



Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis

Diabetes Care 2016;39:833–846 | DOI: 10.2337/dc15-2251

Marta Guasch-Ferré,^{1,2,3} Adela Hruby,¹
 Estefanía Toledo,^{3,4} Clary B. Clish,⁵
 Miguel A. Martínez-González,^{3,4}
 Jordi Salas-Salvadó,^{2,3} and Frank B. Hu^{1,6,7}

OBJECTIVE

To conduct a systematic review of cross-sectional and prospective human studies evaluating metabolite markers identified using high-throughput metabolomics techniques on prediabetes and type 2 diabetes.

RESEARCH DESIGN AND METHODS

We searched MEDLINE and EMBASE databases through August 2015. We conducted a qualitative review of cross-sectional and prospective studies. Additionally, meta-analyses of metabolite markers, with data estimates from at least three prospective studies, and type 2 diabetes risk were conducted, and multivariable-adjusted relative risks of type 2 diabetes were calculated per study-specific SD difference in a given metabolite.

RESULTS

We identified 27 cross-sectional and 19 prospective publications reporting associations of metabolites and prediabetes and/or type 2 diabetes. Carbohydrate (glucose and fructose), lipid (phospholipids, sphingomyelins, and triglycerides), and amino acid (branched-chain amino acids, aromatic amino acids, glycine, and glutamine) metabolites were higher in individuals with type 2 diabetes compared with control subjects. Prospective studies provided evidence that blood concentrations of several metabolites, including hexoses, branched-chain amino acids, aromatic amino acids, phospholipids, and triglycerides, were associated with the incidence of prediabetes and type 2 diabetes. We meta-analyzed results from eight prospective studies that reported risk estimates for metabolites and type 2 diabetes, including 8,000 individuals of whom 1,940 had type 2 diabetes. We found 36% higher risk of type 2 diabetes per study-specific SD difference for isoleucine (pooled relative risk 1.36 [1.24–1.48]; $I^2 = 9.5\%$), 36% for leucine (1.36 [1.17–1.58]; $I^2 = 37.4\%$), 35% for valine (1.35 [1.19–1.53]; $I^2 = 45.8\%$), 36% for tyrosine (1.36 [1.19–1.55]; $I^2 = 51.6\%$), and 26% for phenylalanine (1.26 [1.10–1.44]; $I^2 = 56\%$). Glycine and glutamine were inversely associated with type 2 diabetes risk (0.89 [0.81–0.96] and 0.85 [0.82–0.89], respectively; both $I^2 = 0.0\%$).

CONCLUSIONS

In studies using high-throughput metabolomics, several blood amino acids appear to be consistently associated with the risk of developing type 2 diabetes.

Type 2 diabetes is among the most prevalent chronic diseases, affecting more than 380 million people worldwide in 2014 (1). Risk factors for type 2 diabetes consist of a combination of unhealthy diet, lifestyle, and genetic factors that may interact with each other and with the environment (2). In the last decade, high-throughput metabolomics technologies have provided insights into the pathophysiological

¹Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA

²Human Nutrition Unit, Faculty of Medicine and Health Sciences, Pere Virgili Institute for Health Research, Rovira i Virgili University, Reus, Spain

³CIBER Fisiopatología de la Obesidad y Nutrición, Instituto de Salud Carlos III, Madrid, Spain

⁴Department of Preventive Medicine and Public Health, University of Navarra, Health Research Institute of Navarra, Pamplona, Spain

⁵Broad Institute, Cambridge, MA

⁶Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA

⁷Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

Corresponding author: Frank B. Hu, nhbhf@channing.harvard.edu.

Received 14 October 2015 and accepted 6 February 2016.

This article contains Supplementary Data online at <http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc15-2251/-/DC1>.

This study was registered in PROSPERO registry, reg. no. CRD42015023439, <http://www.crd.york.ac.uk/PROSPERO/>.

© 2016 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered.

pathways and understanding of the disease and its precedents (3). Metabolomics refers to the systematic analysis of metabolites (low molecular weight biochemicals including sugars, amino acids, organic acids, nucleotides, and lipids) in a biological sample.

Several studies have evaluated the relationship between a wide range of metabolites, insulin resistance, and type 2 diabetes, using mainly two different techniques: mass spectrometry (MS) coupled with gas- or liquid-phase chromatography (GC and LC, respectively) and proton (^1H) nuclear magnetic resonance (NMR) spectroscopy (3,4). With these technologies, researchers have taken both targeted approaches that focus on a specific subset of defined metabolites and more agnostic untargeted approaches that analyze a large number of measurable molecules found in a sample, including chemical unknowns (4).

To date, published findings suggest that branched-chain amino acids (BCAAs), identified using high-throughput metabolomics, are associated with insulin resistance (3,5) and type 2 diabetes (5). Other amino acids such as aromatic amino acids, glycine, glutamine, and glutamate have also been related to prediabetes and type 2 diabetes risk (5–7). Metabolomics have revealed as well that several sugar metabolites and gluconeogenesis substrates, including glucose and fructose, are higher in individuals with prediabetes compared with control subjects (8). Finally, lipid subclasses such as phospholipids, sphingomyelins, and triglycerides, and also specific lipids like palmitate and palmitoleate, have all been related to insulin resistance and type 2 diabetes in humans (5,12). Expanding the current knowledge on the physiopathology of type 2 diabetes and identifying novel predictive biomarkers may help to facilitate the detection and management of diabetes. However, no systematic review or meta-analysis on this topic has been published.

We hypothesized that elevated concentrations of some metabolites, such as BCAAs and aromatic amino acids, would be associated with a higher risk of prediabetes and type 2 diabetes, while other metabolites (e.g., glycine and glutamine) would be inversely associated. We conducted a systematic review of human studies assessing

metabolite markers, identified using high-throughput metabolomics, of prediabetes and type 2 diabetes. Additionally, we conducted quantitative meta-analyses for specific metabolite markers of type 2 diabetes risk: BCAAs (isoleucine, leucine, and valine), aromatic amino acids (tyrosine and phenylalanine), glycine, glutamine, alanine, and histidine.

RESEARCH DESIGN AND METHODS

Data Sources and Searches

We followed the *Cochrane Handbook of Systematic Reviews* (<http://handbook.cochrane.org>) and used the MOOSE (Meta-analysis Of Observational Studies in Epidemiology) checklist. The protocol has been registered in the PROSPERO registry (http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015023439). We conducted a systematic search of published literature in any language in two different databases, MEDLINE (<http://www.ncbi.nlm.nih.gov/pubmed>) and EMBASE (<http://gateway.ovid.com>), from the earliest available online indexing year through August 2015, and hand searched reference lists of articles and key journals for human studies evaluating metabolite biomarkers associated with insulin resistance and type 2 diabetes, identified using high-throughput metabolomics techniques. Our search terms combined the exposure (metabolite markers) with several outcomes (insulin resistance, insulin sensitivity, and type 2 diabetes) in humans. The detailed search strategy is presented in Supplementary Data.

Study Selection and Eligibility Criteria

Titles and abstracts were screened in duplicate by two of three authors (M.G.-F., A.H., and E.T.) for eligibility. The third of these authors resolved disagreements. Studies were eligible for inclusion if they were human studies (cohort, case-cohort, case-control, or clinical trials) and if metabolites were identified using high-throughput metabolomics techniques in blood (plasma or serum) or urine samples. We have included studies in which prediabetes and type 2 diabetes were defined as impaired glucose tolerance, impaired fasting glucose, insulin resistance, or impaired insulin sensitivity (HOMA of insulin resistance [HOMA-IR] or HOMA of β -cell function [HOMA-B])

according to standard criteria for diagnosis and classification of diabetes (1,10) (e.g., impaired fasting glucose defined as fasting glucose 100–125 mg/dL [5.6–6.9 mmol/L] and type 2 diabetes defined as fasting glucose ≥ 126 mg/dL [≥ 7.0 mmol/L] or other criterion-justified ranges). We excluded animal studies; studies conducted in children, adolescents, and pregnant women; studies of type 1 diabetes or gestational diabetes mellitus; nonoriginal papers (reviews, commentaries, editorials, or letters); and duplicate publications.

