International Journal of Epidemiology, 2016, 1507–1516
doi: 10.1093/ije/dyw221
Advance Access Publication Date: 30 September 2016
Original article

Metabolomics

Plasma metabolomics identified novel metabolites associated with risk of type 2 diabetes in two prospective cohorts of Chinese adults

Gaokun Qiu,[1] Yan Zheng,[2] Hao Wang,[1] Jie Sun,[3] Hongxia Ma,[3] Yang Xiao,[1] Yizhun Li,[1] Yu Yuan,[1] Handong Yang,[4] Xiulou Li,[4] Xinwen Min,[4] Ce Zhang,[4] Chengwei Xu,[4] Yue Jiang,[3] Xiaomin Zhang,[1] Meian He,[1] Ming Yang,[1] Zhibin Hu,[3] Huiru Tang,[5,6] Hongbing Shen,[3] Frank B Hu,[2] An Pan[1]* and Tangchun Wu[1]*

[1]Department of Occupational and Environmental Health and Department of Epidemiology and Biostatistics, Ministry of Education and State Key Laboratory of Environmental Health, Huazhong University of Science and Technology, Wuhan, China, [2]Department of Nutrition and Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA, [3]Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing, China, [4]Department of Cardiovascular Disease, Dongfeng Central Hospital, Hubei University of Medicine, Shiyan, China, [5]State Key Laboratory of Genetic Engineering, Fudan University, Shanghai, China and [6]CAS Key Laboratory of Magnetic Resonance in Biological Systems, University of Chinese Academy of Sciences, Wuhan, China

*Corresponding author. School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, 13 Hongkong Rd, Wuhan 430030, Hubei, China. E-mail: wut@mails.tjmu.edu.cn, or panan@hust.edu.cn

Accepted 12 July 2016

Abstract

Background: Metabolomics studies in Caucasians have identified a number of novel metabolites in association with the risk of type 2 diabetes (T2D). However, few prospective metabolomic studies are available in Chinese populations. In the present study, we sought to identify novel metabolites consistently associated with incident T2D in two independent cohorts of Chinese adults.

Methods: We performed targeted metabolomics (52 metabolites) of fasting plasma samples by liquid chromatography-mass spectrometry in two prospective case-control studies nested within the Dongfeng-Tongji (DFTJ) cohort and Jiangsu Non-communicable Disease (JSNCD) cohort. After following for 4.61 ± 0.15 and 7.57 ± 1.13 years, respectively, 1039 and 520 eligible participants developed incident T2D in these two cohorts, and controls were 1:1 matched with cases by age ($\pm$ 5 years) and sex. Multivariate conditional logistic regression models were constructed to identify metabolites associated with future T2D risk in both cohorts.

Results: We identified four metabolites consistently associated with an increased risk of developing T2D in the two cohorts, including alanine, phenylalanine, tyrosine and palmitoylcarnitine. In the meta-analysis of two cohorts, the odds ratios (95% confidence intervals, CIs) comparing extreme quartiles were 1.79 (1.32–2.42) for alanine, 1.91 (1.41–2.60) for phenylalanine, 1.85 (1.37–2.48) for tyrosine and 1.63 (1.21–2.20) for palmitoylcarnitine (all $P_{trend} \leq 0.01$).
Conclusions: We confirmed the association of alanine, phenylalanine and tyrosine with future T2D risk and further identified palmitoylcarnitine as a novel metabolic marker of incident T2D in two prospective cohorts of Chinese adults. Our findings might provide new aetiological insight into the development of T2D.

Key words: Type 2 diabetes, targeted metabolomics, nested case-control study, Chinese population

Key Messages

- Few prospective metabolomic studies of T2D risk are available in Chinese populations.
- Alanine, phenylalanine, tyrosine and palmitoylcarnitine were consistently associated with an increased risk of developing T2D in the two cohorts of Chinese adults.
- Palmitoylcarnitine was identified as a novel metabolic marker of incident T2D.

Introduction

Type 2 Diabetes (T2D) is metabolic disorder characterized by impaired insulin sensitivity and increased insulin resistance.[1] The pathogenesis of T2D involves complex genetic and environmental influences among which pronounced ethnic disparities have been reported, with Asians being at higher risks compared with Caucasians.[2,3] In China, the prevalence of T2D has been increasing dramatically over recent decades, from less than 1% in 1980 to 11.6% in 2010,[4] and the pace of the T2D epidemic in China is projected to continue accelerating.[5] The investigation of novel biomarkers of T2D risk may advance the understanding of disease pathophysiology and facilitate targeted preventive care.

Metabolomics is an emerging analytical technology defined as the high-throughput characterization and quantification of molecule metabolites in biological samples.[6,7] The metabolic profile represents the end products of genomic, transcriptomic and proteomic variability as well as environmental stimulations, thereby providing the most integrated profile of biological status and being more relevant to disease phenotypes.[7–9] Several prospective metabolomic studies conducted in Caucasians have identified a number of novel metabolites predictive of T2D risk, including branched-chain amino acids (leucine, isoleucine and valine),[10,11] aromatic amino acids (phenylalanine and tyrosine),[10–12] other amino acids,[11,13,14] acylcarnitines[13] and certain lipids.[12,13,15,16] Likewise, Zhao *et al.*[17] reported novel associations of flavonoids and tetra-peptides with future T2D risk in American Indians, and Tillin *et al.*[18] confirmed most of the aforementioned associations of amino acids in South Asians, particularly that of tyrosine. Moreover, Walford *et al.* found in a multi-ethnic population that betaine was predictive of reduced risk of incident T2D, and an increase in circulating betaine during preventive lifestyle interventions was also associated with lower T2D incidence.[19] Only one prospective metabolomic study has been conducted in a Chinese population investigating T2D risk, though no replication was available and only 73 incident cases were included.[20]

In the present study, we performed targeted metabolomics in two large nested case-control studies within the Dongfeng-Tongji (DFTJ) cohort and Jiangsu Non-communicable Disease (JSNCD) cohort. We focused on amino acids and (acyl) carnitines in particular, as they were promising candidate biomarkers of T2D risk as revealed by previous studies[10–14] and could be easily accommodated within a single analytical run.[21] We aimed to identify novel metabolites consistently associated with T2D risk in both cohorts, and to examine the predictive utility of identified metabolic markers beyond established diabetes risk factors.

