

Serum acylcarnitines and amino acids and risk of type 2 diabetes in a multiethnic Asian population

Samuel H Gunther ¹, Chin Meng Khoo,² E-Shyong Tai,³ Xueling Sim ¹, Jean-Paul Kovalik,⁴ Jianhong Ching,⁴ Jeannette J Lee,¹ Rob M van Dam^{1,5,6}

To cite: Gunther SH, Khoo CM, Tai E-S, *et al*. Serum acylcarnitines and amino acids and risk of type 2 diabetes in a multiethnic Asian population. *BMJ Open Diab Res Care* 2020;**8**:e001315. doi:10.1136/bmjdr-2020-001315

► Supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjdr-2020-001315>).

Received 25 February 2020
Revised 4 June 2020
Accepted 24 August 2020

ABSTRACT

Introduction We evaluated whether concentrations of serum acylcarnitines and amino acids are associated with risk of type 2 diabetes and can improve predictive diabetes models in an Asian population.

Research design and methods We used data from 3313 male and female participants from the Singapore Prospective Study Program cohort who were diabetes-free at baseline. The average age at baseline was 48.0 years (SD: 11.9 years), and participants were of Chinese, Malay, and Indian ethnicity. Diabetes cases were identified through self-reported physician diagnosis, fasting glucose and glycated hemoglobin concentrations, and linkage to national disease registries. We measured fasting serum concentrations of 45 acylcarnitines and 14 amino acids. The association between metabolites and incident diabetes was modeled using Cox proportional hazards regression with adjustment for age, sex, ethnicity, height, and parental history of diabetes, and correction for multiple testing. Metabolites were added to the Atherosclerosis Risk in Communities (ARIC) predictive diabetes risk model to assess whether they could increase the area under the receiver operating characteristic curve (AUC).

Results Participants were followed up for an average of 8.4 years (SD: 2.1 years), during which time 314 developed diabetes. Branched-chain amino acids (HR: 1.477 per SD; 95% CI 1.325 to 1.647) and the alanine to glycine ratio (HR: 1.572; 95% CI 1.426 to 1.733) were most strongly associated with diabetes risk. Additionally, the acylcarnitines C4 and C16-OH, and the amino acids alanine, combined glutamate/glutamine, ornithine, phenylalanine, proline, and tyrosine were significantly associated with higher diabetes risk, and the acylcarnitine C8-DC and amino acids glycine and serine with lower risk. Adding selected metabolites to the ARIC model resulted in a significant increase in AUC from 0.836 to 0.846.

Conclusions We identified acylcarnitines and amino acids associated with risk of type 2 diabetes in an Asian population. A subset of these modestly improved the prediction of diabetes when added to an established diabetes risk model.

INTRODUCTION

Type 2 diabetes mellitus is a significant public health challenge, with prevalence rates increasing globally as a result of urbanization, reduced physical activity, changing dietary patterns, and increasing obesity.¹ Over 60%

Significance of this study

What is already known about this subject?

► Metabolomics studies can identify potential biomarkers of insulin resistance and type 2 diabetes, but data from Asian populations are limited.

What are the new findings?

- A panel of serum acylcarnitines (C4, C8-DC, and C16-OH) and amino acids (alanine, glutamate, glutamine, glycine, isoleucine, leucine, ornithine, phenylalanine, proline, tyrosine, valine, and the alanine to glycine ratio) were associated with type 2 diabetes risk in an Asian population.
- A subset of these metabolites modestly improved diabetes risk prediction when added to an established risk function.
- The associations between alanine and tyrosine and diabetes risk may be partly mediated by glucose metabolism, and those of glutamate and glutamine and diabetes risk by central adiposity.

How might these results change the focus of research or clinical practice?

- Selected acylcarnitines and amino acids may play a role in the development of type 2 diabetes and could be potential therapeutic targets.
- Further research is required to evaluate whether these novel biomarkers provide a clinically significant improvement in diabetes risk prediction.

of all diabetes cases occur in Asia. Individuals of Asian ethnicity experience a higher risk of diabetes, as ethnic predisposition has been observed to interact with other known risk factors such as age and body mass index (BMI).^{2,3} However, prospective studies of incident diabetes often do not include participants of Asian ethnicity or consider ethnic differences.

Metabolomics—the targeted or untargeted analysis and profiling of components of the metabolic pathway—is emerging as a useful method to better understand the pathogenesis and improve the prediction of type 2 diabetes.⁴ Metabolomics studies have



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to Samuel H Gunther; shgunther@u.nus.edu

identified several compounds, including acylcarnitines and amino acids, that are associated with insulin resistance and incident type 2 diabetes.^{4,5} However, data from Asian populations are sparse. Previous studies using targeted metabolomics analysis have implicated amino acid and acylcarnitine signatures associated with future risk of type 2 diabetes in Asian populations.^{6,7} In addition, a panel of amino acids were associated with insulin resistance independent of BMI in a population of ethnic Chinese and Indians.⁸

The aim of this study was therefore to conduct metabolomics analyses to identify risk factors for type 2 diabetes in a multiethnic Asian population. We measured a panel of circulating acylcarnitines and amino acids using targeted metabolomics and evaluated their association with incident diabetes. We also evaluated whether adding these biomarkers to the established Atherosclerosis Risk in Communities (ARIC) predictive diabetes model⁹ would improve the prediction of diabetes.

RESEARCH DESIGN AND METHODS

Study population and design

The study population consisted of participants from the Singapore Prospective Study Program (SP2). The methodology of SP2, a population-based study conducted in Singapore between 2004 and 2007, has been previously described in detail.¹⁰ Briefly, recruited individuals had participated in any of four previous, population-based, cross-sectional surveys: the Thyroid and Heart Study 1982–1984,¹¹ the National Health Survey 1992,¹² the National University of Singapore Heart Study 1993–1995,¹³ and the National Health Survey 1998.¹⁴ Participants were contacted at their homes for an interview and health examination. The interview consisted of standardized questionnaires on lifestyle factors and medical history, while the health examination included physical measurements and collection of fasting blood samples. The health examination did not include thyroid or liver function tests, or measures of blood urea nitrogen. Of the 11 053 original participants of the four studies, 10 445 were eligible for SP2. The 517 original participants who had died, 6 who had emigrated, and 85 who had recorded an error in identity card during the original study were considered ineligible. There were 7742 participants who completed the questionnaire, of whom 5157 subsequently participated in the health examination. The main reason for non-response was being uncontactable during the study period ($n=2673$), with a smaller portion having refused to participate ($n=30$). For this study, relevant baseline data, including sociodemographic information, measures of glycemic control, and serum metabolite measurements, were available for 4451 participants.

Since the outcome of interest was incident diabetes, further exclusion of participants was based on diabetes status at baseline. Of the 4451 participants, 431 reported having diabetes at baseline. Additionally, 159 participants were classified to have undiagnosed diabetes based on their

baseline fasting plasma glucose (FPG ≥ 7.0 mmol/L) and/or baseline glycosylated hemoglobin (HbA_{1c} $\geq 6.5\%$; 48 mmol/mol) measurements.¹⁵ Furthermore, 419 participants did not provide consent for their diabetes status to be tracked over time, leaving 3442 participants available for the study at baseline. A follow-up examination was conducted between 2011 and 2016, which also consisted of both an interview at home and a health examination. Of the baseline participants, 129 were lost to follow-up, leaving 3313 participants with revisit and diabetes outcome data.

Assessment of incident diabetes

Diabetes assessment over the course of the follow-up period was based on one or more of the following criteria: self-reported physician diagnosis of diabetes during follow-up interview, FPG (≥ 7.0 mmol/L) or HbA_{1c} ($\geq 6.5\%$; 48 mmol/mol) concentrations, or diabetes diagnosis reported in a nationwide health database covering subsidized general practitioners, polyclinics, and public hospitals, and updated on a regular basis. The dates of diagnosis through any of these three criteria were linked with initial SP2 interview date to calculate time-to-event.