Data Extraction and Quality Assessment

From each study, we extracted the following information: authors, year of publication and journal, study name and design, location, sample size, setting, participant characteristics, duration of follow-up (if applicable), analytic technique, biological sample, primary and secondary outcomes measured, statistical tests, confounding factors, whether the analyses were corrected for multiple comparisons, and major findings (analyzed metabolites, adjusted relative risks [RRs] or odds ratios [ORs]). We evaluated and scored the quality of the studies independently and in duplicate on a six-point scale based on guidelines adapted from reference 11. The criteria (up to 1 point per criterion) included reporting of study participation and attrition, measurement of exposure and outcome, measurement of and accounting for confounding, and the appropriateness of the statistical analysis. Points were summed, and studies with scores 0–3 were considered to be of low quality, while scores 4–6 were considered of high quality.

Data Synthesis and Analysis

We conducted a qualitative review of findings in nonprospective studies (cross-sectional and case-control). We extracted the relevant information from the identified studies, as detailed above, and described and summarized the findings of these reports in a qualitative manner.

For prospective cohort studies, we conducted a qualitative review and quantitative meta-analyses. For the qualitative review, we extracted and described the information for each study, and for the quantitative review we combined the main results from these

studies using meta-analyses. In our meta-analyses of the associations between amino acids and type 2 diabetes, we considered all prospective studies that provided any multivariable-adjusted effect estimate (OR, RR, or hazard ratio [HR]) with an accompanying measure of uncertainty (CI, SE, or other data to calculate variance). However, we only meta-analyzed the estimates of metabolites that were reported in at least three different prospective studies; i.e., three data points was our minimum threshold for conducting a meta-analysis. Studies tended to report the risk for type 2 diabetes events per 1 SD of a given metabolite; other studies provided means and tests for differences between means. For the studies providing means (SD) in case and control subjects (12,13), we used the Hasselblad and Hedges method to convert the measurements to OR and 95% CI (14,15) prior to inclusion in meta-analyses. In the primary analyses, we only included those studies reporting the risk estimates as OR, RR, or HR with the respective measure of uncertainty. We then conducted analyses that additionally included converted measurements from means. The rationale to include the converted measurements as secondary analyses was that the reported means were typically unadjusted for confounding factors (12,13) in contrast to risk estimates. Summary RRs using both random-effects and fixed-effects models were obtained with the calculation of the logarithm of the RRs and corresponding 95% CIs of the individual studies. Our primary approach was the random-effects model because it incorporates both within- and between-study components of variance. Fixed-effects models were evaluated secondarily. We used inverse variance weighting to derive an overall estimate in the meta-analyses. Forest plots were used to evaluate risk estimates across studies. Heterogeneity among studies was estimated by the Cochran Q test and I^2 statistic, with $>30\%$ considered at least moderate heterogeneity (<http://handbook.cochrane.org>). Heterogeneity was considered statistically significant at $P \leq 0.10$. Potential sources of heterogeneity were explored in meta-regressions and subgroup analyses, such as study design (case-control, nested case-control, cohort), length of follow-up (≤ 7 years, > 7 years), biological sample

(plasma, serum), and analytic technique (MS, NMR). To assess potential heterogeneity by sex, we performed meta-regressions using the proportion of men as the dependent variable. Publication bias was assessed by visual inspection of funnel plots, evaluating skewness (nonsymmetry) of the distribution of SEs around the study-level effect estimates, and the Egger and Begg tests, using a significance level of $P < 0.05$ to indicate significant asymmetry (16,17). All analyses were performed using STATA (version 12.0; StataCorp LP, College Station, TX) with a two-tailed α of 0.05 considered statistically significant.

RESULTS

Literature Search Results

Figure 1 shows the results of the literature search and study selection (metabolites altered in prediabetes/type 2 diabetes and metabolite predictors of prediabetes/type 2 diabetes). Among the 1,019 unique abstracts reviewed independently and in duplicate by two investigators, 943 were excluded (Fig. 1). Among the 76 full text articles reviewed thereafter, 38 were excluded. After final exclusions, and with addition of 8 references identified by hand searching of citations, 46 publications met the inclusion criteria and were included in the present systematic review.

Metabolites Altered in Prediabetes or Type 2 Diabetes (Nonprospective Studies)

We identified 27 cross-sectional and/or case-control publications, encompassing 29 analyses in different studies, which compared metabolite concentrations in those without (healthy control subjects) versus those with prediabetes and/or type 2 diabetes (case subjects) (Supplementary Table 1). Of these, seven focused on prediabetes-related measures (18–24), mainly using 2-h post-oral glucose tolerance glucose, and/or HOMA-IR. Another nine studies (25–33) focused on both prediabetes and type 2 diabetes. The remaining 11 studies (34–44) evaluated type 2 diabetes as the main outcome. Most study designs were case-control studies ($n = 24$), in which blood or urine was collected at the same time as case status was defined. The number of participants ranged from 20 to 7,098.

More than 20 studies evaluated the association between amino acids and

prediabetes/type 2 diabetes with findings that were significant. In particular, in a study carried out in Singapore, HOMA-IR was associated with higher levels of plasma BCAAs, aromatic amino acids, and the ratio of glutamate to glutamine (18). These findings were in accordance with those of Newgard et al. (19) where BCAAs, aromatic amino acids, and glutamate-to-glutamine ratio were positively associated with prediabetes in obese individuals. An untargeted metabolomics approach in 2,204 women from the TwinsUK study found that BCAAs and derivatives were associated with both impaired fasting glucose and type 2 diabetes status (30). At least 10 studies analyzed the relationship between hydroxy acids and prediabetes/type 2 diabetes, most of them finding positive significant associations. Gall et al. (21) observed that α -hydroxybutyrate was elevated in insulin resistant individuals and in those with impaired glucose tolerance. Furthermore, 3-hydroxybutyrate and β -hydroxybutyrate were found to be elevated in individuals with type 2 diabetes from the Cooperative Health Research in the Region of Augsburg (KORA) study.

Findings from these studies also suggest that some acylcarnitines are associated with higher risk of prediabetes/type 2 diabetes. Of note, acetylcarnitine C2 concentrations were significantly higher in participants with type 2 diabetes—by 157%—compared with those without diabetes from Project SuGAR (Sea Island Genetic African American Family Registry) (28). Medium-chain acylcarnitines C6-carnitine, C8-carnitine, and C10-carnitine were higher and propionylcarnitine concentrations were lower in participants with diabetes compared with nondiabetic control subjects (38). Higher concentrations of plasma long-chain acylcarnitines and free carnitines were observed among participants with type 2 diabetes in another case-control study conducted in a U.S. population in which 46 acylcarnitines were targeted (27).