Research Design and Methods

Study population

A detailed description of the baseline profiles of the DFTJ cohort has been published elsewhere.[22] In brief, the DFTJ cohort was launched in 2008 and enrolled retirees of the

Dongfeng Motor Corporation (DMC) who were residents in Hubei Province of central China. A total of 27 009 DMC retirees responded to questionnaires, participated in physical examinations and provided blood samples during 2008–10. Follow-up investigations were conducted during 2013–14, with a follow-up rate of 96.2%. T2D was defined if at least one of the following criteria was met:[23] (i) fasting glucose $\geq$ 7.0 mmol/l; (ii) haemoglobin A1c (HbA1c) $\geq$ 6.5%; (iii) self-reported use of antidiabetic medication. As HbA1c levels were not measured, baseline T2D were diagnosed only with a fasting glucose test and reported use of antidiabetic medications. After 4.61 $\pm$ 0.15 years of follow-up, 1515 participants developed incident T2D. After further exclusion of participants with baseline cardiovascular disease or cancer and those without sufficient blood samples, a total of 1039 incident cases remained. Controls were randomly selected from participants who were free of diabetes, cardiovascular disease and cancer at baseline and were also diabetes-free in the follow-up examinations, and were 1:1 matched for age ($\pm$ 5 years) and sex to incident cases.

The JSNCD cohort was established in 2004–05, and participants were recruited with multi-stage random cluster sampling from residents of Wujin district in Jiangsu Province of east China. All participants were at least 35 years of age and had lived in their current residence for no less than 5 years. In total, 17 723 participants completed the baseline survey. Follow-up investigations were performed during 2008–09 and 2012–13, and the follow-up rates were 91.8% and 92.1%, respectively. As HbA1c levels were not measured in this cohort, incident T2D cases were diagnosed only with a fasting glucose test and reported use of antidiabetic medications. After 7.57 $\pm$ 1.13 years of follow-up, 779 participants developed incident T2D, and a total of 520 incident cases remained after excluding participants with baseline cardiovascular disease or cancer and those without sufficient blood samples. Controls were selected and matched to incident cases with the same criteria as aforementioned.

In both cohorts, participants were interviewed by trained investigators using semi-structured questionnaires to collect information on socio-demographic factors, lifestyle habits, health status and medical history. Standing height, body weight and waist circumference were measured by trained personnel with participants being in light indoor clothing without shoes. Body mass index (BMI) was calculated as weight in kg divided by the square of height in metres. Participants who reported regular exercise for at least 30 min on no less than 5 days per week were defined as physically active. Blood pressures were measured on the left upper arm with the participants in a seated position after a brief rest, and hypertension was defined if the participant had a blood pressure $\geq$ 140/90 mmHg, or reported use of antihypertensive medication. All blood samples were drawn after an overnight fast for at least 8 h, and stored at -80°C until analysis. Serum lipids and glucose levels were measured with Architect Ci8200 analyser (Abbott Laboratories, Abbott Park, United States) in the DFTJ cohort and with OLYMPUS AU640 analyser (Olympus Diagnostic Systems, Southall, Middlesex, UK) in the JSNCD cohort. All participants gave written informed consent. The study protocol was approved by the Ethics and Human Subject Committee of Tongji Medical College and Nanjing Medical University.

Metabolic profiling by high-performance liquid chromatography–mass spectrometry

Metabolic profiling by high-performance liquid chromatography–mass spectrometry (HPLC-MS) was performed based on the methods described by Wang *et al.*[10] and Kalim *et al.*[21] with modifications. In brief, reference standards (all from Sigma-Aldrich, St Louis, MO) were used to determine chromatographic retention times, multiple reaction–monitoring (MRM) transitions, fragmentors and collision energies for all concerned metabolites. The LC-MS/MS system consisted of an Agilent 1200 Series HPLC equipped with a 6400 triple quadrupole mass spectrometer (Agilent Technologies). Plasma samples (10 µl) were extracted using 90 µl of 74.9:24.9:0.2 (v/v/v) acetonitrile/methanol/formic acid containing stable isotope-labelled internal standards (2.5 µmol/l valine-d8, 5 µmol/l phenylalanine-d8 and 1 µmol/l carnitine-d9; all from Cambridge Isotope Laboratories, Andover, MA), briefly vortexed and then centrifuged (10 min, 12000*g*, 4°C). The supernatants (10 µl) were injected onto a Cortecs HILIC column (100 x 2.1 mm, Waters Corp., Milford, MA) which was eluted isocratically at a flow rate of 0.5 ml/min with 5% mobile phase A (10 mmol/l ammonium formate and 0.1% formic acid in water) for 0.5 min followed by a linear gradient to 40% mobile phase B (acetonitrile with 0.1% formic acid) over 5 min. MS analyses were carried out using electrospray ionization and dynamic MRM scans in the positive ion mode which enables the parallel analysis of 52 metabolites. Internal standard peak areas were monitored for quality control, and individual samples with peak areas deviating from the mean by more than two standard deviations (SD) during one day's analysis were re-analysed. Raw peak areas of each metabolite were normalized relative to the pooled plasma reference samples that were analysed in the sample queue periodically after each set of

20 samples, to account for the instrument drift occurring over time.[10,24] All samples were analysed in randomized and blinded order with the same instrument in the central laboratory at the School of Public Health, Tongji Medical College, Huazhong University of Science and Technology. The ion spray voltage was 4.5 kV and the source temperature was 350°C.