Assessment of serum metabolites

For extraction of acylcarnitines and amino acids, 50 μ L of serum was spiked with deuterium-labeled acylcarnitine and amino acid standards, including D3-C2, D3-C3, D3-C4, D9-C5, D3-C8, and D3-C16 carnitines (Cambridge Isotope Laboratories, Andover, Massachusetts, USA), and alanine-D4, arginine-D5, citrulline-D2, glutamic acid-D3, glycine-D2, histidine-D6, leucine-D3, methionine-D3, ornithine-D2, phenylalanine-D5, proline-D7, serine-D3, tryptophan-D5, and valine-D8 (Sigma Aldrich, USA). The mixture was extracted using methanol. The extracts were then derivatized with 3M hydrochloric acid in methanol or butanol (Sigma Aldrich), respectively, dried, and reconstituted in methanol for analysis in liquid chromatography-mass spectrometry (LC-MS).

Acylcarnitine measurements were performed without a column, using flow injection tandem mass spectrometry on the Agilent 6430 Triple Quadrupole LC/MS System (Agilent Technologies, California, USA). The sample analysis was carried out at 0.4 mL/min of 80/20 methanol/water as mobile phase, and injection of 4 μ L of sample. Data acquisition and analysis were performed on Agilent MassHunter Workstation V.B.06.00 software. Amino acids were separated using a C18 column (Phenomenex, 100 \times 2.1 mm, 1.6 μ m, Luna Omega) on an Agilent 1290 Infinity LC System (Agilent Technologies) coupled with quadrupole-ion trap mass spectrometer (QTRAP 5500, AB Sciex, DC, USA). Mobile phase A (water) and mobile phase B (acetonitrile), both containing 0.1% formic acid, were used for chromatography separation. The LC run was performed at a flow rate of 0.4 mL/min with initial gradient of 2% B for 0.8 min, then increased to 15% B in 0.1 min, 20% B in 5.7 min, 50% B in 0.5 min, and 70% B in 0.5 min, followed by re-equilibration of the column to the initial

run condition of 2% B for 0.9 min. All compounds were ionized in positive mode using electrospray ionization. The chromatograms were integrated using MultiQuant V.3.0.3 software (AB Sciex).

Absolute quantification of both acylcarnitines and amino acids was done by comparing the ratios of the metabolites with their respective internal standards, against an external calibration curve. External calibration curves consisted of C2, C3, C4, C5, C6, C8, C10, C12, C14, C16, and C18 carnitines and all reported amino acids.

Quality control (QC) was performed on the metabolite data prior to statistical analysis. To reduce potential batch effect, data were log₂-transformed and normalized in two steps: first by adjusting the intensity levels of each metabolite according to the QC sample runs in each batch, and second by equalizing the average log₂ concentration across the batches. Next, all metabolites with a coefficient of variation (CV)% of more than 20% for duplicate measurements were excluded from further analysis. The proportion of values either missing or below the limit of detection was determined for each remaining metabolite, and all metabolites with more than 5% missing data were also excluded from further analysis. Finally, missing values for the remaining metabolites were imputed using k-nearest neighbor imputation, whereby missing values for a given metabolite were replaced with an average of the non-missing values for that metabolite from participants similar in terms of the other variables. See online supplemental table 1 for detailed QC metrics for each metabolite.

Prior to QC, the metabolite panel consisted of 59 metabolites: 45 acylcarnitine and 14 amino acid measurements. As no column separation was used for the acylcarnitine analysis, certain pairs of isomeric acylcarnitines were reported as the sum of both species. Isoleucine and leucine, and glutamate and glutamine, were similarly measured in aggregate. Following QC, 37 acylcarnitines and all 14 amino acids remained available for analysis. Five acylcarnitine measurements, namely the C12-OH/C10-DC, C18:1-OH/C16:1-DC, and C18-OH/C16-DC aggregate measures, in addition to C4-OH and C22, were removed due to having a CV% greater than 20%. Three more, namely C5, C8:1-DC, and the C8:1-OH/C6:1-DC aggregate, were removed due to having too many values either missing or below the limit of detection.

The 37 remaining acylcarnitine measurements were combined into three functional groups based on chain length: short chain (species with carbon chain length 8 or shorter), medium chain (10–14) and long chain (16 or longer). Similarly, certain amino acids were combined into two functional groups: aromatic amino acids and branched-chain amino acids. Additionally, we examined the Fischer ratio (ratio of the sum of isoleucine, leucine, and valine to the sum of tyrosine and phenylalanine) and the ratio of alanine to glycine, two established metabolite ratios associated with diabetes.^{16 17} Online supplemental table 2 provides details of the full metabolite panel, functional groupings, and ratios, including those metabolites excluded from statistical analysis.

Assessment of covariates

Covariates in this study included age (years), sex (male/female), ethnicity (Chinese/Malay/Indian), height (cm), waist circumference (cm), parental history of type 2 diabetes (yes/no), systolic blood pressure (SBP, mm Hg), FPG (mmol/L), serum triglycerides (TG, mmol/L), high-density lipoprotein (HDL)-cholesterol (mmol/L), and fasting insulin (mU/L). Ethnicity was based on participants' national identity cards and recorded as one of three categories representing the three main ethnic groups in Singapore, namely Chinese, Malay, and Indian; participants of mixed ethnicity were categorized as the primary ethnicity listed on their identity card. Height was measured without shoes and with the head in the Frankfurt plane position using a portable stadiometer (SECA, Model 782-2321009; Vogel & Halke, Hamburg, Germany). Waist circumference at the midpoint between the last rib and the iliac crest was measured using stretch-resistant tape. SBP was measured twice using the Dinamap Carescape Monitor (Woodley Equipment, Horwich, Bolton, UK) and the average of the two values was used for analysis. Blood samples were analyzed on the day of collection for FPG, TG, HDL-cholesterol, and fasting insulin at the biochemistry laboratories of the National University Hospital and Singapore General Hospital. Measurements were calibrated between the two sites and the calibrated values used for analysis. See online supplemental table 3 for detailed methodology and QC metrics for biosample analysis.

Data on additional variables of potential interest in relation to diabetes incidence were also collected during the SP2 interview and health examination. These included BMI (kg/m²), serum creatinine (μmol/L), regular use of lipid-lowering or blood pressure-lowering medication (yes/no), and daily intake of the key nutrients carbohydrates, protein, total fat, saturated fat, monounsaturated fat, and polyunsaturated fat (% of total energy).

Statistical analysis

Metabolite data were first log₂-transformed to test various metabolite ratios in our data set using the p-gain statistics, which determines whether a ratio of metabolites contains more information than the constituent metabolites alone.¹⁸ The significance level for the p-gain test was set at 0.05. Concentrations were then exponentiated back to their original value to be standardized and converted into Z-scores. The ratios were also standardized and converted to Z-scores for consistency of interpretation. The association between metabolites and time to diabetes development was modeled using Cox regression. Two models were used for the main analysis: an unadjusted model, and a multivariable model adjusting for the non-modifiable risk factors age, sex, ethnicity, height, and parental history of diabetes. β-values from the models were exponentiated to produce HR. Each metabolite, aggregate pair, or ratio was evaluated in both models, and multiple testing correction was applied to all results using the Benjamini-Hochberg procedure, controlling the false discovery rate at an α-value of 0.05.

Several additional Cox regression models were developed for secondary analyses in order to assess the impact of modifiable risk factors on the observed associations and to explore potential biological mechanisms underlying the development of diabetes. For these analyses, the multivariable Cox model used in the main analysis was further adjusted for SBP, waist circumference, levels of blood lipids (TG and HDL-cholesterol), and levels of glycemic markers (FPG and fasting insulin).

Metabolites that remained significantly associated with incident diabetes in the main multivariable model were identified as potential biomarkers and incorporated into the ARIC risk model to determine whether they could improve predictive ability. The ARIC model was selected for this study based on a previous investigation by Chin *et al.*¹⁹ in which the ARIC model was shown to better predict diabetes risk in the Singapore population than the San Antonio Health Study model and the Framingham model. In addition to sociodemographic factors, the ARIC model includes the risk factors waist circumference, SBP, FPG, TG, and HDL-cholesterol. In accordance with the binary way in which ethnicity is identified in the ARIC model, participants belonging to the minority Malay and Indian ethnic groups were combined into a single 'non-Chinese' category during risk prediction.