At least 20 analyses included metabolomics profiling of lipids. Several lipid classes (plasma phospholipids, cholesterol esters, triglycerides, glycerophospholipids, and sphingolipids), identified using high-throughput techniques, were found to be different between those with and without prediabetes/type 2

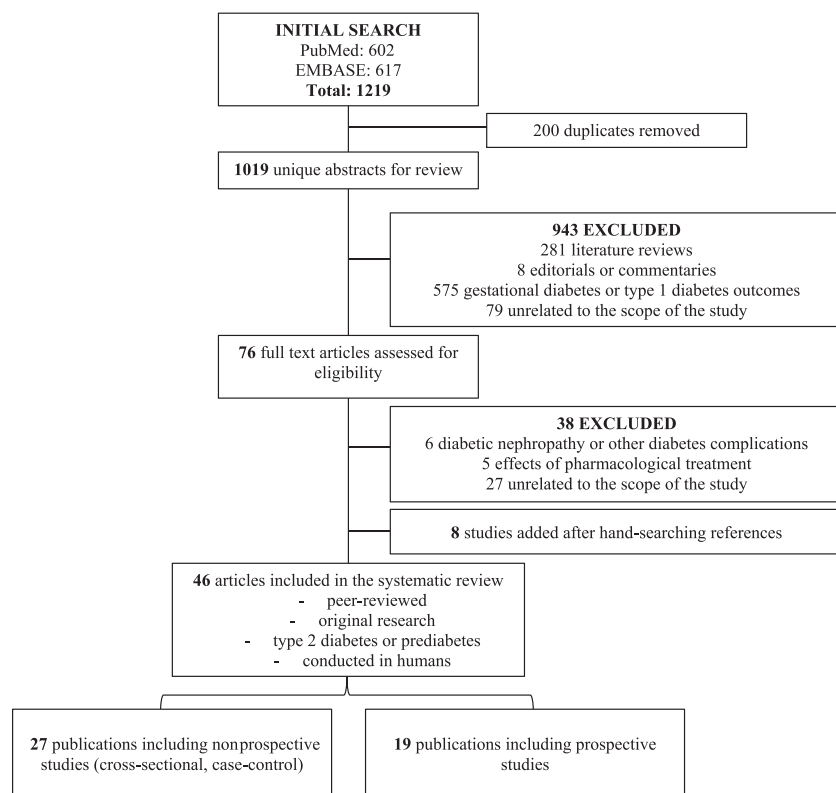


Figure 1—Flow diagram of literature search and study selection for metabolite markers of prediabetes and type 2 diabetes.

diabetes. Low carbon number–saturated lipids (myristic, palmitic, and stearic acids) were elevated in individuals with impaired fasting glucose and type 2 diabetes compared with control subjects, while some glycerophospholipids and sphingomyelins were reduced (31). In another study, lysophosphatidylcholines (LPCs) and lysophosphatidylethanolamines (LPEs) were significantly higher in men with type 2 diabetes, as were other fatty acids (e.g., dodecanoic and myristic acid) compared with nondiabetic control subjects (42).

Sugar metabolites including glucose, dihexose, mannose, arabinose, and fructose and also glycolipids showed positive associations with the prevalence of prediabetes and/or type 2 diabetes in >10 studies. Case-control results from the KORA study showed that plasma 1,5-anhydroglucitol was 37.8% lower in participants with diabetes compared with the control group, while concentrations of glucose, mannose, desoxyhexose, and dihexose were higher (40). Consistently, glucose, mannose, and fructose were associated with impaired fasting glucose and type 2 diabetes, whereas 1,5-anhydroglucitol was

lower in case subjects compared with nondiabetic control subjects (30). Finally, some organic acids like lactate, maleic acid, dimethyl ester, and acetic acid (34,36), and other compounds like purines (23) and urea cycle metabolites (citrulline, ornithine, and arginine), were associated with prediabetes and type 2 diabetes prevalence in isolated studies.

Metabolites Predicting Prediabetes or Type 2 Diabetes (Prospective Studies)

Of the 33 prospective analyses, 22 analyses were focused on type 2 diabetes as the main outcome (10 nested case-control, 1 case-cohort, and 11 cohort studies) (5–7,9,12,13,45–54); 11 analyses evaluated prediabetes (2 case-control and 9 cohort studies) (5,8,9,13,45,47,48,50,53,55,56). Prospective replications in independent cohorts were conducted in five of the studies (5,6,46,49,57) (Table 1). The number of participants in these studies ranged from 124 to 6,607 (total of 4,167 diabetes case subjects and 29,926 participants without diabetes). Follow-up after sample collection ranged from 6 months to 19 years. Eighteen analyzed plasma and 15 serum samples, and 24 studies conducted MS

and nine NMR. Most studies used targeted metabolomics analysis ($n = 20$) analyzing between 50 and 200 metabolites, except for five analyses that focused on fewer metabolites (between 1 and 23). Two studies applied an untargeted approach (one assessing >4,500 MS features and the other >11,000) (Table 1). Most of the prospective studies included as covariates in their models age, sex, BMI, and fasting glucose. Some studies included further adjustment for smoking, alcohol intake, physical activity, education, coffee intake, waist circumference, blood cholesterol, triglycerides, fasting insulin, HbA_{1c}, and dietary variables. In meta-analyses (see below), we used the estimate from the most adjusted model reported in a given study. All the prospective studies included had high-quality scores.

Amino Acids

Of the 33 analyses, 19 evaluated the association between amino acid metabolites and prediabetes/type 2 diabetes. Besides BCAAs and aromatic amino acids (16 analyses), other common amino acids metabolites analyzed were glycine, glutamine, and glutamate. Seventeen studies used a targeted approach, targeting on average 15–20 amino acids, and two studies used untargeted approaches. LC-MS was the most common analytic technique used. Most analyses reported significant associations for BCAAs and aromatic amino acids. The findings from the Piekämäki cohort, which targeted 134 metabolites, reported that BCAAs and phenylalanine were associated with fasting and 2-h glucose, while alanine and tyrosine were associated with post-oral glucose tolerance test (OGTT) glucose after adjustment for fasting insulin and postload glucose (8). The association between BCAAs and prediabetes was confirmed in four other population-based analyses conducted in both American and European populations (5,50,55,56).

Of all the analyses, at least 10 analyzed the amino acid glycine, with most suggesting an inverse association with prediabetes/type 2 diabetes. Glutamine was evaluated in at least eight analyses, but only three reported significant inverse associations (6,12,46). In addition, the glutamine-to-glutamate ratio was found to predict insulin resistance after 5 years in participants of the Insulin Resistance Atherosclerosis Study (IRAS) (50).

Table 1—Characteristics of studies investigating prospective associations of metabolites and/or type 2 diabetes

Reference	Study population location	Study design	N, follow-up time	Technique and metabolite targets	Biological sample	Outcome	Covariates in fully adjusted model	Key findings	Other substudies included	Quality score†
Rhee et al., 2011, J Clin Invest (9)	FHS Offspring, U.S.	Nested case-control	378 (189 no diabetes/189 T2D), 12 yrs	LC-MS/MS; targeted (>100 lipids: TAGs, CEs, LPCs, PCs, LPEs, DAGs, SMs)	Plasma	Prediabetes (fasting and post-OGTT insulin and glucose, HOMA-IR)	Age, sex, BMI, fasting glucose, fasting insulin, TG, HDL-C	Lipids: (†) TAGs of lower carbon no. and double bond correlated with HOMA-IR; TAGs of higher carbon no. and double bond not correlated with HOMA-IR 15 lipid metabolites associated with incident T2D; (†) TAGs of lower carbon no. and double bond, LPEs 18:2, SM 22:0, PC 36:2; (†) TAGs of higher carbon no. and double bond, LPC 22:6, PC 38:6	Diet, pharmacological, and acute exercise test substudies	5
Wang et al., 2011, Nat Med (5)	FHS Offspring, U.S.	Nested case-control	378 (189 no diabetes/189 T2D), 12 yrs	LC-MS/MS; targeted (61 amino acids, biogenic amines, other polar plasma metabolites)	Plasma	Prediabetes (fasting and post-OGTT insulin and glucose, HOMA-IR, HOMA-β)	Age, sex, BMI, fasting glucose, insulin resistance or secretion measure*	BCAAs, aromatic amino acids: (†) 5 amino acids (isoleucine, leucine, valine, tyrosine, phenylalanine) correlated with fasting insulin, HOMA-IR, HOMA-β (†) 5 amino acids (isoleucine, leucine, valine, tyrosine, phenylalanine); (†) combination of isoleucine, tyrosine, phenylalanine; (–) changes in the 5 amino acids during OGTT not associated with T2D (†) leucine, valine, tyrosine, phenylalanine; (–) isoleucine; (†) combination of isoleucine, tyrosine, and phenylalanine	Analysis in 400 control subjects randomly selected from FHS Offspring, confirming T2D risk based on amino acid combination	5.5
Würtz et al., 2012, Diabetes Care (8)	Pieksämäki cohort, Finland	Cohort, prospective, population based	618 free of diabetes and with OGTT data, 6.5 yrs	NMR; targeted (prospectively, 19 metabolites cross-sectionally associated from 134 targeted amino acids, gluconeogenic substrates, and lipids)	Serum	Prediabetes (fasting and post-OGTT glucose)	Age, sex, BMI, systolic blood pressure, TG, HDL-C, fasting insulin, fasting and post-OGTT glucose	BCAAs, gluconeogenesis substrates: (†) BCAAs (leucine, isoleucine, valine), phenylalanine, α-/acid glycoprotein predicted fasting and post-OGTT glucose; (†) alanine, lactate, pyruvate, tyrosine predicted only post-OGTT glucose; (–) fatty acids not associated	Cross-sectional study (Pieksämäki cohort and Health 2000 Study, n = 1,873; 19 metabolites significant, moved forward to prospective analysis); dynamic study over 6.5 yrs of metabolite changes vs. changes in fasting and post-OGTT glucose: changes in fatty acids associated	