Statistical analysis

Before analysis, metabolites data were natural-log transformed and then standardized to z scores in each cohort. Pearson partial correlation coefficients were calculated across metabolite pairs in controls, adjusting for age and sex. Participants were then categorized into four categories according to quartile cutoffs of each metabolite in controls. We constructed multivariate conditional logistic models for each metabolite (as categorical or continuous variables) adjusting for age, BMI, smoking status, drinking status, education level, physical activity, systolic blood pressure, serum high-cholesterol lipoprotein (HDL) cholesterol and triglycerides, fasting glucose, family history of T2D and metabolomics batch. Metabolites with the same direction of association in both cohorts were selected, and false discovery rates (FDRs) were calculated considering all metabolites showing the same direction of association in both cohorts, to account for multiple testing. Metabolites achieving FDR < 0.1 in both cohorts were identified, and their associations with T2D risk were pooled with inverse-variance weighted meta-analysis. A metabolomic score was created by summing the quartile ranks of the identified metabolites to assess their composite effects.[25] We then conducted a sensitivity analysis with further adjustments of dietary variables, including intakes of fruit and vegetable, meat, fish and seafoods, dairy products and soybeans, and a meta-analysis of other metabolites with the same direction of associations in both cohorts regardless of the FDR values.

We examined the predictive ability of the identified metabolites by comparing diabetes prediction models using traditional risk factors with and without all the identified metabolic markers in the present study. We calculated the area under the receiver operating characteristic curve (AUC), the net reclassification improvement (NRI) and the integrated discrimination improvement (IDI) to assess the incremental value of these metabolic markers for risk prediction beyond traditional risk factors.[26] Considering the fact that no criteria were established for interpretation of the magnitude of the IDI, we also calculated the relative IDI.[27] All statistical analyses were performed with SAS version 9.3 (SAS Institute Inc.).

Results

Table 1 presents the baseline demographic and clinical characteristics of the study participants from the two cohorts according to case-control status. In both cohorts, incident T2D cases had higher baseline levels of BMI, waist circumference, blood pressures, triglycerides and fasting glucose, whereas higher levels of HDL cholesterol were observed in controls (all P-values < 0.01). Cases were also more likely to report a family history of diabetes in both cohorts (both P-values < 0.01).

A total of 52 metabolites were detected in our metabolomics platform, including 26 amino acids, 12 carnitine and acylcarnitines, two cholines, two amines, two purine derivatives, two B vitamins, two indole derivatives and four other metabolites. MRM transitions and retention times of these metabolites are provided in Supplementary Table 1, available as Supplementary data at *IJE* online. We then assessed age- and sex-adjusted pairwise Pearson correlations between baseline levels of metabolites (Figure 1). Mean correlations within amino acids and (acyl) carnitines were modest in the DFTJ and JSNCD cohorts [r = 0.15 and 0.16 for amino acids, respectively; r = 0.27 and 0.31 for (acyl) carnitines, respectively].

We observed 20 metabolites demonstrating associations in the same direction with T2D risk in both cohorts (Figure 2; Supplementary Tables 2 and 3, available as Supplementary data at *IJE* online). Among these metabolites, alanine, phenylalanine, tyrosine and palmitoylcarnitine were identified with FDRs < 0.1 in both cohorts. The associations of these four metabolites with diabetes risk remained unchanged after further adjustments of dietary variables (data not shown). In fixed effect pooled analysis of these four metabolites (Figure 3; Supplementary Table 4, available as Supplementary data at *IJE* online), the odds ratios (95% confidence intervals, CIs) comparing extreme quartiles were 1.79 (1.32–2.42) for alanine, 1.91 (1.41–2.60) for phenylalanine, 1.85 (1.37–2.48) for tyrosine and 1.63 (1.21–2.20) for palmitoylcarnitine (all $P_{trend} \leq 0.01$), and participants in the highest quartile of the metabolomic score had 2.44-fold odds of developing T2D (95% CI: 1.79–3.32; $P_{trend} = 2.42E-09$) compared with those in the lowest quartile.

We also performed an exploratory analysis of pooling other metabolites with the same direction of association in both cohorts (Supplementary Table 4). An additional set of 12 metabolites achieved FDR < 0.1, including six amino acids (betaine, glutamate, leucine/isoleucine, ornithine, proline and valine) and six other metabolites (acetylcholine, α-glycerophosphocholine, creatinine, indoleacetate, inosine and trimethylamine-N-oxide).

We then assessed predictive performance of the four identified metabolites in both cohorts (Figure 4; Supplementary Table 5, available as Supplementary data

Table 1. Baseline characteristics of the study populations

Variables	The DFTJ cohort			The JSNCD cohort		
	Cases	Matched controls	*P*	Cases	Matched controls	*P*
N	1039	1039		520	520	
Age, years	62.82 ± 7.23	62.93 ± 7.32	0.736	53.82 ± 10.25	53.74 ± 10.18	0.896
Male sex, %	44.7	44.7	1	34.8	34.8	1
BMI, kg/m^2	25.73 ± 3.34	23.64 ± 3.07	< 0.001	25.53 ± 3.42	23.70 ± 3.22	< 0.001
Waist circumference, cm	85.88 ± 9.84	80.38 ± 8.39	< 0.001	86.09 ± 9.53	80.57 ± 9.13	< 0.001
Current smoker, %	19.0	19.6	0.850	22.7	23.3	0.888
Current drinker, %	22.5	24.0	0.019	19.04	20.8	0.662
Physical activity (yes, %)	70.6	71.3	0.386	59.6	61.9	0.242
Systolic blood pressure, mmHg	131.07 ± 18.29	124.10 ± 17.23	< 0.001	132.87 ± 21.15	126.59 ± 20.14	< 0.001
Diastolic blood pressure, mmHg	78.94 ± 11.10	74.98 ± 10.06	< 0.001	83.43 ± 10.98	80.59 ± 10.24	< 0.001
Hypertension (%)	54.4	30.4	< 0.001	52.5	36.7	< 0.001
HDL cholesterol, mmol/L	1.42 ± 0.47	1.49 ± 0.43	< 0.001	1.45 ± 0.46	1.55 ± 0.63	0.003
LDL cholesterol, mmol/L	3.04 ± 0.78	2.95 ± 0.73	0.008	2.36 ± 0.97	2.36 ± 0.95	0.996
Total cholesterol, mmol/L	5.24 ± 0.93	5.09 ± 0.89	< 0.001	4.59 ± 1.05	4·48 ± 1·04	0.094
Total triglycerides, mmol/L	1.61 ± 0.98	1.23 ± 0.72	< 0.001	1.89 ± 1.28	1.49 ± 1·20	< 0.001
Fasting glucose, mmol/L	5.99 ± 0.60	5.48 ± 0.53	< 0·001	5.68 ± 0.80	5.03 ± 0.61	< 0.001
Family history (%)	5.0	2.3	0·001	11.0	5.8	0.002