Predictive ability was assessed using area under the receiver operating characteristic curve (AUC), with 95% CI for model AUC estimated using the DeLong method for correlated receiver operating characteristic curves. Metabolites were individually added to the ARIC model, and a final, parsimonious model was identified using a backward stepwise approach. An ARIC model was constructed containing all potential biomarkers, with subsequent models removing metabolites according to the lowest β -value, until a final model yielding the most improved predictive value with the best model fit was identified. Models were trained on a randomly selected half of the data set and tested on the other, and model calibration and goodness of fit were assessed using the Akaike information criterion (AIC). As a secondary analysis, models were also assessed in terms of their net reclassification improvement indices, which illustrate the ability of successive predictive models to accurately reclassify cases and non-cases. Finally, sensitivity analyses were conducted using the participants of Chinese ethnicity to identify potential ethnic differences in diabetes prediction. All analyses were performed using R V.3.5.1 (R Core Team, Vienna, Austria).

RESULTS

Of the 3313 study participants in the revisit, 314 (9.5%) developed incident diabetes during follow-up. The mean (SD) duration of follow-up was 8.4 (2.1) years. **Table 1** summarizes the sociodemographic and clinical characteristics of the study population according to diabetes development. Participants who developed diabetes were significantly older, shorter, had a greater waist circumference and BMI, were more likely to have a parental

history of diabetes, and were more likely to be taking lipid-lowering and blood pressure-lowering medication at baseline as compared with non-cases. They also had higher SBP, FPG, TG, fasting insulin, serum creatinine, and HbA_{1c} concentrations, and lower HDL-cholesterol concentrations. A higher proportion of cases were of Malay or Indian ethnicity as compared with non-cases, who were more likely to be of Chinese ethnicity.

Table 2 shows the association between metabolite concentrations and incident diabetes after adjustment for age, sex, ethnicity, height, and parental history of diabetes (see online supplemental table 4 for unadjusted associations). In the multivariable model, three acylcarnitines (C4, C8-DC, and C16-OH), 10 amino acid measurements (alanine, glutamate/glutamine, glycine, isoleucine/leucine, ornithine, phenylalanine, proline, serine, tyrosine, and valine), and the alanine to glycine ratio, as well as the sum molar of the aromatic and branched-chain groups, were significantly associated with incident diabetes. Of these metabolites, C8-DC, glycine, and serine were inversely associated with diabetes risk, whereas the other metabolites were directly associated with diabetes risk.

To investigate the collinearity of the statistically significant biomarkers, which could potentially cause model overfitting, we generated a heatmap of the pairwise partial correlations between the biomarkers, controlling for other diabetes risk factors (**figure 1**; see online supplemental figure 1 for unadjusted correlations). The highest degree of collinearity occurred among the branched-chain amino acids isoleucine/leucine and valine ($r=0.866$). The branched-chain species were also moderately correlated with phenylalanine (isoleucine/leucine $r=0.687$, valine $r=0.644$). The pairwise correlations among the remaining biomarkers were not strong (r for all <0.600).

We also evaluated whether the observed associations between metabolites and diabetes incidence could be explained by established diabetes risk factors that may act as mediators (**table 3**). Building on the multivariable Cox model used for the main analysis, four models were constructed additionally adjusting for SBP, waist circumference, blood lipids (TG and HDL-cholesterol), or glycemic markers (FPG and fasting insulin). Attenuation of associations was defined as a change in the β -value of at least 10% toward the null. Addition of SBP to the model did not substantially attenuate any of the associations. Adjustment for waist circumference attenuated associations for glutamate/glutamine, the branched-chain amino acids (isoleucine/leucine and valine individually and combined), and the alanine to glycine ratio. Adjustment for blood lipids attenuated associations for alanine, glutamate/glutamine, the branched-chain species, and the alanine to glycine ratio. Finally, adjusting for glycemic markers attenuated associations for alanine, tyrosine, and the branched-chain species. Some of the acylcarnitines (C4 and C8-DC) and amino acids (proline and ornithine) with weaker associations in the basic models lost statistical significance after further adjustments even though their β -values changed less than 10%.

Table 1 Baseline characteristics of participants who did and did not develop diabetes mellitus (DM)

	Overall (N=3313)	Follow-up DM cases (n=314)	Follow-up non-cases (n=2999)	P value
Age (years), mean (SD)	48.0 (11.9)	53.5 (11.2)	47.0 (11.6)	<0.001
Sex (female), n (%)	1795 (54.2)	162 (51.6)	1633 (54.5)	0.333
Ethnicity, n (%)				
Chinese	2577 (77.8)	209 (66.6)	2368 (79.0)	<0.001
Malays	422 (12.7)	62 (19.7)	360 (12.0)	
Indians	314 (9.5)	43 (13.7)	271 (9.0)	
Height (cm), mean (SD)	163.4 (8.8)	162.0 (8.7)	163.5 (8.8)	<0.001
Waist circumference (cm), mean (SD)	82.7 (11.6)	90.8 (12.1)	81.7 (11.2)	<0.001
BMI (kg/m ²), mean (SD)	23.4 (4.0)	26.3 (4.8)	23.1 (3.8)	<0.001
Dietary intake (energy %/day)				
Carbohydrates, mean (SD)	53.4 (6.8)	53.3 (7.0)	53.5 (6.7)	0.949
Protein, mean (SD)	15.6 (2.3)	15.6 (2.4)	15.5 (2.3)	0.922
Total fat, mean (SD)	30.8 (5.8)	31.0 (6.1)	30.9 (5.8)	0.788
Saturated fat, mean (SD)	11.8 (3.2)	11.8 (3.4)	11.9 (3.2)	0.922
Monounsaturated fat, mean (SD)	10.5 (2.7)	10.5 (2.9)	10.6 (2.7)	0.826
Polyunsaturated fat, mean (SD)	6.4 (2.7)	6.6 (3.0)	6.4 (2.7)	0.284
Parental history of diabetes (yes), n (%)	930 (27.0)	106 (33.8)	824 (27.5)	0.018
Lipid-lowering and blood pressure-lowering medications (yes), n (%)	760 (22.9)	124 (39.5)	636 (21.2)	<0.001
Systolic blood pressure (mm Hg), mean (SD)	130.5 (19.8)	141.9 (21.0)	128.9 (19.2)	<0.001
Fasting plasma glucose (mmol/L), mean (SD)	4.73 (0.49)	5.15 (0.56)	4.68 (0.46)	<0.001
Triglycerides (mmol/L), mean (SD)	1.31 (0.82)	1.74 (0.98)	1.26 (0.79)	<0.001
High-density lipoprotein (mmol/L), mean (SD)	1.44 (0.36)	1.31 (0.31)	1.46 (0.36)	<0.001
Fasting insulin (mU/L), mean (SD)	7.42 (6.14)	10.67 (6.33)	7.04 (6.02)	<0.001
Glycated hemoglobin (DCCT %), mean (SD)	5.48 (0.41)	5.82 (0.46)	5.44 (0.39)	<0.001
Serum creatinine (µmol/L), mean (SD)	9.35 (6.46)	10.18 (6.53)	9.25 (6.46)	0.018

BMI, body mass index; DCCT, Diabetes Control and Complications Trial units.

We next evaluated whether adding acylcarnitine and amino acid metabolites to the ARIC model could improve the prediction of incident diabetes based on AUC and AIC metrics (table 4). The ARIC model alone had an AUC of 0.836 and an AIC of 1559.4. Addition of the metabolites individually did not result in significant improvement in AUC. However, addition of all 14 metabolite measurements to the ARIC model resulted in a modest but statistically significant improvement in AUC to 0.847 (p=0.013) and a lower AIC (1541.9). Step-wise removal of the biomarkers produced an ARIC model supplemented with seven metabolite measurements (C8-DC, C16-OH, isoleucine/leucine, ornithine, proline, serine, and the alanine to glycine ratio), which had the best model fit (AIC: 1530.3). This model also represented a significant improvement in AUC (0.846, p=0.022) as compared with the ARIC model alone, and recorded a net improvement of 39.8% in reclassification of diabetes cases and non-cases.