Continued on p. 838

Table 1—Continued

Reference	Study location	Study design	N, follow-up time	Technique and metabolite targets	Biological sample	Outcome	Covariates in fully adjusted model	Key findings	Other substudies included	Quality score†
Shah et al., 2012, <i>Diabetologia</i> (65)	WLM, U.S.	Cohort, prospective, substudy of participants from a clinical trial	500 participants who lost ≥4 kg weight in the clinical trial, baseline and 6 months	MS/MS; targeted (60 metabolites, 45 acylcarnitines, 15 amino acids)	Plasma	Prediabetes (HOMA-IR at baseline and at 6 months and 6-month changes)	Sex, education, income, weight, change in weight, Healthy Eating Index*	BCAAs: (†) from factor analysis, factor of baseline BCAAs and related catabolites correlated with baseline HOMA-IR, 6-month HOMA-IR, changes in HOMA-IR, (–) other factors composed of acylcarnitines, ketone-related, and urea cycle metabolites not associated	Replication of significant results in 12 obese patients with diabetes before and 1 month after gastric bypass surgery and 10 patients before and 1 month after low-calorie diet	4.5
Stancáková et al., 2012, <i>Diabetes</i> (12)	METSIM, Finland	Nested case-control	526 men (375 normoglycemic/151 T2D), 4.7 yrs	NMR: targeted (8 amino acids: alanine, phenylalanine, valine, leucine, isoleucine, tyrosine, histidine, glutamine)	Serum	T2D	Age, BMI, 0–30 min postload insulin/glucose*	BCAAs, amino acids: (†) alanine, leucine, isoleucine, tyrosine, phenylalanine; (†) glutamine; (–) histidine, valine	Cross-sectional substudy (9,369 without diabetes); gene expression substudy: single GCKR variant of 43 loci tested that associated with amino acids evaluated	4.5
Wang-Sattler et al., 2012, <i>Mol Syst Biol</i> (45)	KORA, Germany	Cohort, prospective, population based	876 no diabetes at baseline/91 incident T2D, 7 yrs (also nested case-control: 91 T2D/91 propensity-matched control, 7 yrs); 641 NGT at baseline/118 incident IG, 7 yrs	LC-FIA-MS; targeted (131 metabolites initially targeted in cross-sectional analyses: hexose, acylcarnitines, amino acids, biogenic amines, SMs, PCs, LPCs; 3 carried forward in prospective analyses; BCAAs also tested)	Serum	Prediabetes (incident IG)	Age, sex, BMI, physical activity, alcohol intake, smoking, blood pressure, HDL-C, HbA _{1c} , fasting glucose, insulin*	Amino acids, lipids: (†) glycine, LPC 18:2; combination of (†) glycine, LPC 18:2 and (†) acetylcarmitine C2; (–) BCAAs (isoleucine, leucine, valine, tyrosine, phenylalanine) not associated	Cross-sectional substudy (1,206 no diabetes/91 with diabetes); cross-sectional replication in EPIC-Potsdam (1,253 no diabetes/133 with diabetes at baseline); gene expression study associated	5.5

Continued on p. 839

Table 1—Continued

Reference	Study, population location	Study design	N, follow-up time	Technique and metabolite targets	Biological sample	Outcome	Covariates in fully adjusted model	Key findings	Other substudies included	Quality score†
Cheng et al., 2012, Circulation (46)	FHS Offspring, U.S.	Cohort, prospective, population based	601 no diabetes/T2D NR, 12 yrs	LC-MS/MS; targeted 2 selected metabolites based on cross-sectional analysis; glutamine and glutamate	Plasma	T2D	Age, sex, BMI, baseline fasting glucose	Amino acids: (†) glutamate; (‡) glutamine and glutamine-to-glutamate ratio	Cross-sectional study of 45 metabolites (FHS, n = 1,015, and MDC, n = 746, without diabetes); FHS: († HOMA) glutamate, alanine, asparagine, betaine, carnitine, dimethylglycine, kynurenic acid, BCAAs, phenylalanine, proline, tyrosine, valine; (‡ HOMA) glutamine and glutamine-to-glutamate ratio, glycine	4.5
MDC, Sweden		Cohort, population based	409 no diabetes, T2D NR, 12 yrs	LC-MS/MS; targeted 2 selected metabolites based on cross-sectional analysis; glutamine and glutamate	Plasma	T2D	Age, sex, BMI, baseline fasting glucose	(‡) Glutamine; nonsignificant associations for glutamate and glutamine-to-glutamate ratio		
Ferrannini et al., 2013, Diabetes (13)	RISC, Europe (13 countries)	Cohort, population based	1,261 (779 no diabetes/123 incident IGT, IFG, or T2D), 3 yrs	LC-MS/MS; targeted 145 metabolites, amino acids, urea cycle, nucleotide metabolism	Plasma	Prediabetes (HOMA)	Age, sex, BMI, parental history of diabetes, fasting glucose	α-HB, L-GPC, BCAAs, amino acids: (†) α-HB; (‡) L-GPC	Cross-sectional substudy: 2,030 participants in RISC and 2,453 in Botnia); in vitro substudy	5.5
Botnia, Finland		Cohort, prospective, population based	2,580 (2,429 no diabetes/151 T2D), 9.5 yrs; 542 (412 no diabetes/130 T2D), 9.5 yrs	UHPLC-MS/MS; targeted (α-HB, L-GPC; ≥23 other amino acids and fatty acids in 542 additional participants)	Plasma	T2D	Age, sex, BMI, parental history of diabetes, fasting glucose	(†) α-HB; (‡) L-GPC; in 542 participants: (†) BCAAs (leucine, isoleucine, valine), glutamate, alanine, arginine; (‡) glycine		
Mahendran et al., 2013, Diabetes Care (47)	METSIM, Finland	Cohort, prospective, population based	4,335 men (4,059 no diabetes/276 T2D), 4.5 yrs	NMR; targeted (lipids, glycerol, free fatty acids, serum fatty acid profile)	Serum	Prediabetes (fasting and postload glycemia with OGTT)	Age, BMI, smoking, physical activity, Matsuda ISI	Lipids: (†) glycerol, total triglycerides, MUFA, FFAs, SFAs, and n-7 and n-9; (‡) n-6 and linoleic acid separately; (–) nonsignificant associations for n-3 triglycerides, MUFA, FFAs, SFAs, n-7 and n-9; (‡) n-3, n-6, linoleic acid separately	Cross-sectional substudy (9,398 men from the METSIM)	4.5
Mahendran et al., 2013, Diabetes (48)	METSIM, Finland	Cohort, prospective, population based	4,335 men (4,059 no diabetes/276 T2D), 5 yrs	NMR; targeted (β-hydroxybutyrate and acetoacetate)	Serum	Prediabetes (glucose AUC)	Age, BMI, smoking, physical activity	(†) β-Hydroxybutyrate, acetoacetate	Gene expression study	4.5
						T2D	Age, BMI, smoking, physical activity	(†) Acetoacetate		