Continuous variables were presented as mean ± SD and compared with one-way analysis of variance (ANOVA); categorical variables were compared with chi-square tests.

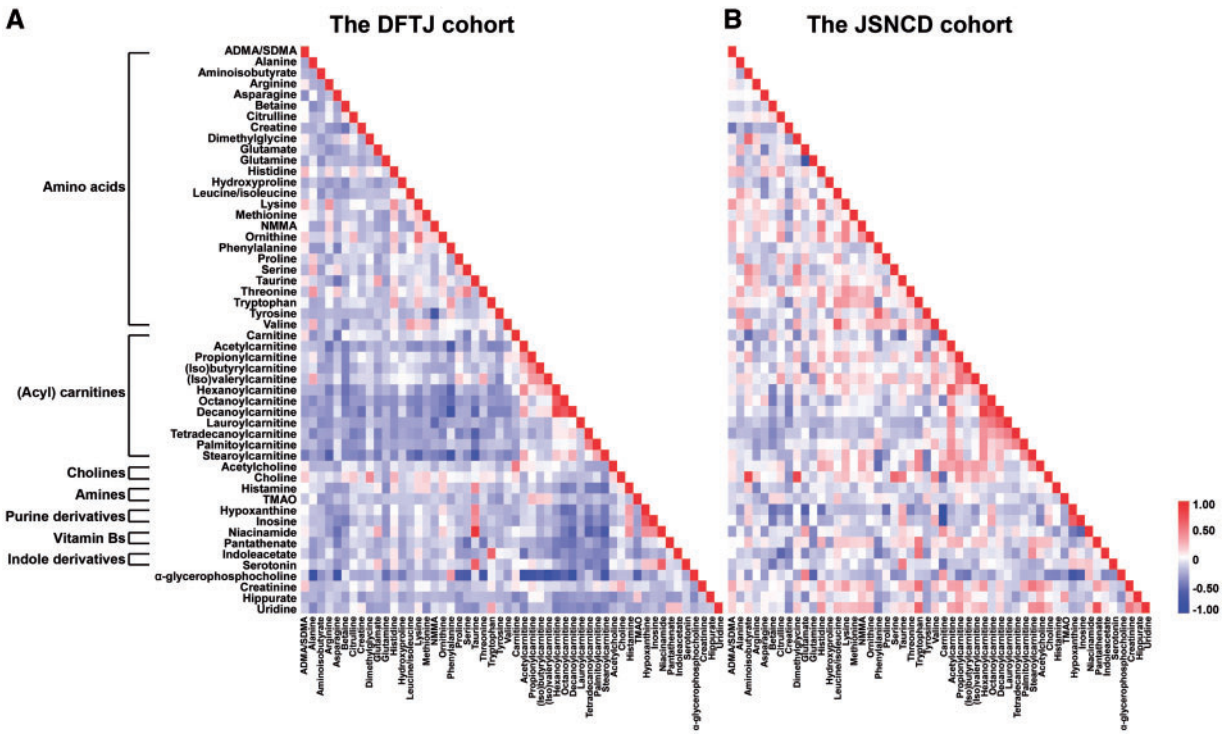

Figure 1. Correlation matrix of plasma metabolite levels in controls of the DFTJ cohort (A) and the JSNCD cohort (B). Age- and sex-adjusted Pearson correlation coefficients are presented.

at *IJE* online). In the JSNCD cohort, the addition of these four metabolites increased the c-statistic from 0.673 (reference model without fasting glucose) to 0.718 ($P = 2.88E-4$), and from 0.777 (reference model with fasting glucose) to 0.794 ($P = 0.007$). The relative IDI and NRI were estimated at 54.0% and 11.6%, respectively, in the latter scenario (*P*-values both < 0.0001). However, risk prediction improvement in the DFTJ cohort was modest. The