Finally, we conducted sensitivity analyses in the 3095 participants who were not pre-diabetic (pre-diabetes defined by the American Diabetes Association criteria of FPG >5.6 mmol/L),¹⁵ which produced similar results (data not shown). In an additional sensitivity analysis, we only included ethnic Chinese, who made up 77.8% of the study participants. All of the more significant associations in the original analysis (p<0.01 after multiple testing correction) remained significant in the ethnic Chinese subgroup (online supplemental table 5).

DISCUSSION

In this prospective cohort study, we identified several serum acylcarnitine and amino acid metabolites with the potential to serve as biomarkers of type 2 diabetes in Asian populations. Higher concentrations of the acylcarnitines C4 and C16-OH, alanine, glutamate/glutamine, ornithine, proline, the branched-chain amino acids isoleucine/leucine and valine, and the aromatic amino acids

Table 2 Associations between metabolites and incident diabetes, adjusted for age, sex, ethnicity, height, and parental history of diabetes

Acylcarnitines	HR (95% CI)*	P value	P value†	Amino acids	HR (95% CI)	P value	P value†
C2	1.041 (0.932 to 1.162)	0.478	0.603	Alanine	1.377 (1.236 to 1.534)	<0.001	<0.001
C3	1.050 (0.941 to 1.171)	0.384	0.557	Arginine	0.888 (0.791 to 0.996)	0.043	0.1
C4	1.146 (1.037 to 1.265)	0.007	0.027	Citrulline	0.909 (0.808 to 1.022)	0.111	0.195
C4-DC	1.112 (0.995 to 1.242)	0.061	0.126	Glutamate/glutamine	1.421 (1.281 to 1.576)	<0.001	<0.001
C5:1	1.093 (0.980 to 1.218)	0.111	0.195	Glycine	0.710 (0.621 to 0.810)	<0.001	<0.001
C5-DC	1.072 (0.966 to 1.190)	0.192	0.318	Histidine	0.937 (0.836 to 1.050)	0.142	0.242
C5-OH/C3-DC	1.021 (0.914 to 1.140)	0.712	0.753	Isoleucine/leucine	1.507 (1.349 to 1.683)	<0.001	<0.001
C6	1.042 (0.936 to 1.160)	0.452	0.601	Methionine	1.076 (0.958 to 1.207)	0.216	0.348
C7-DC	0.872 (0.777 to 0.979)	0.021	0.061	Ornithine	1.196 (1.071 to 1.336)	0.001	0.005
C8	0.860 (0.755 to 0.980)	0.024	0.063	Phenylalanine	1.268 (1.142 to 1.407)	<0.001	<0.001
C8:1	1.124 (1.020 to 1.240)	0.019	0.058	Proline	1.181 (1.059 to 1.317)	0.003	0.013
C8-DC	0.853 (0.754 to 0.966)	0.012	0.044	Serine	0.837 (0.742 to 0.944)	0.004	0.017
C8-OH/C6-DC	0.869 (0.769 to 0.980)	0.023	0.063	Tyrosine	1.338 (1.208 to 1.483)	<0.001	<0.001
C10	0.848 (0.742 to 0.969)	0.016	0.055	Valine	1.419 (1.275 to 1.579)	<0.001	<0.001
C10:1	0.875 (0.773 to 0.990)	0.035	0.088	Aromatic	1.200 (1.082 to 1.332)	0.001	0.005
C10:2	0.984 (0.879 to 1.102)	0.78	0.796	Branched-chain	1.477 (1.325 to 1.647)	<0.001	<0.001
C10:3	1.099 (0.998 to 1.209)	0.056	0.12	Fischer ratio	1.028 (0.917 to 1.153)	0.633	0.734
C12	0.896 (0.797 to 1.008)	0.068	0.136	Alanine:glycine	1.572 (1.426 to 1.733)	<0.001	<0.001
C12:1	0.882 (0.784 to 0.993)	0.037	0.089				
C14	0.976 (0.873 to 1.092)	0.675	0.751				
C14:1	0.942 (0.840 to 1.056)	0.305	0.466				
C14:1-OH	0.957 (0.854 to 1.073)	0.455	0.601				
C14:2	0.900 (0.801 to 1.013)	0.08	0.155				
C14-OH/C12-DC	1.052 (0.940 to 1.176)	0.379	0.557				
C16	1.096 (0.986 to 1.219)	0.091	0.17				
C16:1	0.980 (0.877 to 1.094)	0.714	0.753				
C16:1-OH/C14:1-DC	1.012 (0.905 to 1.130)	0.839	0.839				
C16:2	0.960 (0.858 to 1.074)	0.474	0.603				
C16-OH	1.176 (1.060 to 1.304)	0.002	0.01				
C18	0.969 (0.866 to 1.084)	0.582	0.718				
C18:1	1.016 (0.910 to 1.134)	0.782	0.796				
C18:1-DC	1.023 (0.916 to 1.142)	0.686	0.751				
C18:2	1.029 (0.922 to 1.149)	0.606	0.72				
C18:2-OH	0.957 (0.852 to 1.075)	0.456	0.601				
C18-DC/C20-OH	1.025 (0.918 to 1.144)	0.665	0.751				
C20	0.863 (0.763 to 0.975)	0.018	0.058				
C20:4	1.061 (0.951 to 1.183)	0.292	0.458				
Short chain	1.044 (0.935 to 1.166)	0.443	0.601				
Medium chain	0.885 (0.783 to 1.000)	0.05	0.112				
Long chain	1.029 (0.922 to 1.150)	0.608	0.72				

Significant associations following multiple testing are in bold.

*As metabolite data were converted to Z-scores, HR is measured in per SD.

†Adjusted for multiple testing using the Benjamini-Hochberg false discovery rate.

tyrosine and phenylalanine were associated with higher diabetes risk. A higher alanine to glycine ratio was also associated with higher diabetes risk. In contrast, higher concentrations of serine, glycine, and C8-DC acylcarnitine

were associated with lower diabetes risk. Adjustment for known metabolic risk factors (blood lipids and glycemic markers) partially explained the associations with diabetes risk for alanine, glutamate/glutamine, tyrosine,