Continued on p. 840

Table 1—Continued

Reference	Study, population location	Study design	N, follow-up time	Technique and metabolite targets	Biological sample	Outcome	Covariates in fully adjusted model	Key findings	Other substudies included	Quality score†
Floegel et al., 2013, Diabetes (6)	EPIC-Potsdam, Germany	Case cohort	3,082 (2,282 no diabetes/800 T2D), 7 yrs	FIA-MS/MS; targeted (163 metabolites, acylcarnitines, amino acids, hexose, glycerophospholipids, SMs)	Serum	T2D	Age, sex, BMI, waist circumference, alcohol intake, smoking, physical activity, education, coffee, red meat, whole-grain bread, hypertension*	BCAAs, amino acids, lipids, sugars; 14 metabolites significantly associated with T2D; (†) hexose, phenylalanine, diacyl-PC (C32:1, c36:1, C38:3 C40:5), factor combining these metabolites (factor 1); BCAAs; (‡) glycine, SM C16:1, lysophosphatidylcholine C18:2 acyl-alkyl-PC C34:3, C40:6, C42:5, C44:4, C44:5, factor combining these metabolites (factor 2)	Cross-sectional substudy of prediabetics (76 Caucasians from TuF study)	5.5
Wang et al., 2013, J Clin Invest (49)	FHS Offspring, U.S.	Nested case-control	376 (188 no diabetes/188 T2D), 12 yrs	HILIC and LC-MS/MS; targeted (70 metabolites)	Plasma	T2D	Age, sex, BMI and fasting glucose, insulin measures*	(†) Hexose, factor combining PCs and amino acids (factor 2)	Additional substudy in 1,561 randomly selected from FHS; substudy in mice and cell tissues	5.5
Wurtz et al., 2013, Diabetes Care (56)	Cardiovascular Risk in Young Finns, Finland	Cohort, prospective, population based	1,680 young adults without diabetes, 6 yrs	NMR spectroscopy and isocratic and reverse-phase HPLC; targeted (11 amino acids [BCAAs, aromatic, gluconogenic, and others])	Serum	Prediabetes (HOMA-IR)	Age, sex, BMI and fasting glucose, insulin measures*	BCAAs, aromatic amino acids; (†) isoleucine, leucine, valine, phenylalanine, tyrosine predicted HOMA-IR in men	Cross-sectional substudy	4
Padberg et al., 2014, PLoS One (7)	Bavarian Red Cross, Germany	Nested case-control	124 (28 T2D/96 healthy control), 6 yrs	GC-SIM-MS, GC-MS, and LC-MS/MS techniques; targeted (196 lipids, amino acids, and sugars)	Plasma	T2D	Age, sex, BMI, diabetes, IFG, center*	Sugar metabolites: (†) mannose, 2-hydroxybutyrate and 3-hydroxybutyrate	Cross-sectional substudy in 325 patients (59 T2D)	5
Palmer et al., 2015, J Clin Endocrinol Metab (50)	IRAS, U.S. (European-, Hispanic-, and African American)	Cohort, prospective, population based	147 participants (72 high insulin sensitivity/75 low insulin sensitivity); 146 participants (70 no diabetes/76 T2D), 5 yrs	MS/MS; targeted (69 metabolites, amino acids, acylcarnitines)	Plasma	Prediabetes-insulin resistance (OGTT)	Age, sex, BMI, ethnicity	(‡) Glycine and combined aspartate/asparagine; (†) valine, leucine or isoleucine, phenylalanine, glutamine/glutamate; (†) 3-hydroxybutyryl carnitine C4-OH, C10:3 and C8:1	None	4.5

Continued on p. 841

Table 1—Continued

Reference	Study, population location	Study design	N, follow-up time	Technique and metabolite targets	Biological sample	Outcome	Covariates in fully adjusted model	Key findings	Other substudies included	Quality score†
Drogan et al., 2015, Clin Chem (51)	EPIC-Potsdam, Germany	Nested case-control	600 (300 no diabetes/300 T2D), 6 yrs	UPLC-MS; untargeted (>4,500 metabolite features)	Serum	T2D	Age, BMI, waist circumference, physical activity, smoking, education, hypertension, alcoholic beverages, coffee, red meat, whole-grain bread*	(†) Hexose sugars (glucose, fructose, inositol) isopentenyladenosine-5'-monophosphate (purine nucleotide), hydroxymethylbutenyl diphosphate, PC (22:47:dm18:0)/PC(O-18:0/22:5), PC(38:4)/PC(36:1)/PE(O-18:0/O-18:0)/PE(O-20:0/O-16:0)	None	4.5
Zhao et al., 2015, Diabetes Care (52)	SHFS, U.S. (Native American)	Nested case-control	431 (298 no diabetes/133 T2D), 5.5 yrs	LC-MS; untargeted (>11,000 metabolite features)	Plasma	T2D	Age, sex, center, BMI, eGFR, HDL-C, TG, fasting glucose, HOMA-IR*	Lipids: 7 metabolites (5 known, 2 unknown); (†) 2-hydroxybiphenyl, an unknown metabolite, and the combination of these metabolites; (‡) PC (22:6/20:4), (3S)-7-hydroxy-2,3,4,5,8-pentamethoxy-isoflavan, or Met-Glu-Ile-Arg and Leu-Asp-Tyr-Arg tetrapeptides and an unknown metabolite and the combination of these metabolites	None	5
Fizelova et al., 2015, Atherosclerosis (53)	METSIM, Finland	Cohort, prospective, population based	6,607 men (6,221 no diabetes/386 T2D), 5 yrs	NMR; targeted (lipoproteins and apolipoprotein)	Serum	Prediabetes (glucose AUC)	Age, BMI, smoking, physical activity	(†) ApoA1, ApoB, LDL, small LDL particles, small VLDL particles, large VLDL particles, ApoA1-to-HDL-C ratio, ApoB-to-LDL-C ratio, ApoB-to-non-HDL-C ratio; (‡) HDL-C, large HDL particles (†) Small VLDL particles, large VLDL particles, ApoA1-to-HDL-C ratio, ApoB-to-non-HDL-C ratio, ApoB-to-non-HDL-C ratio; (‡) HDL-C, large HDL particles	None	3.5
Tiffin et al., 2015, Diabetologia (54)	SABRE, England (South Asian and European)	Cohort, prospective, population based	1,007 South Asian men (780 no diabetes/227 T2D) and 801 European men (688 no diabetes/113 T2D), 19 yrs	NMR; targeted (9 amino acids)	Serum	T2D	Age, waist-to-height ratio, truncal skinfold thickness, Matsuda-IR, HDL, smoking, alcohol consumption, fasting glucose, BMI	BCAAs, aromatic amino acids in men; (-) nonsignificant associations for amino acids in European men; (†) isoleucine, leucine, valine, phenylalanine, tyrosine, alanine, combination of isoleucine, leucine, and tyrosine and combination of BCAAs and aromatic amino acids in South Asians	Cross-sectional substudy	4.5

(†), positive association (eg, higher metabolite, higher risk); (-), no association; (‡), inverse association (eg, lower metabolite, lower risk) with prediabetes traits or type 2 diabetes. Apo, apolipoprotein; AUC, area under the curve; CE, cholesterol; DAG, diacylglycerols; FFAs, free fatty acids; FHS, Framingham Heart Study; FIA, flow-injection analysis; α -HB, α -hydroxybutyrate; HDL-C, HDL cholesterol; HILIC, hydrophilic interaction LC; HPLC, high-performance LC; IGT, impaired glucose tolerance; ISI, insulin sensitivity index; LDL-C, LDL cholesterol; LPC, lysophosphatidylcholines; Matsuda-IR, Matsuda index of insulin resistance; MDC, Malmö Diet and Cancer; METSIM, Metabolic Syndrome in Men; MUFA, monounsaturated fatty acids; NGT, normal glucose tolerance; NR, nonreported; PC, phosphatidylcholine; SABRE, Southall And Brent Revisited; SFAs, saturated fatty acids; SHS, Strong Heart Family Study; SM, sphingomyelin; T2D, type 2 diabetes; TG, triglycerides; TUF, Tubingen Family; UHPLC, ultra-high-performance LC; UPLC-MS, ultra-performance LC coupled with MS; WUM, Weight Loss Maintenance; yrs, years. * Analyses were corrected for multiple comparison. † Study quality was assessed using 6 criteria (up to 1 point per criterion), including participation (1 point if characteristics of the study population were described and inclusion and exclusion criteria reported and if there was a record of sampling, period, and location of recruitment), attrition (1 point if completeness of follow-up described and adequate), exposure characteristics (1 point if description of metabolomics method and metabolites analyzed), validated outcome (1 point if criteria for diabetes/prediabetes were defined), control of confounding (1 point if the models were adjusted at least for age, sex, BMI, and fasting glucose), and analysis (1 point if risk estimate determination and statistical approaches were appropriate for study design and multiple comparison testing was applied). Scores were summed; studies with scores from 0 to 3 and 4 to 6 were considered to be of lower and higher quality, respectively.