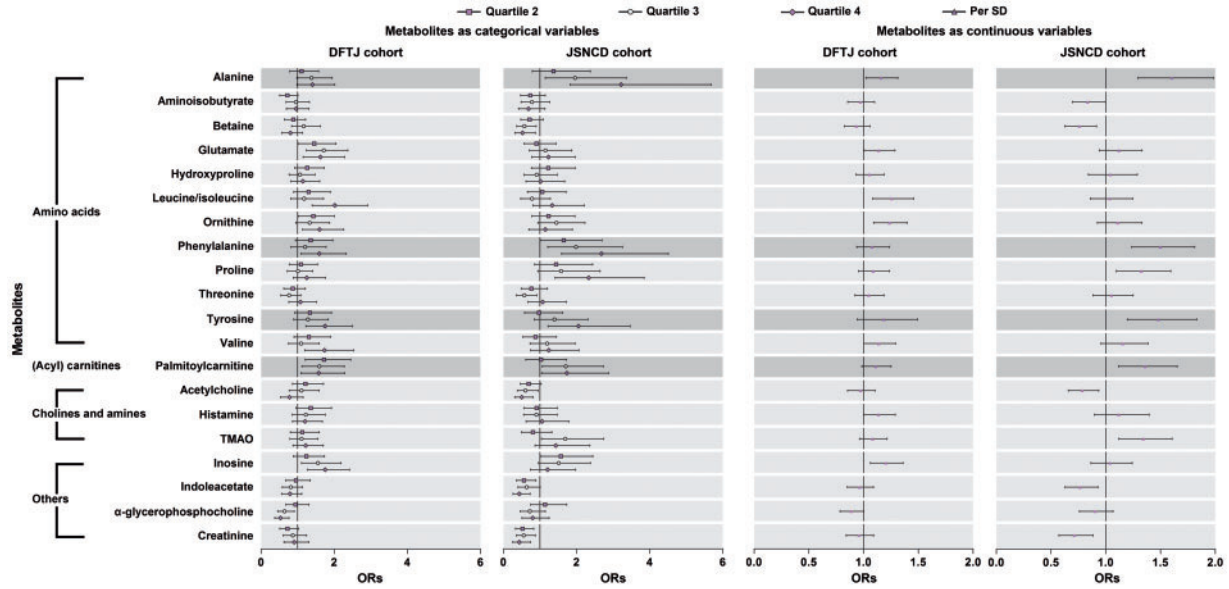

Figure 2. Metabolites demonstrating associations in the same direction with the risk of future diabetes in both cohorts. ORs were obtained with conditional logistic models adjusting for age, BMI, smoking and drinking status, education level, physical activity, systolic blood pressure, serum HDL cholesterol and triglycerides, fasting glucose, family history of diabetes and metabolomics batch. Metabolites shaded in darker grey were identified as being associated with future T2D risk: alanine (FDR = 0.077 and 4.39E-04); phenylalanine (FDR = 0.077 and 0.002); tyrosine (FDR = 0.009 and 0.008); palmitoylcarnitine (FDR = 0.077 and 0.020).

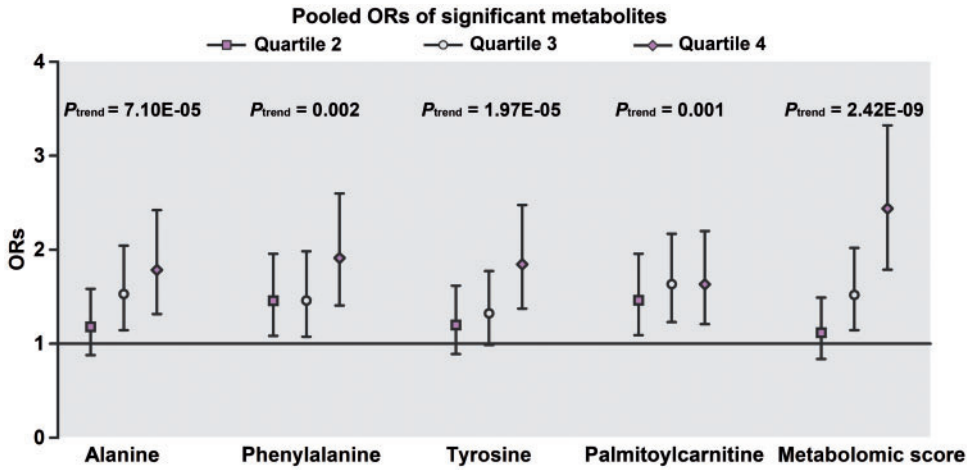

Figure 3. Pooled ORs of the four identified metabolites and the metabolomics score. ORs were pooled with a fixed effect meta-analysis.

increase of c-statistic was 0.017 for the reference model without fasting glucose ($P = 0.002$) and 0.004 for the reference model with fasting glucose ($P > 0.05$). Nevertheless, the relative IDI and NRI in comparison with the latter reference model suggested predictive improvement, which were estimated at 12.6% and 5.0%, respectively (P values both < 0.0001).

Discussion

In this prospective investigation in two independent nested case-control studies of targeted metabolomics and T2D risk, we identified four metabolites consistently associated with the risk of developing T2D, including alanine, phenylalanine, tyrosine and palmitoylcarnitine. We also observed a notable composite effect of the combination of these four metabolites. In the assessment of predictive performance, we found that only in one cohort did these metabolites modestly improve risk prediction of future T2D beyond established diabetes risk factors.

Among the three amino acids identified as associated with incident T2D, alanine is non-essential to the human body and its association with diabetes risk has been reported in some studies. As a major hepatic substrate for

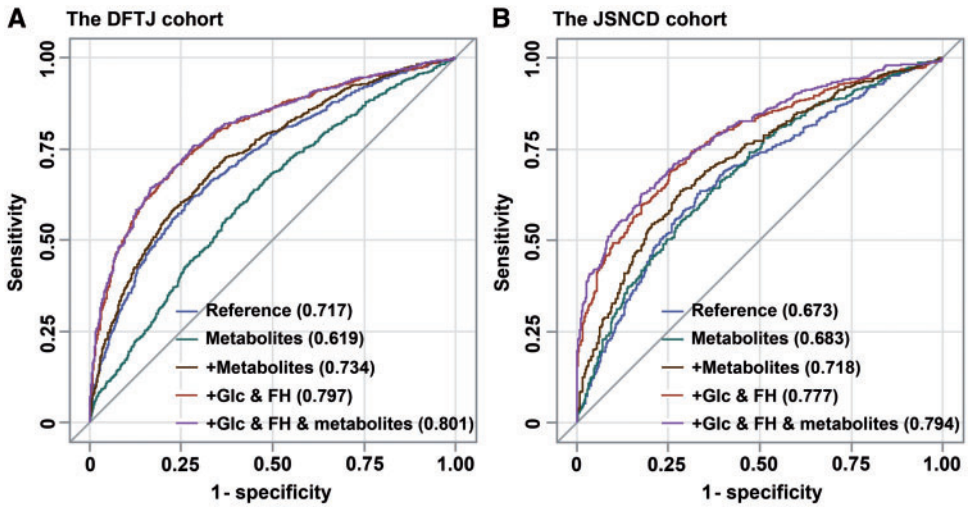

Figure 4. ROC curves for risk of future diabetes in the DFTJ cohort and JSNCD cohort. The reference model included age, gender, BMI, smoking status, drinking status, physical activity, systolic blood pressure, serum HDL cholesterol and triglycerides. Subsequent models include the basic clinical variables plus the identified metabolic predictors as indicated. Glc, fasting glucose; FH, family history.