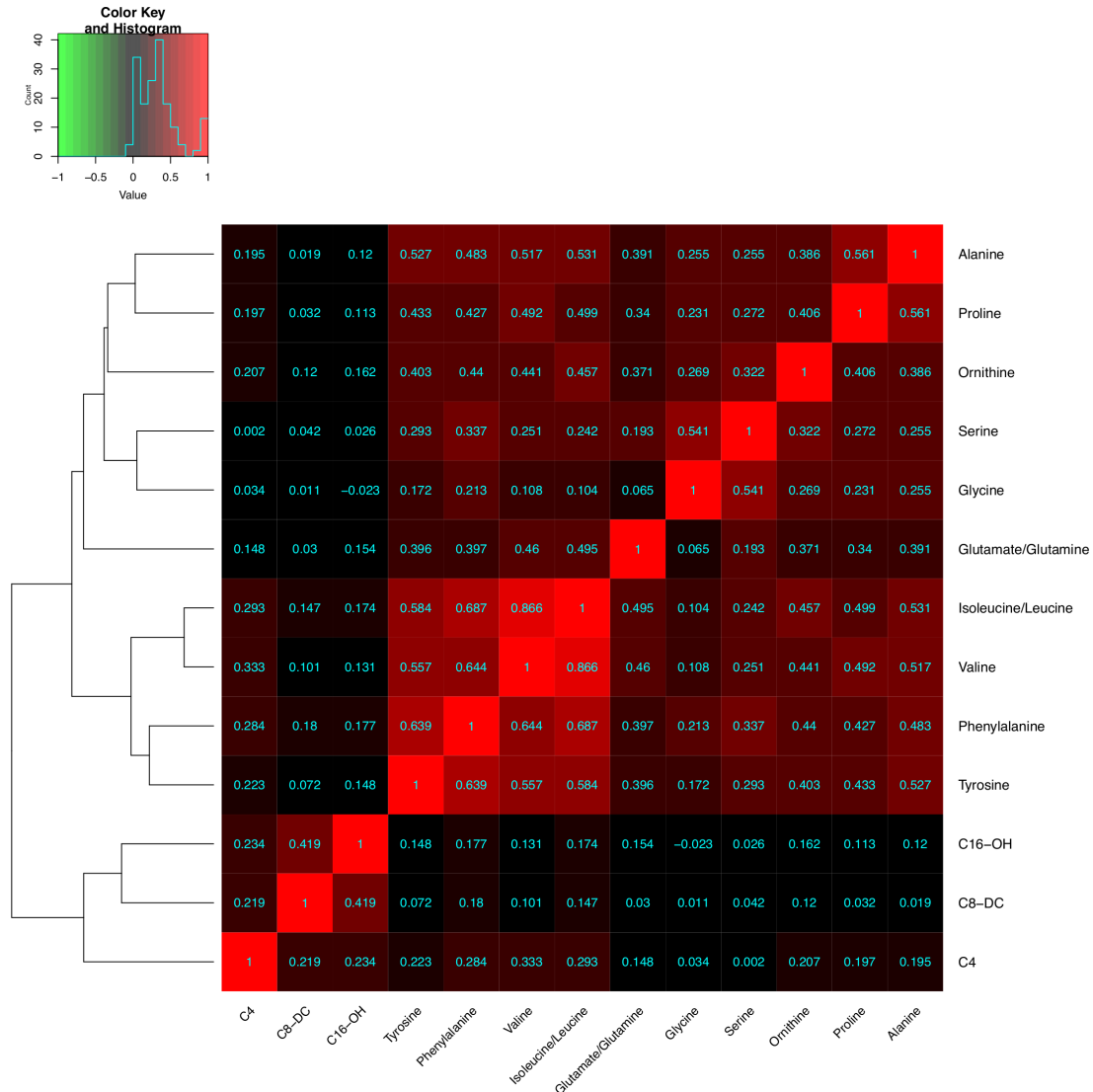


Figure 1 Heatmap of the pairwise partial correlations between the 13 potential biomarkers, adjusted for age, sex, ethnicity, height, waist circumference, parental history of diabetes, systolic blood pressure, fasting plasma glucose, serum triglycerides, high-density lipoprotein cholesterol, and fasting insulin.

the branched-chain species, and the alanine to glycine ratio. Adding a panel of metabolites (C8-DC, C16-OH, isoleucine/leucine, ornithine, proline, serine, and the alanine to glycine ratio) to the ARIC model with established diabetes risk factors led to a modest but statistically significant improvement in the prediction of diabetes.

Previous studies have implicated several of the amino acids associated with diabetes in our study as potential biomarkers of insulin resistance and type 2 diabetes. Higher levels of branched-chain amino acids have previously been linked to higher diabetes risk in populations of European, Hispanic, African, and Asian ancestry.^{5 20–22} Furthermore, a large-scale Mendelian randomization analysis identified genetic instruments reflective of higher levels of circulating branched-chain species that were also associated with higher diabetes risk, suggesting a causal role of branched-chain amino acid metabolism in diabetes development.²³ These findings are consistent with knowledge of biological mechanisms involved in diabetes development.

Branched-chain species play a central role in the PI3K-AKT-mTOR signaling pathway by regulating expression of genes and phosphorylation of kinases involved in glucose and lipid metabolism.²⁴ Metabolic imbalance and overexpression of branched-chain species lead to phosphorylation of insulin receptor substrate (IRS)-1, which interferes with insulin signaling and over time leads to insulin resistance.⁵ A related amino acid group to the branched-chain species is the aromatic amino acids, consisting of phenylalanine, tyrosine, and tryptophan (not measured in this study). The aromatic and branched-chain species share a transmembrane protein,²⁵ and higher levels of the five amino acids have been observed to be associated with higher diabetes risk in multiple previous studies as well as our own.^{20 22 26 27} It has been proposed that tyrosine can inhibit glucose transport and phosphorylation,²⁸ a hypothesis supported by our finding that additional adjusting for FPG and fasting insulin attenuates the association between tyrosine and diabetes risk.

Table 3 Associations between serum biomarkers and type 2 diabetes after adjustment for potential mediators

Metabolite	Model 1*	Model 1 and systolic blood pressure	Model 1 and waist circumference	Model 1 and blood lipids†	Model 1 and glycemic markers‡
	HR§ (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
C4	1.146 (1.037 to 1.265)	1.128 (1.021 to 1.245)	1.101 (0.996 to 1.218)	1.061 (0.957 to 1.177)	1.120 (1.006 to 1.246)
C8-DC	0.853 (0.754 to 0.966)	0.824 (0.728 to 0.934)	0.888 (0.783 to 1.006)	0.852 (0.755 to 0.962)	0.911 (0.802 to 1.035)
C16-OH	1.176 (1.060 to 1.304)	1.140 (1.027 to 1.265)	1.167 (1.053 to 1.294)	1.117 (1.007 to 1.238)	1.146 (1.029 to 1.277)
Alanine	1.377 (1.236 to 1.534)	1.353 (1.213 to 1.509)	1.317 (1.178 to 1.472)	1.256 (1.121 to 1.408)	1.236 (1.098 to 1.392)
Glutamate/glutamine	1.421 (1.281 to 1.576)	1.401 (1.261 to 1.555)	1.303 (1.171 to 1.449)	1.307 (1.173 to 1.456)	1.327 (1.187 to 1.484)
Glycine	0.710 (0.621 to 0.810)	0.727 (0.637 to 0.831)	0.777 (0.680 to 0.889)	0.751 (0.657 to 0.858)	0.705 (0.613 to 0.811)
Isoleucine/leucine	1.507 (1.349 to 1.683)	1.492 (1.334 to 1.669)	1.405 (1.251 to 1.577)	1.338 (1.191 to 1.503)	1.399 (1.242 to 1.577)
Ornithine	1.196 (1.071 to 1.336)	1.174 (1.050 to 1.314)	1.212 (1.084 to 1.355)	1.203 (1.073 to 1.349)	1.120 (0.995 to 1.260)
Phenylalanine	1.268 (1.142 to 1.407)	1.241 (1.117 to 1.379)	1.212 (1.089 to 1.349)	1.215 (1.091 to 1.354)	1.222 (1.093 to 1.365)
Proline	1.181 (1.059 to 1.317)	1.168 (1.047 to 1.302)	1.146 (1.026 to 1.280)	1.120 (0.999 to 1.256)	1.090 (0.969 to 1.227)
Serine	0.837 (0.742 to 0.944)	0.851 (0.755 to 0.960)	0.884 (0.782 to 0.999)	0.876 (0.776 to 0.989)	0.837 (0.737 to 0.951)
Tyrosine	1.338 (1.208 to 1.483)	1.300 (1.172 to 1.443)	1.243 (1.118 to 1.381)	1.260 (1.132 to 1.403)	1.235 (1.104 to 1.381)
Valine	1.419 (1.275 to 1.579)	1.392 (1.250 to 1.551)	1.309 (1.171 to 1.462)	1.290 (1.154 to 1.442)	1.281 (1.143 to 1.436)
Aromatic	1.200 (1.082 to 1.332)	1.174 (1.057 to 1.303)	1.157 (1.040 to 1.288)	1.152 (1.033 to 1.285)	1.147 (1.026 to 1.282)
Branched-chain	1.477 (1.325 to 1.647)	1.454 (1.302 to 1.623)	1.364 (1.217 to 1.528)	1.325 (1.182 to 1.485)	1.341 (1.193 to 1.507)
Alanine:glycine	1.572 (1.426 to 1.733)	1.522 (1.378 to 1.680)	1.451 (1.306 to 1.611)	1.447 (1.302 to 1.607)	1.494 (1.340 to 1.664)

Values in italics represent an attenuation of at least 10% toward the null in relation to model 1 values.

