbiota-derived lipopolysaccharide uptake and trafficking to adipose tissue: implications for inflammation and obesity. Obesity Rev J Int Study Obesity. 2016;17(4):297–312. doi:10.1111/obr.12370.

10. Koh A, Molinaro A, Stahlman M, Khan MT, Schmidt C, Manneras-Holm L, Wu H, Carreras A, Jeong H, Olofsson LE, et al. Microbially produced imidazole propionate impairs insulin signaling through mtorc1. Cell. 2018;175(4):947–961 e917. doi:10.1016/j.cell.2018.09.055.

11. Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. BMJ. 2018;361:k2179. doi:10.1136/bmj.k2179.

12. Gao Z, Yin J, Zhang J, Ward RE, Martin RJ, Lefevre M, Cefalu WT, Ye J. Butyrate improves insulin sensitivity and increases energy expenditure in mice. Diabetes. 2009;58(7):1509–1517. doi:10.2337/db08-1637.

13. Erridge C, Attina T, Spickett CM, Webb DJ. A high-fat meal induces low-grade endotoxemia: evidence of a novel mechanism of postprandial inflammation. Am J Clin Nutr. 2007;86(5):1286–1292. doi:10.1093/ajcn/86.5.1286.

14. Cani PD, Amar J, Iglesias MA, Poggi M, Knauf C, Bastelica D, Neyrinck AM, Fava F, Tuohy KM, Chabo C, et al. Metabolic endotoxemia initiates obesity and insulin resistance. Diabetes. 2007;56(7):1761–1772. doi:10.2337/db06-1491.

15. Cani PD, Knauf C, Iglesias MA, Drucker DJ, Delzenne NM, Burcelin R. Improvement of glucose tolerance and hepatic insulin sensitivity by oligofructose requires a functional glucagon-like peptide 1 receptor. Diabetes. 2006;55(5):1484–1490. doi:10.2337/db05-1360.

16. de Mello VD, Paananen J, Lindstrom J, Lankinen MA, Shi L, Kuusisto J, Pihlajamaki J, Auriola S, Lehtonen M, Rolandsson O, et al. Indolepropionic acid and novel lipid metabolites are associated with a lower risk of type 2 diabetes in the finnish diabetes prevention study. Scientific Reports. 2017;7(1):46337. doi:10.1038/srep46337.

17. Schwiertz A, Taras D, Schafer K, Beijer S, Bos NA, Donus C, Hardt PD. Microbiota and scfa in lean and overweight healthy subjects. Obesity (Silver Spring). 2010;18(1):190–195. doi:10.1038/oby.2009.167.

18. Wilmanski T, Rappaport N, Earls JC, Magis AT, Manor O, Lovejoy J, Omenn GS, Hood L, Gibbons SM, Price ND. Blood metabolome predicts gut microbiome alpha-diversity in humans. Nat Biotechnol. 2019. doi:10.1038/s41587-019-0233-9.

19. Gareth J, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: with applications in r. New York: Springer Publishing Company; 2014.

20. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. Am J Hum Genet. 2011;88(5):586–598. doi:10.1016/j.ajhg.2011.04.014.

21. Brown DG, Borresen EC, Brown RJ, Ryan EP. Heat-stabilised rice bran consumption by colorectal cancer survivors modulates stool metabolite profiles and metabolic networks: A randomised controlled trial. Br J Nutr. 2017;117(9):1244–1256. doi:10.1017/S0007114517001106.

22. Knorr S, Sinn M, Galetskiy D, Williams RM, Wang C, Muller N, Mayans O, Schleheck D, Hartig JS. Widespread bacterial lysine degradation proceeding via glutarate and l-2-hydroxyglutarate. Nat Commun. 2018;9(1):5071. doi:10.1038/s41467-018-07563-6.

23. Aron-Wisnewsky J, Prifti E, Belda E, Ichou F, Kayser BD, Dao MC, Verger EO, Hedjazi L, Bouillot JL, Chevallier JM, et al. Major microbiota dysbiosis in severe obesity: fate after bariatric surgery. Gut. 2019;68(1):70–82. doi:10.1136/gutjnl-2018-316103.

24. Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC, Siuzdak G. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. Proc Natl Acad Sci U S A. 2009;106(10):3698–3703. doi:10.1073/pnas.0812874106.

25. Menni C, Hernandez MM, Vital M, Mohney RP, Spector TD, Valdes AM. Circulating levels of the anti-oxidant indoleproprionic acid are associated with higher gut microbiome diversity. Gut Microbes. 2019;10(6):688–695.

26. Morris DJ, Brem AS. Role of gut metabolism of adrenal corticosteroids and hypertension: clues gut-cleansing antibiotics give us. Physiol Genomics. 2019;51(3):83–89. doi:10.1152/physiolgenomics.00115.2018.

27. Anderson RA, Baillie TA, Axelson M, Cronholm T, Sjovall K, Sjovall J. Stable isotope studies on steroid metabolism and kinetics: sulfates of 3 alpha-hydroxy-5 alpha-pregnane derivatives in human pregnancy. Steroids. 1990;55(10):443–457. doi:10.1016/0039-128X(90)90013-2.

28. Wang M, Seippel L, Purdy RH, Backstrom T. Relationship between symptom severity and steroid variation in women with premenstrual syndrome: study on serum pregnenolone, pregnenolone sulfate, 5 alpha-pregnane-3,20-dione and 3 alpha-hydroxy-5 alpha-pregnan-20-one. J Clin Endocrinol Metab. 1996;81(3):1076–1082. doi:10.1210/jcem.81.3.8772579.