gluconeogenesis[28] and a stimulator of glucagon secretion,[29] circulating alanine was found to be correlated cross-sectionally[11] and prospectively[30] with impaired insulin sensitivity. Two prospective studies have reported that elevated alanine level was positively associated with incident diabetes in Finnish males[11] and South Asian males living in UK.[18] However, these two studies were modest in sample sizes (151 incident cases and 227 incident cases, respectively) and included only male participants. Moreover, no replication samples were available in these studies and therefore the robustness of their results remained unclear. In our study, we found a relation of alanine with increased risk of developing T2D in two large independent cohorts, which constituted the first report of this association in the Chinese population. Given the fact that some studies have failed to observe an association between alanine and T2D risk,[10,12] our results need further replication, particularly in Chinese populations.

In the present study, we also found positive relations of phenylalanine and tyrosine with increased T2D risk, which were consistent with previous prospective reports in Caucasians and South Asians.[10–12] Phenylalanine is essential to the human body, whereas tyrosine is semi-essential and is synthesized from phenylalanine.[31] Postulated pathways linking this association involves the inhibition of glucose transport/phosphorylation[32] and induction of insulin resistance through phosphorylation of the insulin receptor substrate 1.[11,33] Previous metabolomics studies also found that phenylalanine and tyrosine were associated with the development of insulin resistance,[18,30,34,35] supporting their roles in the pathogenesis of T2D. Notably, a recent study revealed that tyrosine was more strongly predictive of future diabetes in South Asians compared with their European counterparts,[18] suggesting potential ethnic differences in metabolic disturbances related to diabetes risk. However, it remains to be elucidated whether tyrosine and phenylalanine could contribute to the mechanisms underlying the ethnic difference of T2D risk between Asians and Europeans.[2,3] Our finding confirmed the association of tyrosine and its precursor, phenylalanine, with future T2D risk for the first time in a Chinese population, and further highlighted the involvement of aromatic amino acid metabolism in the pathogenesis of T2D, especially in Asians.

Besides amino acids, a long-chain acylcarnitine, palmitoylcarnitine was also found to be associated with increased T2D risk in our study. Acylcarnitines are generated through the esterification of fatty acids, a process required for the transportation of long-chain fatty acids (>14 carbon atoms) across the mitochondrial membrane for β oxidation.[36] The 16-carbon fatty acid, palmitic acid, is the most abundant saturated fatty acid in human serum.[37] A number of studies have reported a positive association between palmitic acid levels and future T2D risk;[38–40] however, whether palmitoylcarnitine might also play a role in T2D pathogenesis remains unknown. A recent study reported that palmitoylcarnitine treatment could directly reduce insulin sensitivity in human myotubes,[41] supporting the aetiological involvement of palmitoylcarnitine in insulin resistance. Some cross-sectional studies also revealed that circulating palmitoylcarnitine was positively correlated with obesity measures[42] and was elevated in pre-diabetic and diabetic patients.[43,44] Our study provided the first prospective evidence to date relating palmitoylcarnitine to future T2D risk. This novel association needs to be replicated in other studies, and the underlying mechanisms warrant further investigation.

A number of prospective metabolomics studies have identified branched-chain amino acids (leucine, isoleucine and valine) as being associated with the development of insulin resistance and T2D.[10,11,18] Despite the fact that the associations of these amino acids failed to reach the criterion of FDR < 0.1 in both cohorts, the directions of associations were consistent in these two cohorts, and the meta-analysis suggested that they could be potentially related to incident T2D. Likewise, we found several other metabolites possibly related to T2D risk as suggested by the meta-analysis, some of which have displayed associations with insulin resistance, prevalent or incident T2D, such as ornithine,[45] proline[45–47] and betaine.[19]

We observed that the associations of the four identified metabolites with T2D risk and the predictive improvement were more pronounced in the JSNCD cohort, whose mean age was nearly 10 years younger than that of the DFTJ cohort. We therefore suspected that these metabolites might play a more important role in the development of early-onset T2D, whereas at advanced age, their effects might be obscured by the marked decreases in insulin sensitivity associated with ageing.[48] Besides, the different definitions of T2D used in the two cohorts might also be accountable. Alternatively, this observation might be a chance finding and needs to be clarified in further investigations.

Incremental risk prediction of metabolic biomarkers beyond traditional diabetes risk factors has been evaluated in some previous studies. The Framingham Offspring Study found that the c-statistic was barely improved by the identified metabolites associated with T2D risk in the study sample in which controls were randomly selected,[10] and the EPIC-Postdam study also reported a very modest improvement of c-statistic.[12] In our study, the incremental predictive ability of the four identified metabolites beyond traditional risk factors was also modest. Nevertheless, as reviewed by Sattar *et al.*,[49] the incremental usefulness of novel biomarkers for predicting future T2D is generally limited, in contrast to the remarkable predictive accuracy yielded by established risk factors. Despite the limited improvement in risk prediction, our findings are still important in providing new pathogenic insights in T2D development.

The strength of our study largely lies in its prospective nature, the inclusion of two independent population-based cohorts to reduce the possibility of chance findings, and the considerable sample size. To our best knowledge, the present study included the largest number of incident T2D cases among prospective investigations of metabolic profiles and T2D risk. Moreover, we were able to confirm previously reported associations of three metabolites (alanine, phenylalanine and tyrosine) with incident T2D while further discovering a novel one (palmitoylcarnitine) with a stringent criterion (FDR < 0.1 in both cohorts). Finally, incident T2D cases were diagnosed by a fasting glucose test in both cohorts, and plus HbA1c levels in the DFTJ cohort, rather than by self-reported medical history alone, therefore minimizing the contamination of undiagnosed and misdiagnosed cases.