*Adjusted for age, sex, ethnicity, height, and parental history of diabetes.

†Serum triglycerides and high-density lipoprotein.

‡Fasting plasma glucose and fasting insulin.

§As metabolite data were converted to Z-scores, HR is measured in per SD.

Branched-chain species serve as nitrogen donors for alanine, glutamate, and glutamine,²⁴ which may partially explain why higher levels of these amino acids were also significantly associated with diabetes risk in our study. That being said, higher serum levels of alanine, but not the branched-chain species, were consistently associated with higher diabetes risk in two Chinese cohorts,⁶ which suggests the link between alanine and diabetes risk is not necessarily due to branched-chain species metabolism. This is also consistent with biological mechanisms, as alanine stimulates glucagon secretion,²⁹ and alterations in alanine metabolism as a manifestation of non-alcoholic fatty liver disease have been linked to higher diabetes

risk.³⁰ The association for alanine was attenuated by further adjusting for blood lipids and glycemic markers in our study, which is consistent with its biological role in gluconeogenesis. Furthermore, higher concentrations of aggregate glutamate/glutamine were associated with insulin resistance and diabetes development in ethnic Chinese and Indian SP2 participants,⁸ and in the Insulin Resistance Atherosclerosis Study,²¹ while higher concentrations of glutamate by itself were associated with insulin resistance phenotypes in the Framingham Heart Study and Malmö Diet and Cancer Study cohorts.³¹ Glutamate has also been shown to stimulate glucagon secretion and gluconeogenesis,³² and serves as a metabolic precursor to

Table 4 Comparisons of AUC, AIC, and NRI for the ARIC model* and models supplemented with the potential biomarkers

Model	AUC (95% CI)	P value†	AIC	NRI‡ (%)
ARIC	0.836 (0.814 to 0.858)	Reference	1559.4 (reference)	Reference
ARIC+C4	0.836 (0.814 to 0.858)	0.794	1561.3	-3.3§
ARIC+C8-DC	0.837 (0.815 to 0.859)	0.462	1558.8	16.0
ARIC+C16-OH	0.837 (0.815 to 0.859)	0.575	1558.6	-4.5
ARIC+alanine	0.836 (0.814 to 0.858)	0.692	1558	7.7
ARIC+glutamate/glutamine	0.837 (0.815 to 0.859)	0.398	1557.1	20.7
ARIC+glycine	0.839 (0.817 to 0.861)	0.208	1550.1	21.4
ARIC+isoleucine/leucine	0.838 (0.816 to 0.860)	0.293	1553.1	24.4
ARIC+ornithine	0.837 (0.815 to 0.859)	0.487	1558.8	5.1
ARIC+phenylalanine	0.837 (0.815 to 0.859)	0.420	1558.8	5.0
ARIC+proline	0.836 (0.814 to 0.858)	0.865	1561.3	4.9
ARIC+serine	0.836 (0.814 to 0.858)	0.853	1559.1	13.8
ARIC+tyrosine	0.836 (0.814 to 0.858)	0.648	1559.6	9.9
ARIC+valine	0.837 (0.815 to 0.858)	0.626	1558.1	14.1
ARIC+alanine to glycine ratio	0.840 (0.819 to 0.862)	0.132	1539.9	32.6
ARIC+C4+C8-DC+C16-OH+alanine+glutamate/glutamine+ glycine+isoleucine/leucine+ornithine+phenylalanine+proline +serine+tyrosine+valine+ alanine to glycine ratio	0.847 (0.825 to 0.868)	0.013	1541.9	45.7
ARIC+C8-DC+C16-OH+isoleucine/ leucine+ornithine+proline+serine+alanine to glycine ratio	0.846 (0.824 to 0.867)	0.022	1530.3	39.8

*Includes age, sex, ethnicity, height, waist circumference, parental history, systolic blood pressure, fasting plasma glucose, serum triglycerides, and high-density lipoprotein cholesterol.

†P value compares the AUC of each model with the reference model.

‡The net improvement in reclassification of diabetes cases and non-cases compared with the ARIC model alone.

§A negative NRI value indicates the supplemented model performs worse at classification of diabetes cases and non-cases than the ARIC model alone.

AIC, Akaike information criterion; ARIC, Atherosclerosis Risk in Communities; AUC, area under the receiver operating characteristic curve; NRI, net reclassification improvement.

α -ketoglutarate, a keto acid with anticatabolic effects on protein metabolism,³³ which again suggests a biological link between glutamate and diabetes risk independent of branched-chain amino acids. In our study, the association for glutamate/glutamine was attenuated by further adjusting for waist circumference and blood lipids, which potentially implicates central adiposity as a mediator in the link between these species and diabetes.

A growing body of evidence supports our finding of an inverse association between circulating glycine levels

and diabetes risk.³⁴ Lower serum concentrations of glycine were associated with higher insulin resistance and diabetes risk in the Insulin Resistance Atherosclerosis Study, the Framingham Heart Study, the Malmö Diet and Cancer Study, the European Prospective Investigation into Cancer and Nutrition - Potsdam, and the Relationship of Insulin Sensitivity to Cardiovascular Risk study cohorts,^{21 31 35 36} while in a Japanese prospective cohort study, baseline concentrations of glycine were lower in participants who developed diabetes compared with

those who did not.³⁷ Furthermore, a Mendelian randomization analysis reported genetic instruments reflecting higher levels of circulating glycine were associated with a lower diabetes risk, suggesting a causal protective effect of glycine on diabetes risk.³⁸ This is also consistent with biological mechanisms, as glycine plays key metabolic roles as a neurotransmitter, in the synthesis of heme and the antioxidant glutathione, and in the regulation of one-carbon metabolism.³⁴ Dysregulation of these pathways from overexpression of glycine is proposed to contribute to insulin resistance by increasing oxidative stress in pancreatic cells, compromising mitochondrial function, and disrupting glucose homeostasis.³⁴ We also observed a significant association between a higher alanine to glycine ratio and diabetes risk. The alanine to glycine ratio was strongly associated with insulin sensitivity measured using a hyperglycemic clamp and incident diabetes in the Cooperative Health Research in the Region of Augsburg S4_to_F4 cohort.¹⁷ Analysis of metabolite ratios is an emerging field that can provide additional information in association studies by reducing overall biological variability in a given study population and better representing biochemical pathways,¹⁸ and our results provide further evidence of their value.

Additionally, we observed significant associations between ornithine and proline concentrations and diabetes risk, a finding also reported in a Japanese study.³⁷ The biological mechanisms underlying this putative relationship are not well understood. Both ornithine and proline are produced by arginase activity during the urea cycle, and upregulated arginase activity, resulting in higher ornithine and proline levels, can decrease nitric oxide bioavailability and lead to metabolic complications including diabetes.³⁹ However, this pathway is mediated by arginine and also results in citrulline biosynthesis, and neither of those species were significantly associated with diabetes risk in our study. Conversely, arginine was associated with diabetes risk in a Japanese cohort,³⁷ and ornithine levels were inversely associated with diabetes risk in a Chinese study.⁴⁰ Likewise, the role of serine in diabetes development is understudied, although the Japanese study did report lower concentrations of serine in participants who developed diabetes compared with those who did not.³⁷ Serine is synthesized by glycine activity, and it is possible that depressed levels in those with higher diabetes risk are reflective of depressed glycine levels and the consequent metabolic imbalances.⁴¹ Enzymes involved in serine biosynthesis have been linked to insulin signaling and sensitivity in animal studies, while a lack of serine in cancer cells results in altered mitochondrial metabolism akin to metabolic disturbances resulting in insulin resistance.⁴¹ Further research into the roles of ornithine, proline and serine in diabetes development is required to clarify these inconsistencies and whether these species contribute to or merely indicate higher diabetes risk.