Another study identified a novel predictor of type 2 diabetes, 2-aminoadipic acid (2-AAA) (49). In a nested case control of the Framingham Offspring Study, individuals in the top quartile of 2-AAA concentrations had 4.5-fold higher risk of type 2 diabetes independently of fasting glucose after 12 years (49); this was replicated in the Malmö Cancer and Diet Study, with a follow-up of 13 years, with similar results of fourfold higher risk of diabetes (49). Moreover, these findings were validated in mice and tissue experiments in which 2-AAA modulated glucose homeostasis in vivo, and, further, in vitro experiments indicated that 2-AAA had an effect on insulin secretion in pancreatic β -cells and isolated islets (49).

Meta-analyses of Amino Acids. Eight studies examining the prospective association between several amino acids and type 2 diabetes were included in meta-analyses (5,6,12,13,45,46,50,54). Figure 2 shows the pooled RR for incident type 2 diabetes per study-specific SD difference in the amino acid analyzed. Forest plots for each individual amino acid meta-analysis are shown in Supplementary Fig. 1. The pooled RRs for incident type 2 diabetes per study-specific SD difference in

the BCAAs isoleucine, leucine, and valine were 1.36 (95% CI 1.24–1.48), 1.36 (1.17–1.58), and 1.35 (1.19–1.53), respectively, with no significant evidence of between-study heterogeneity for isoleucine and leucine ($I^2 = 9.5%$, P for heterogeneity = 0.36, and $I^2 = 37.4%$, P for heterogeneity = 0.172, respectively) and moderate heterogeneity for valine ($I^2 = 45.8%$; P for heterogeneity = 0.086) (Supplementary Fig. 1).

Aromatic amino acids, tyrosine and phenylalanine (per study-specific SD increment), were also significantly associated with higher risk of type 2 diabetes in pooled analyses (RR 1.36 [95% CI 1.19–1.55], $I^2 = 51.6%$, P for heterogeneity = 0.053) and (1.26 [1.10–1.44], $I^2 = 56%$, P for heterogeneity = 0.034) (Fig. 2), respectively. The pooled analyses for isoleucine, phenylalanine, and tyrosine combined resulted in 43% greater risk of type 2 diabetes (1.43 [1.13–1.80]), but the heterogeneity was high ($I^2 = 74.5%$, P for heterogeneity = 0.003). Glycine and glutamine were associated with 11 and 15% lower risk of diabetes, respectively, with no evidence of heterogeneity (0.89 [0.81–0.96] and 0.85 [0.82–0.89]), respectively, both $I^2 = 0.0%$, both P for heterogeneity >0.6) (Fig. 1).

The pooled RRs for incident type 2 diabetes per study-specific SD difference in alanine and histidine were 1.19 (95% CI 0.99–1.42) and 0.98 (0.91–1.06), respectively, with no significant evidence of between-study heterogeneity ($I^2 = 34.8%$ and $I^2 = 11.3%$, respectively, both P for heterogeneity >0.2) (Supplementary Fig. 1). Arginine, ornithine, and methionine were also associated with a higher risk of type 2 diabetes, although two of the three estimates were unadjusted for confounding factors (arginine 1.19 [1.14–1.25], $I^2 = 9.9%$, P for heterogeneity = 0.330; ornithine 1.10 [1.05–1.15], $I^2 = 84.3%$, P for heterogeneity = 0.002; and methionine 1.45 [1.38–1.52], $I^2 = 92.7%$, P for heterogeneity <0.001). Serine was not significantly associated with type 2 diabetes (0.97 [0.91–1.03], $I^2 = 0.0%$, P for heterogeneity = 0.527) (data not shown).

For all metabolites, findings were similar in secondary analyses using a fixed-effects approach and also when including the unadjusted estimates converted from means. All results, except for histidine, were also consistent across subgroup analyses by study design, sample analyzed, and technique, with all showing significant positive associations

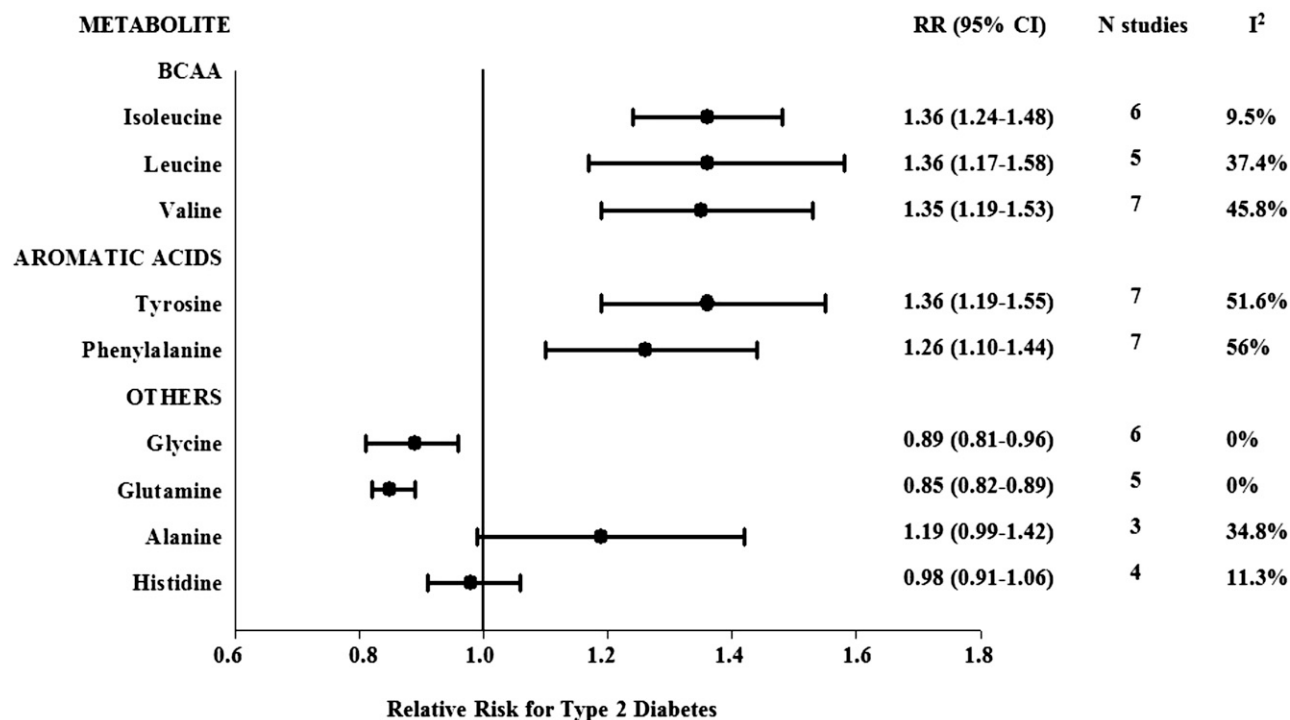


Figure 2—Pooled estimates of type 2 diabetes risk associated per study-specific SD difference in each amino acid from prospective studies. Overall estimates obtained from forest plots and random-effects meta-analysis of studies evaluating BCAAs and other amino acids and incidence of type 2 diabetes. Estimates were derived from the most fully adjusted model in each included analysis. Closed circles and horizontal bars represent the overall estimates and 95% CIs.

between these amino acids and type 2 diabetes. For histidine, results differed by study design (significant for case-control studies but not for population-based studies) and biological sample (significant for plasma but not for serum). Meta-regressions did not show significant differences in effect estimates by sex for any of the metabolites analyzed (all meta-regression $P > 0.10$). However, the power to detect heterogeneity across any of these factors may be limited given to the relatively small number of studies. Finally, the Egger and Begg tests did not indicate the presence of publication bias. Visual inspections of the funnel plots were in agreement with the statistical test, with no apparent asymmetry.