Our study also has several limitations. First, our metabolomics platform only covered a limited number of metabolite targets, therefore not being able to capture the full metabolic profile of the study population. However, amino acids and (acyl) carnitines were well characterized in the present study, which provided information on protein and lipid metabolism, two important metabolic pathways in human body.[50,51] Second, although the associations between branched-chain amino acids and T2D risk did not reach the criterion of FDR < 0.1 in both cohorts, supportive evidence was obtained in meta-analysis of two cohorts. Third, participants in the current study consisted of middle-aged and elderly Chinese; as a result, caution must be taken when generalizing our findings to other populations.

Conclusions

In the present study, we confirmed the association of alanine, phenylalanine and tyrosine with future T2D risk, and further identified palmitoylcarnitine as a novel metabolic marker in two prospective cohorts of Chinese adults. If replicated, the identified metabolites may provide novel insights into the pathophysiology of T2D which has become a major health concern in China. Further studies are urgently warranted to elucidate the underlying mechanisms and explore related intervention strategies to reduce T2D risk.

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

This work was supported by: the Natural National Scientific Foundation of China [81230069 and 81390542 to T.W., and 81390543 to H.S.]; the National Key Basic Research and Development Program (973 project)[2011CB503806 to T.W.]; the 111 Project [to T.W.]; the Program for Changjiang Scholars and Innovative Research Team in University [to T.W.]; the Fundamental Research Funds for the Central Universities, HUST [to T.W.]; the Priority Academic Program for the Development of Jiangsu Higher Education Institutions (Public Health and Preventive Medicine)[to H.S.]; and the Flagship Major Development of Jiangsu Higher Education Institutions (to H.S.). The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Acknowledgements

We acknowledge the dedication, commitment and contribution of the working staff in Dongfeng Central Hospital and Dongfeng Motor Corporation. We thank all participants of the DFTJ cohort and JSNCD cohort. We confirm that those acknowledged have confirmed their agreement.

Author Contributions

G.Q. took part in the study design, data collection and the metabolomics analysis of plasma samples, did data analysis and drafted the manuscript. Y.Z. was involved in data analysis and commented on the drafting the manuscript. H.W., Y.X., Y.L. and Y.Y. took part in data collection and the metabolomics analysis. T.W. and A.P. were principal investigators who were responsible for the overall supervision and acted as the guarantors for the paper. F.B.H., H.S., Z.H., Y.J., X.Z. and M.H. took part in the study design and drafting of the manuscript. H.T. instructed the metabolomics analysis. H.M., J.S., H.Y., X.L., X.M., C.Z. and C.X. were involved in data collection. All authors had access to the data, commented on the report drafts and approved the final submitted version.

Conflict of interest: None.

References

1. Defronzo RA. Banting Lecture. From the triumvirate to the ominous octet:a new paradigm for the treatment of type 2 diabetes mellitus. *Diabetes* 2009;**58**:773–95.
2. McNeely MJ, Boyko EJ. Type 2 diabetes prevalence in Asian Americans. *Diabetes Care* 2004;**27**:66–69.
3. Hu FB. Globalization of diabetes. The role of diet, lifestyle, and genes. *Diabetes Care* 2011;**34**:1249–57.
4. Xu Y, Wang L, He J *et al*. Prevalence and control of diabetes in Chinese adults. *JAMA* 2013;**310**:948–59.
5. Wang S, Marquez P, Langenbrunner J. *Toward a Healthy and Harmonious Life in China: Stemming the Rising Tide of Noncommunicable Diseases*. Washington, DC: World Bank, 2011.
6. Wishart DS, Tzur D, Knox C *et al*. HMDB: the human metabolome database. *Nucleic Acids Res* 2007;**35**:D521–26.
7. Wishart DS, Jewison T, Guo AC *et al*. HMDB 3.0 - The human metabolome database in 2013. *Nucleic Acids Res* 2013;**41**:D801–07.
8. Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB. Metabolomics applied to diabetes research moving from information to knowledge. *Diabetes* 2009;**58**:2429–43.
9. Nishiumi S, Shinohara M, Ikeda A *et al*. Serum metabolomics as a novel diagnostic approach for pancreatic cancer. *Metabolomics* 2010;**6**:518–28.
10. Wang TJ, Larson MG, Vasan RS *et al*. Metabolite profiles and the risk of developing diabetes. *Nat Med* 2011;**17**:448–53.
11. Stancakova A, Civelek M, Saleem NK *et al*. Hyperglycemia and a common variant of GCKR are associated with the levels of eight amino acids in 9,369 Finnish men. *Diabetes* 2012;**61**:1895–902.
12. Floegel A, Stefan N, Yu ZH *et al*. Identification of serum metabolites associated with risk of Type 2 Diabetes using a targeted metabolomic approach. *Diabetes* 2013;**62**:639–48.
13. Wang-Sattler R, Yu ZH, Herder C *et al*. Novel biomarkers for prediabetes identified by metabolomics. *Mol Syst Biol* 2012;**8**:615.
14. Wang TJ, Ngo D, Psychogios N *et al*. 2-Aminoadipic acid is a biomarker for diabetes risk. *J Clin Invest* 2013;**123**:4309–17.
15. Rhee EP, Cheng S, Larson MG *et al*. Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *J Clin Invest* 2011;**121**:1402–11.
16. Drogan D, Dunn WB, Lin W *et al*. Untargeted metabolic profiling identifies altered serum metabolites of type 2 diabetes mellitus in a prospective, nested case control study. *Clin Chem* 2015;**61**:487–97.
17. Zhao JY, Zhu Y, Hyun N *et al*. Novel metabolic markers for the risk of diabetes development in American Indians. *Diabetes Care* 2015;**38**:220–27.
18. Tillin T, Hughes AD, Wang Q *et al*. Diabetes risk and amino acid profiles: cross-sectional and prospective analyses of ethnicity, amino acids and diabetes in a South Asian and European cohort from the SABRE (Southall And Brent REvisited) Study. *Diabetologia* 2015;**58**:968–79.
19. Walford GA, Ma Y, Clish C *et al*. Metabolite profiles of diabetes incidence and intervention response in the Diabetes Prevention Program. *Diabetes* 2016;**65**:1424–33.
20. Yu D, Moore S, Matthews C *et al*. Plasma metabolomic profiles in association with type 2 diabetes risk and prevalence in Chinese adults. *Metabolomics* 2015;**12**:1–11.
21. Kalim S, Clish CB, Wenger J *et al*. A plasma long-chain acylcarnitine predicts cardiovascular mortality in incident dialysis patients. *J Am Heart Assoc* 2013;**2**:e000542.
22. Wang F, Zhu J, Yao P *et al*. Cohort Profile: The Dongfeng-Tongji cohort study of retired workers. *Int J Epidemiol* 2013;**42**:731–40.
23. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2014;**37**(**Suppl 1**):S81–90.
24. Dunn W B, Broadhurst D, Begley P *et al*. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat Protoc* 2011;**6**:1060–83.
25. Zheng Y, Yu B, Alexander D *et al*. Associations between metabolomic compounds and incident heart failure among African Americans:The ARIC Study. *Am J Epidemiol* 2013;**178**:534–42.
26. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;**27**:157–72; discussion 207–12.
27. Parikh CR, Thiessen-Philbrook H. Key Concepts and limitations of statistical methods for evaluating biomarkers of kidney disease. *J Am Soc Nephrol* 2014;**25**:1621–29.
28. Felig P, Wahren J, Sherwin R, Palaiologos G. Amino acid and protein metabolism in diabetes mellitus. *Arch Intern Med* 1977;**137**:507–13.
29. Muller WA, Faloona GR, Unger RH. The effect of alanine on glucagon secretion. *J Clin Invest* 1971;**50**:2215–18.
30. Wurtz P, Tiainen M, Makinen VP *et al*. Circulating metabolite predictors of glycemia in middle-aged men and women. *Diabetes Care* 2012;**35**:1749–56.