In addition to amino acids, we observed three acylcarnitines, C4, C8-DC, and C16-OH, to be associated with diabetes risk. Acylcarnitines are primarily produced from

mitochondrial fatty acid β -oxidation, and their accumulation may indicate incomplete fatty acid oxidation and downstream metabolic disturbances, including depletion of tricarboxylic acid cycle intermediates and activation of pathways that interfere with insulin action.^{42–44} Short-chain species, such as C4, are intermediate products of β -oxidation, and their accumulation in participants with type 2 diabetes may indicate generalized dysfunction at the interface of fatty acid oxidation and the electron transport chain.⁴ Dicarboxylic species, including C8-DC, are produced when long-chain fatty acids undergo ω -oxidation, a compensatory pathway activated when β -oxidation is disturbed. Reduced concentrations of these species in those with a high risk of diabetes could indicate a disturbance of β -oxidation if the ω -oxidation rescue pathway was also impaired. This would lead to accumulation of fatty acid fuels in the mitochondria and contribute to insulin resistance via the mismatching of fuel and ATP demand.⁴⁵ While we did not observe significant associations between medium-chain species and diabetes risk following multiple testing correction, it has been suggested that the accumulation of medium-chain species results in activation of the proinflammatory NF κ B pathway, which in turn promotes insulin resistance.⁴⁴ The accumulation of long-chain species, such as C16-OH, is similarly thought to be reflective of impaired tricarboxylic acid cycle activity, as they are the initial products of β -oxidation.⁴⁶

The link between acylcarnitines and diabetes is controversial, and there is lack of consensus over whether elevated or depressed levels of specific short-chain, medium-chain, and long-chain species are associated with diabetes risk.^{4 7 42–46} In a Chinese cohort, fasting serum concentrations of C4 were higher in diabetes cases than in non-cases, but the investigators did not find an association with C8-DC or C16-OH.⁷ The authors also described a panel of long-chain acylcarnitines that were significantly associated with diabetes risk and increased the AUC of a predictive diabetes risk model, although C16-OH was not part of the panel. In a US study, fasting concentrations of both C4 and C16-OH were higher in participants with type 2 diabetes compared with lean participants without diabetes.⁴ A German study, however, reported higher concentrations of C16-OH but not C4 in participants with diabetes compared with those with normal glucose tolerance.⁴² A Mexican study reported elevated concentrations of C4 in obese participants without diabetes compared with their counterparts with diabetes,⁴³ while a US study reported no difference in C4 levels between participants of these categories.⁴⁴ In our study, the association between C4 and diabetes risk was not attenuated after additional adjusting for waist circumference, which suggests the mechanism may not be mediated by body fatness. To our knowledge, this is the first study to report an inverse association between serum C8-DC levels and diabetes risk, although an animal study reported higher concentrations of C8-DC in insulin-resistant mice.⁴⁷ Further research is required

to clarify the role of acylcarnitines in the development of diabetes in humans.

Strengths of our study included the prospective design and the Asian study population, a population that is more susceptible to diabetes than populations of European ancestry.³ Our study also had several potential limitations. First, we had substantial non-response during follow-up. This is a common issue in large cohort studies, and we addressed it by using a nationwide clinical registry to ascertain incident diabetes in addition to reported diagnosis and fasting glucose and HbA_{1c} measurements during follow-up. However, there remains some potential for cases to have gone undetected, for instance if participants were diagnosed at a private clinic. Second, metabolite profiles were measured only once during follow-up, resulting in potentially inaccurate measurements of long-term biomarker levels and potential attenuation of observed associations. While the targeted metabolomic approach facilitated identification of potential biomarkers, the panel of metabolites was not exhaustive and concentrations of other clinically important species such as lysine and tryptophan were not recorded. Measuring certain metabolites, including glutamate and glutamine, in aggregate may also have weakened our findings, as the two species play separate biological roles and have displayed markedly differential associations with diabetes risk when measured separately.^{31 32} Additionally, while we based our multivariable analyses on an established diabetes risk model, there remains a potential for residual confounding due to risk factors not included in the ARIC model. Finally, our findings apply to a multi-ethnic Asian population and may not necessarily generalize to other populations or ethnic groups.

Our results provide further evidence of the role of specific acylcarnitines, amino acids, and amino acid ratios in the development of type 2 diabetes in Asian populations. A predictive model containing a panel of acylcarnitines and amino acids improved classification of both diabetes cases and non-cases as compared with a model containing solely the established risk factors included in the ARIC model. The increasing availability and affordability of profiling technologies mean they could feasibly be applied in the clinical setting. However, it remains unclear whether measurement of novel metabolites leads to sufficient improvement in the identification of high-risk groups to warrant use in clinical practice. Further research is warranted to establish whether specific acylcarnitines and amino acids play a causal role in the etiology of diabetes and could be targets for preventive interventions.

Author affiliations

¹Saw Swee Hock School of Public Health, National University of Singapore, Singapore

²Department of Medicine, Yong Loo Lin School of Medicine, National University Health System, Singapore

³Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore

⁴Programme in Cardiovascular and Metabolic Disorders, Duke-NUS Medical School, Singapore

⁵Yong Loo Lin School of Medicine, National University of Singapore, Singapore

⁶Department of Nutrition, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, USA

Acknowledgements The authors thank the members of the SP2 cohort for their cooperation and participation. They also thank the SP2 data collection and management team, Hai Ning Wee and Kee Voon Chua of Duke-NUS, and Milly Ng and Yueheng Hong of SSHSPH.

Contributors SHG contributed to study design and conduct, data analysis, and wrote the manuscript. XS contributed to study design and data analysis and reviewed the manuscript. E-ST, CMK, J-PK, JC, JLL, and RMvD contributed to study design and conduct, data collection, and reviewed the manuscript. All authors approved the final version of the manuscript.

Funding This work was supported by grants from the Biomedical Research Council (grant 03/1/27/18/216), National Medical Research Council (grants 0838/2004 and 1111/2007), and National Research Foundation (through the Biomedical Research Council, grants 05/1/21/19/425 and 11/1/21/19/678) of the Republic of Singapore.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval All participants provided written informed consent before taking part in SP2 and the follow-up examinations. SP2 and the follow-up examinations were approved by the National University of Singapore IRB (reference no. 12-282).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data collected and analyzed in this study are available upon request from the National University of Singapore. Inquiries can be directed to SSHSPHDataRequest@nus.edu.sg.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Samuel H Gunther <http://orcid.org/0000-0002-5914-0345>

Xueling Sim <http://orcid.org/0000-0002-1233-7642>

REFERENCES

- 1 Guariguata L, Whiting DR, Hambleton I, *et al*. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract* 2014;103:137–49.
- 2 Nakagami T, Qiao Q, Carstensen B, *et al*. Age, body mass index and type 2 diabetes-associations modified by ethnicity. *Diabetologia* 2003;46:1063–70.
- 3 Ramachandran A, Snehalatha C, Shetty AS, *et al*. Trends in prevalence of diabetes in Asian countries. *World J Diabetes* 2012;3:110–7.
- 4 Mihalik SJ, Goodpaster BH, Kelley DE, *et al*. Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipototoxicity. *Obesity* 2010;18:1695–700.
- 5 Newgard CB, An J, Bain JR, *et al*. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab* 2009;9:311–26.
- 6 Qiu G, Zheng Y, Wang H, *et al*. Plasma metabolomics identified novel metabolites associated with risk of type 2 diabetes in two prospective cohorts of Chinese adults. *Int J Epidemiol* 2016;45:1507–16.
- 7 Sun L, Liang L, Gao X, *et al*. Early prediction of developing type 2 diabetes by plasma acylcarnitines: a population-based study. *Diabetes Care* 2016;39:1563–70.
- 8 Tai ES, Tan MLS, Stevens RD, *et al*. Insulin resistance is associated with a metabolic profile of altered protein metabolism in Chinese and Asian-Indian men. *Diabetologia* 2010;53:757–67.
- 9 Schmidt MI, Duncan BB, Bang H, *et al*. Identifying individuals at high risk for diabetes: the Atherosclerosis risk in Communities study. *Diabetes Care* 2005;28:2013–8.