29. Alexander D, Lombardi R, Rodriguez G, Mitchell MM, Marian AJ. Metabolomic distinction and insights into the pathogenesis of human primary dilated cardiomyopathy. Eur J Clin Invest. 2011;41(5):527–538. doi:10.1111/j.1365-2362.2010.02441.x.

30. Li X, Ning Y, Liu D, Yan A, Wang Z, Wang S, Miao M, Zhu H, Jia Y. Metabolic mechanism of phenyllactic acid naturally occurring in chinese pickles. Food Chem. 2015;186:265–270. doi:10.1016/j.foodchem.2015.01.145.

31. Tuberoso CI, Bifulco E, Caboni P, Sarais G, Cottiglia F, Floris I. Lumichrome and phenyllactic acid as chemical markers of thistle (galactites tomentosa moench) honey. J Agric Food Chem. 2011;59(1):364–369. doi:10.1021/jf1039074.

32. Zhu YB, Xu Y, Wang LM, Qi B. Biosynthesis of (r)-2-hydroxy-3-phenylpropionic acid using whole recombinant escherichia coli cells in an aqueous/n-octane biphasic system. J Zhejiang Univ Sci B. 2018;19(4):285–292. doi:10.1631/jzus.B1700040.

33. Gutierrez-Diaz I, Fernandez-Navarro T, Salazar N, Bartolome B, Moreno-Arribas MV, Lopez P, Suarez A, de Los Reyes-gavilan CG, Gueimonde M, Gonzalez S. Could fecal phenylacetic and phenylpropionic acids be used as indicators of health status? J Agric Food Chem. 2018;66(40):10438–10446. doi:10.1021/acs.jafc.8b04102.

34. McMurray F, Patten DA, Harper ME. Reactive oxygen species and oxidative stress in obesity-recent findings and empirical approaches. Obesity (Silver Spring). 2016;24(11):2301–2310. doi:10.1002/oby.21654.

35. Adisakwattana S. Cinnamic acid and its derivatives: mechanisms for prevention and management of diabetes and its complications. Nutrients. 2017;9(2):2. doi:10.3390/nu9020163.

36. Suchy-Dicey AM, Laha T, Hoofnagle A, Newitt R, Sirich TL, Meyer TW, Thummel KE, Yanez ND, Himmelfarb J, Weiss NS, et al. Tubular secretion in ckd. J Am Soc Nephrol JASN. 2016;27(7):2148–2155. doi:10.1681/ASN.2014121193.

37. Obrenovich ME, Tima M, Polinkovsky A, Zhang R, Emancipator SN, Donskey CJ. Targeted metabolomics analysis identifies intestinal microbiota-derived urinary biomarkers of colonization resistance in antibiotic-treated mice. Antimicrob Agents Chemother. 2017;61(8):8. doi:10.1128/AAC.00477-17.

38. Moayyeri A, Hammond CJ, Valdes AM, Spector TD. Cohort profile: twinsUK and healthy ageing twin study. Int J Epidemiol. 2013;42(1):76–85. doi:10.1093/ije/dyr207.

39. Rebholz CM, Yu B, Zheng Z, Chang P, Tin A, Kottgen A, Wagenknecht LE, Coresh J, Boerwinkle E, Selvin E. Serum metabolomic profile of incident diabetes. Diabetologia. 2018;61(5):1046–1054. doi:10.1007/s00125-018-4573-7.

40. Selvin E, Steffes MW, Zhu H, Matsushita K, Wagenknecht L, Pankow J, Coresh J, Brancati FL. Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. N Engl J Med. 2010;362(9):800–811. doi:10.1056/NEJMoa0908359.

41. Prevention CfDCa. National diabetes statistics report, 2020. In: Services UDoHaH. Atlanta, USA: Centers for Disease Control and Preventio; 2020.

42. Long T, Hicks M, Yu HC, Biggs WH, Kirkness EF, Menni C, Zierer J, Small KS, Mangino M, Messier H, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet. 2017;49(4):568–578. doi:10.1038/ng.3809.

43. Donia MS, Fischbach MA. Human microbiota. Small molecules from the human microbiota. Science. 2015;349(6246):1254766. doi:10.1126/science.1254766.

44. Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, Spector TD, Bell JT, Clark AG, Ley RE. Genetic determinants of the gut microbiome in UK twins. Cell Host Microbe. 2016;19(5):731–743. doi:10.1016/j.chom.2016.04.017.

45. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. Uchime improves sensitivity and speed of chimera detection. Bioinformatics. 2011;27(16):2194–2200. doi:10.1093/bioinformatics/btr381.

46. Jackson MA, Bell JT, Spector T, Steves C. A heritability-based comparison of methods used to cluster 16s rrna gene sequences into operational taxonomic units. Peer J Preprints. 2016. doi:10.7717/peerj.2341.

47. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, et al. Qiime allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335–336. doi:10.1038/nmeth.f.303.

48. Hair JF, Sarstedt M, Ringle CM, Mena JA. An assessment of the use of partial least squares structural equation modeling in marketing research. J Acad Market Sci. 2012;40(3):414–433. doi:10.1007/s11747-011-0261-6.

49. Cohen J. Statistical power analysis for the behavioral sciences. Associates LE, editor. NJ; USA: Hillsdale; 1988.

50. Pavlou PA, Fygenson M. Understanding and predicting electronic commerce adoption: an extension of the theory of planned behavior. Mis Quart. 2006;30(1):115–143. doi:10.2307/25148720.

51. Nitzl C, Roldan JL, Cepeda G. Mediation analysis in partial least squares path modeling helping researchers discuss more sophisticated models. Ind Manage Data Syst. 2016;116(9):1849–1864. doi:10.1108/IMDS-07-2015-0302.