31. Matthews DE. An overview of phenylalanine and tyrosine kinetics in humans. *J Nutr* 2007;**137**:1549–55s.
32. Krebs M, Krssak M, Bernroider E *et al*. Mechanism of amino acid-induced skeletal muscle insulin resistance in humans. *Diabetes* 2002;**51**:599–605.
33. Tremblay F, Brule S, Um SH *et al*. Identification of IRS-1 Ser-1101 as a target of S6K1 in nutrient- and obesity-induced insulin resistance. *Proc Natl Acad Sci U S A* 2007;**104**:14056–61.
34. 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.
35. Wurtz P, Makinen VP, Soininen P *et al*. Metabolic signatures of insulin resistance in 7,098 young adults. *Diabetes* 2012;**61**: 1372–80.
36. Rinaldo P, Matern D, Bennett MJ. Fatty acid oxidation disorders. *Annu Rev Physiol* 2002;**64**:477–502.
37. Yu Y, Cai Z, Zheng J *et al*. Serum levels of polyunsaturated fatty acids are low in Chinese men with metabolic syndrome, whereas serum levels of saturated fatty acids, zinc, and magnesium are high. *Nutr Res* 2012;**32**:71–77.
38. Wang L, Folsom AR, Zheng ZJ, Pankow JS, Eckfeldt JH; ARIC Study investigators. Plasma fatty acid composition and incidence of diabetes in middle-aged adults:the Atherosclerosis Risk in Communities (ARIC) Study. *Am J Clin Nutr* 2003;**78**:91–98.
39. Patel PS, Sharp SJ, Jansen E *et al*. Fatty acids measured in plasma and erythrocyte-membrane phospholipids and derived by food-frequency questionnaire and the risk of new-onset type 2 diabetes: a pilot study in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Norfolk cohort. *Am J Clin Nutr* 2010;**92**:1214–22.
40. Forouhi NG, Koulman A, Sharp SJ *et al*. Differences in the prospective association between individual plasma phospholipid saturated fatty acids and incident type 2 diabetes: the EPIC-InterAct case-cohort study. *Lancet Diabetes Endocrinol* 2014;**2**:810–18.
41. Aguer C, McCoin CS, Knotts TA *et al*. Acylcarnitines: potential implications for skeletal muscle insulin resistance. *FASEB J* 2015;**29**:336–45.
42. 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 glucolipotoxicity. *Obesity (Silver Spring)* 2010;**18**:1695–700.
43. Mai M, Tonjes A, Kovacs P, Stumvoll M, Fiedler GM, Leichtle AB. Serum levels of acylcarnitines are altered in prediabetic conditions. *PLoS One* 2013;**8**:e82459.
44. Zhang X, Zhang C, Chen L, Han X, Ji L. Human serum acylcarnitine profiles in different glucose tolerance states. *Diabetes Res Clin Pract* 2014;**104**:376–82.
45. 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.
46. Gogna N, Krishna M, Oommen AM, Dorai K. Investigating correlations in the altered metabolic profiles of obese and diabetic subjects in a South Indian Asian population using an NMR-based metabolomic approach. *Mol Biosyst* 2015;**11**:595–606.
47. Zhou Y, Qiu L, Xiao Q *et al*. Obesity and diabetes related plasma amino acid alterations. *Clin Biochem* 2013;**46**:1447–52.
48. Paolisso G, Scheen A, Lefevre P. Glucose handling, diabetes and ageing. *Horm Res* 1995;**43**:52–57.
49. Sattar N, Wannamethee SG, Forouhi NG. Novel biochemical risk factors for type 2 diabetes: pathogenic insights or prediction possibilities? *Diabetologia* 2008;**51**:926–40.
50. Munro HN (ed). *Mammalian Protein Metabolism*. Vol. 3. New York, NY: Acacemic Press, 2013.
51. Snyder F. *Lipid Metabolism in Mammals*. New York, NY: Springer, 2012.