- 10 Nang EEK, Khoo CM, Tai ES, *et al.* Is there a clear threshold for fasting plasma glucose that differentiates between those with and without neuropathy and chronic kidney disease?: the Singapore prospective study program. *Am J Epidemiol* 2009;169:1454–62.
- 11 Hughes K, Yeo PP, Lun KC, *et al.* Cardiovascular diseases in Chinese, Malays, and Indians in Singapore. II. Differences in risk factor levels. *J Epidemiol Community Health* 1990;44:29–35.
- 12 Tan CE, Emmanuel SC, Tan BY, *et al.* Prevalence of diabetes and ethnic differences in cardiovascular risk factors. The 1992 Singapore National health survey. *Diabetes Care* 1999;22:241–7.
- 13 Hughes K, Aw TC, Kuperan P, *et al.* Central obesity, insulin resistance, syndrome X, lipoprotein(a), and cardiovascular risk in Indians, Malays, and Chinese in Singapore. *J Epidemiol Community Health* 1997;51:394–9.
- 14 Cutter J, Tan BY, Chew SK. Levels of cardiovascular disease risk factors in Singapore following a national intervention programme. *Bull World Health Organ* 2001;79:908–15.
- 15 American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2014;37(Suppl 1):S81–90.
- 16 Libert DM, Nowacki AS, Natowicz MR. Metabolomic analysis of obesity, metabolic syndrome, and type 2 diabetes: amino acid and acylcarnitine levels change along a spectrum of metabolic wellness. *PeerJ* 2018;6:e5410.
- 17 Molnos S, Wahl S, Haid M, *et al.* Metabolite ratios as potential biomarkers for type 2 diabetes: a DIRECT study. *Diabetologia* 2018;61:117–29.
- 18 Petersen A-K, Krumsiek J, Wägele B, *et al.* On the hypothesis-free testing of metabolite ratios in genome-wide and metabolome-wide association studies. *BMC Bioinformatics* 2012;13:120.
- 19 Chin CW-L, Chia EHS, Ma S, *et al.* The ARIC predictive model reliably predicted risk of type II diabetes in Asian populations. *BMC Med Res Methodol* 2012;12:48.
- 20 Wang TJ, Larson MG, Vasani RS, *et al.* Metabolite profiles and the risk of developing diabetes. *Nat Med* 2011;17:448–53.
- 21 Palmer ND, Stevens RD, Antinozzi PA, *et al.* Metabolomic profile associated with insulin resistance and conversion to diabetes in the insulin resistance atherosclerosis study. *J Clin Endocrinol Metab* 2015;100:E463–8.
- 22 Chen T, Ni Y, Ma X, *et al.* Branched-chain and aromatic amino acid profiles and diabetes risk in Chinese populations. *Sci Rep* 2016;6:20594.
- 23 Lotta LA, Scott RA, Sharp SJ, *et al.* Genetic predisposition to an impaired metabolism of the branched-chain amino acids and risk of type 2 diabetes: a Mendelian randomisation analysis. *PLoS Med* 2016;13:e1002179.
- 24 Nie C, He T, Zhang W, *et al.* Branched chain amino acids: beyond nutrition metabolism. *Int J Mol Sci* 2018;19:954.
- 25 Chen T, Zheng X, Ma X, *et al.* Tryptophan predicts the risk for future type 2 diabetes. *PLoS One* 2016;11:e0162192.
- 26 Ruiz-Canela M, Guasch-Ferré M, Toledo E, *et al.* Plasma branched chain/aromatic amino acids, enriched Mediterranean diet and risk of type 2 diabetes: case-cohort study within the PREDIMED trial. *Diabetologia* 2018;61:1560–71.
- 27 Magnusson M, Lewis GD, Ericson U, *et al.* A diabetes-predictive amino acid score and future cardiovascular disease. *Eur Heart J* 2013;34:1982–9.
- 28 Krebs M, Krssak M, Bernroider E, *et al.* Mechanism of amino acid-induced skeletal muscle insulin resistance in humans. *Diabetes* 2002;51:599–605.
- 29 Müller WA, Faloona GR, Unger RH. The effect of alanine on glucagon secretion. *J Clin Invest* 1971;50:2215–8.
- 30 Sattar N, Scherbakova O, Ford I, *et al.* Elevated alanine aminotransferase predicts new-onset type 2 diabetes independently of classical risk factors, metabolic syndrome, and C-reactive protein in the West of Scotland coronary prevention study. *Diabetes* 2004;53:2855–60.
- 31 Cheng S, Rhee EP, Larson MG, *et al.* Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation* 2012;125:2222–31.
- 32 Cabrera O, Jacques-Silva MC, Speier S, *et al.* Glutamate is a positive autocrine signal for glucagon release. *Cell Metab* 2008;7:545–54.
- 33 Cynober LA. The use of alpha-ketoglutarate salts in clinical nutrition and metabolic care. *Curr Opin Clin Nutr Metab Care* 1999;2:33–7.
- 34 Alves A, Bassot A, Bulteau A-L, *et al.* Glycine metabolism and its alterations in obesity and metabolic diseases. *Nutrients* 2019;11:E1356.
- 35 Floegel A, Stefan N, Yu Z, *et al.* Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* 2013;62:639–48.
- 36 Gall WE, Beebe K, Lawton KA, *et al.* α -hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS One* 2012;5:e10883.
- 37 Yamakado M, Nagao K, Imaizumi A, *et al.* Plasma free amino acid profiles predict four-year risk of developing diabetes, metabolic syndrome, dyslipidemia, and hypertension in Japanese population. *Sci Rep* 2015;5:11918.
- 38 Merino J, Leong A, Liu C-T, *et al.* Metabolomics insights into early type 2 diabetes pathogenesis and detection in individuals with normal fasting glucose. *Diabetologia* 2018;61:1315–24.
- 39 Moon J, Kim OY, Jo G, *et al.* Alterations in circulating amino acid metabolite ratio associated with arginase activity are potential indicators of metabolic syndrome: the Korean genome and epidemiology study. *Nutrients* 2017;9:E740.
- 40 Cao Y-F, Li J, Zhang Z, *et al.* Plasma levels of amino acids related to urea cycle and risk of type 2 diabetes mellitus in Chinese adults. *Front Endocrinol* 2019;10:50.
- 41 Holm LJ, Buschard K. L-Serine: a neglected amino acid with a potential therapeutic role in diabetes. *APMIS* 2019;127:655–9.
- 42 Mai M, Tönjes A, Kovacs P, *et al.* Serum levels of acylcarnitines are altered in prediabetic conditions. *PLoS One* 2013;8:e82459.
- 43 Villarreal-Pérez JZ, Villarreal-Martínez JZ, Lavallo-González FJ, *et al.* Plasma and urine metabolic profiles are reflective of altered beta-oxidation in non-diabetic obese subjects and patients with type 2 diabetes mellitus. *Diabetol Metab Syndr* 2014;6:129.
- 44 Adams SH, Hoppel CL, Lok KH, *et al.* Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid beta-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. *J Nutr* 2009;139:1073–81.
- 45 Liu J-J, Ghosh S, Kovalik J-P, *et al.* Profiling of Plasma Metabolites Suggests Altered Mitochondrial Fuel Usage and Remodeling of Sphingolipid Metabolism in Individuals With Type 2 Diabetes and Kidney Disease. *Kidney Int Rep* 2017;2:470–80.
- 46 Koves TR, Ussher JR, Noland RC, *et al.* Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance. *Cell Metab* 2008;7:45–56.
- 47 Bloom K, Mohsen A-W, Karunanidhi A, *et al.* Investigating the link of ACAD10 deficiency to type 2 diabetes mellitus. *J Inher Metab Dis* 2018;41:49–57.