Lipid Metabolites and Acylcarnitines

The relationships between lipid metabolites and risk of prediabetes/type 2 diabetes were evaluated in 16 prospective studies (Table 1). Most of the studies found significant associations between at least some of the lipid metabolites and prediabetes/type 2 diabetes; however, because of the wide range of lipid metabolites, the results differed across studies. No single lipid metabolite had a reported estimate of association with type 2 diabetes in at least three studies. Therefore, according to our pre-specified criterion of at least three estimates for a given metabolite required for conducting meta-analysis, it was not feasible to meta-analyze any of the lipid metabolites. The great majority of the studies analyzing lipid metabolites applied targeted approaches, except for two that were untargeted (51,52). On average, 40–50 lipid metabolites were targeted. The most common analytic technique was MS, and the most commonly analyzed lipid classes were triglycerides, phosphatidylcholines, and sphingomyelins. As an example, findings from the Framingham cohort in which participants were followed for 12 years revealed that out of >100 lipid metabolites analyzed, a total of 15 were associated with incident diabetes (9). Authors concluded that lipids of lower carbon number and lower double-bond content, especially triacylglycerols (TAGs), were associated with higher risk of prediabetes and type 2 diabetes, while those with higher carbon number and more double bonds were associated with lower risk (9).

Further, individuals in the upper quartile of the combination of the strongest positive and negative predictors, TAG 50:0 and TAG 58:0, respectively, had 4.30-fold higher risk of type 2 diabetes compared with the lowest quartile, independently of fasting insulin and glucose (9). Baseline lower levels of LPC 18:2 predicted impaired glucose tolerance and type 2 diabetes after 7 years of follow-up in the KORA study (45). This study also reported that the combination of higher acetylcarnitine C2 with lower LPC 18:2 and glycine were predictors of diabetes (45). Ferrannini et al. (13) observed that lineoleoylglycerophosphocholine (L-GPC) was inversely associated with the risk of incident dysglycemia in the Relationship between Insulin Sensitivity and Cardiovascular disease (RISC) study (OR per SD 0.64 [95% CI 0.48–0.85]) and the risk of type 2 diabetes in the Botnia study (0.67 [0.54–0.84]) after adjustment for potential confounders including fasting glucose. In a case-cohort study in the framework of the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study, replicated in KORA, diacyl-phosphatidylcholines (C32:1, C36:1, C38:3, and C40:5) were associated with higher risk of type 2 diabetes, while sphingomyelin and acylalkyl-phosphatidylcholine C18:2 were associated with lower risk (6). Other relevant findings include the association of α -hydroxybutyrate with incident dysglycemia in the RISC study (1.25 [1.00–1.60]) and incident type 2 diabetes in the Botnia study (1.26 [1.07–1.48]) (13). With respect to acylcarnitines, five targeted and two untargeted analyses investigated some of these metabolites; three analyses found significant associations with prediabetes/type 2 diabetes (6,45,50). In IRAS, of the 45 acylcarnitines analyzed, greater concentrations of carnitine C4-OH and C10:3 and lower concentrations of carnitine C5:1 and C20 were associated with insulin resistance (50).

Carbohydrate Metabolites

Seven of the targeted and two of the untargeted analyses included sugar metabolites or gluconeogenesis substrates (Table 1). Of these, five studies reported significant associations for different sugar metabolites (6–8,51). In a nested case-cohort study conducted in EPIC-Potsdam,

in which 163 total metabolites were analyzed (34 were associated with type 2 diabetes), principal components analysis identified that a factor including hexose (along with diacyl-phosphatidylcholines, BCAAs, aromatic amino acids, and propionylcarnitine) was significantly associated with 3.82 times the risk of diabetes. In replication analyses presented in the same article, higher concentration of hexose was associated with type 2 diabetes in EPIC-Potsdam and in the replication in KORA (6). In a second study, a separate case-control study nested in EPIC-Potsdam that used an untargeted approach, hexose sugars (e.g., glucose, fructose, and inositol) were strongly associated with higher risk of type 2 diabetes. In contrast, individuals in the highest tertile of sugar alcohols and deoxyhexose sugars had substantially lower risk of type 2 diabetes (51). Finally, in the Bavarian Red Cross study, elevated concentrations of mannosamine and mannose were prospectively associated with future type 2 diabetes (7).

CONCLUSIONS

This systematic review and these meta-analyses identified a number of plasma metabolites prospectively associated with prediabetes and type 2 diabetes in humans, when analyzed using comprehensive high-throughput metabolomics techniques. Current evidence suggests that carbohydrate (glucose and fructose), lipid (phospholipids, sphingomyelins, and triglycerides), and amino acid (BCAAs, aromatic amino acids, glycine, and glutamine) metabolites not only are altered in individuals with type 2 diabetes but also exhibit significant prospective associations with prediabetes and/or type 2 diabetes. Our results from prospective studies support robust positive associations of BCAAs (leucine, isoleucine, and valine), aromatic amino acids (tyrosine and phenylalanine), and inverse associations of glycine and glutamine with incident type 2 diabetes. The present meta-analyses provide the most comprehensive analysis to date of the associations between BCAAs and other amino acids and the risk of type 2 diabetes. These results suggest that alterations in blood BCAAs and aromatic amino acids identified using metabolomics techniques may be useful in identifying novel biomarkers of type 2 diabetes.

Several potential mechanisms may underlie these observations. Circulating

amino acids may directly promote insulin resistance, possibly via disruption of insulin signaling in skeletal muscle (5,58); there could also be a decreased uptake and increased release of BCAAs from skeletal muscle due to increased protein catabolism in insulin resistance (12). Furthermore, BCAAs can modulate insulin secretion (58) and promote diabetes via hyperinsulinemia, leading to pancreatic β -cell exhaustion (5). Together, current evidence suggests that elevations of blood BCAAs may be early signals of deterioration of glycemic control and insulin sensitivity. Other specific amino acid findings, such as increased alanine and glutamate and decreased glutamine and glycine concentrations, were also identified as potential predictors of insulin resistance and type 2 diabetes. These findings suggest pathways specific to amino acid catabolism that may play an important role in the development of type 2 diabetes (58).

Lipidomics have also revealed that a number of lipids may be predictive of prediabetes and type 2 diabetes. However, inconsistent results have been reported in different studies, suggesting that it is also important to take into account the specific lipid fractions where lipids have been analyzed, and some lipid metabolites have only been reported in a single study and have not been replicated. In general, medium- and long-chain saturated and unsaturated fatty acids were reported to be elevated in prediabetes and type 2 diabetes case subjects, while short-chain fatty acids were found to be depleted in type 2 diabetes case subjects compared with healthy nondiabetic control subjects (30,37). Reduced sphingomyelin synthesis was associated with increased reactive oxygen species and reduced insulin secretion (59). Free fatty acids impair the action of insulin via mechanisms including accumulation of intracellular lipid derivatives (e.g., diacylglycerol and ceramides), oxidative stress, inflammation, and mitochondrial dysfunction (31). Some acylcarnitines have been reported to be associated with higher likelihood of prediabetes and type 2 diabetes; namely, acylcarnitine C2 was higher and propionylcarnitine was lower in type 2 diabetes case compared with control subjects (38). Carnitine-*O*-acetyltransferase is a mitochondrial matrix enzyme that produces acylcarnitine

from carnitine and acetyl-CoA; higher expression of carnitine-*O*-acetyltransferase could explain the elevated levels of acylcarnitine C2 in individuals with prediabetes or diabetes (60). Furthermore, alterations of lipids are common in prediabetes states, possibly reflecting alterations in triglyceride lipolysis, which could either contribute to or be a result of the dysregulation of glucose metabolism (30).

The hexose sugars (e.g., glucose, fructose, and inositol) were also associated with future risk of prediabetes and diabetes in prospective studies (6–8,51). These results may partly be explained by the fact that type 2 diabetes is itself typically defined by hyperglycemia (61). Other sugar metabolites, including desoxyhexose, dihexose, mannose, arabinose, fructose, and also certain glycolipids, were associated with prediabetes and/or type 2 diabetes status in cross-sectional or case-control studies (30,40). In addition, 1,5-anhydroglucitol was found to be lower in patients with type 2 diabetes (40); this metabolite is correlated with HbA_{1c} and has been previously proposed as a short-term marker of glycemic control (62).

Since lifestyle and dietary interventions have a great impact on type 2 diabetes prevention (63), future studies should focus on metabolic biomarkers that can modulate the effects of dietary intake, and vice versa, and their relation to insulin resistance and diabetes. In addition, certain biomarkers are directly derived from the digestion and gut absorption of food constituents; therefore, future studies relating metabolites derived from gut microbiota and type 2 diabetes may be of interest. The applications of these new approaches will contribute to future progress in the field of molecular nutritional epidemiology (64).

Several limitations of our study should be acknowledged. Although our search strategy was not limited to English, we may nevertheless not have identified studies that were unpublished or published in languages or in journals not indexed in PubMed or EMBASE; however, the main scientific journals are published in English and indexed in these databases. Misclassification of the studies and publication bias should also be acknowledged as a potential limitation. To avoid this, we had three authors independently review the studies. A meta-analysis of observational data cannot fully control for

residual or unmeasured confounding. Nevertheless, we included adjusted estimates from multivariable models from each contributing study. In contrast to amino acid metabolites, meta-analyses of lipid and carbohydrate metabolites were not feasible because too few estimates were reported for any given metabolite. Thus, conclusions regarding these metabolites are drawn from qualitative review and should be interpreted with caution.

The strengths of the current study include that our data were derived from a systematic and comprehensive approach, thus minimizing the possibility that any major published report was missed. The inclusion of meta-analyses of BCAAs and amino acids predictive of type 2 diabetes provides the best available evidence to date regarding these associations. The results of our meta-analyses were consistent regardless of whether random-effects or fixed-effect models were used and also in subgroup analyses intended to identify potential sources of heterogeneity. There was no evidence to suggest publication bias after exclusion of the single study that accounted for the heterogeneity of some of our results. However, other potential sources of heterogeneity, some difficult to ascertain, could have affected the results (65). For example, the use of fasting versus nonfasting blood; its appropriate collection, handling, and storage; how diabetes was ascertained; and the quality-control calibration conducted for metabolomics analysis may differ between studies. Future studies with higher statistical power to assess sources of heterogeneity need to account for these factors.

Summary

Taken together, the present systematic review provides evidence that several metabolites, including BCAAs and aromatic amino acids, hexoses, phospholipids, and triglycerides, are associated with the risk of prediabetes and type 2 diabetes. Our data indicate that the BCAAs isoleucine, leucine, and valine and the aromatic amino acids tyrosine and phenylalanine are positively associated with the risk of type 2 diabetes, whereas glycine and glutamine are inversely associated with diabetes. Further research is needed on the metabolic biomarkers that are modulated by dietary

intake and gut microbiota and their relationships with insulin resistance and diabetes. Use of these and other new approaches may contribute to future progress in the field of diabetes, molecular nutritional epidemiology, and personalized medicine.

Funding. This work was supported by National Institutes of Health research grants R01-DK-102896, R01-HL-118264, P30-DK-46200, and HL-60712 and an American Diabetes Association Mentor-Based Postdoctoral Fellowship.

The funding sources played no role in the design, collection, analysis, or interpretation of the data or in the decision to submit the manuscript for publication.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. M.G.-F., E.T., M.A.M.-G., J.S.-S., and F.B.H. designed the research. M.G.-F., A.H., E.T., C.B.C., M.A.M.-G., J.S.-S., and F.B.H. conducted the research. M.G.-F., A.H., and E.T. analyzed data. M.G.-F. drafted the manuscript. M.G.-F., A.H., E.T., C.B.C., M.A.M.-G., J.S.-S., and F.B.H. made critical revisions to the manuscript for important intellectual content. E.T., C.B.C., M.A.M.-G., J.S.-S., and F.B.H. obtained funding. All authors read and approved the final manuscript.

References

- International Diabetes Federation. *IDF Diabetes Atlas, 6th ed.* Brussels, Belgium, International Diabetes Federation, 2013
- Hu FB, Satija A, Manson JE. Curbing the diabetes pandemic: the need for global policy solutions. *JAMA* 2015;313:2319–2320
- Roberts LD, Koulman A, Griffin JL. Towards metabolic biomarkers of insulin resistance and type 2 diabetes: progress from the metabolome. *Lancet Diabetes Endocrinol* 2014;2:65–75
- Nicholson JK, Lindon JC. Systems biology: Metabonomics. *Nature* 2008;455:1054–1056
- Wang TJ, Larson MG, Vasani RS, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med* 2011;17:448–453
- 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;62:639–648
- Padberg I, Peter E, González-Maldonado S, et al. A new metabolomic signature in type-2 diabetes mellitus and its pathophysiology. *PLoS One* 2014;9:e85082
- Würtz P, Tiiainen M, Mäkinen V-P, et al. Circulating metabolite predictors of glycemia in middle-aged men and women. *Diabetes Care* 2012;35:1749–1756
- 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;121:1402–1411
- American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2011;34(Suppl. 1):S62–S69
- Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006;144:427–437
- Stancáková 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;61:1895–1902
- Ferrannini E, Natali A, Camastra S, et al. Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. *Diabetes* 2013;62:1730–1737
- Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psychol Bull* 1995;117:167–178
- da Costa BR, Rutjes AWS, Johnston BC, et al. Methods to convert continuous outcomes into odds ratios of treatment response and numbers needed to treat: meta-epidemiological study. *Int J Epidemiol* 2012;41:1445–1459
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–634
- Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088–1101
- 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;53:757–767
- 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;9:311–326
- Lucio M, Fekete A, Weigert C, et al. Insulin sensitivity is reflected by characteristic metabolic fingerprints—a Fourier transform mass spectrometric non-targeted metabolomics approach. *PLoS One* 2010;5:e13317
- Gall WE, Beebe K, Lawton KA, et al.; RISC Study Group. Alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS One* 2010;5:e10883
- Würtz P, Mäkinen V-P, Soininen P, et al. Metabolic signatures of insulin resistance in 7,098 young adults. *Diabetes* 2012;61:1372–1380
- Ho JE, Larson MG, Vasani RS, et al. Metabolite profiles during oral glucose challenge. *Diabetes* 2013;62:2689–2698
- Geidenstam N, Spégel P, Mulder H, Filipsson K, Ridderstråle M, Danielsson APH. Metabolite profile deviations in an oral glucose tolerance test—a comparison between lean and obese individuals. *Obesity (Silver Spring)* 2014;22:2388–2395
- Haus JM, Kashyap SR, Kasumov T, et al. Plasma ceramides are elevated in obese subjects with type 2 diabetes and correlate with the severity of insulin resistance. *Diabetes* 2009;58:337–343
- Zhang X, Wang Y, Hao F, et al. Human serum metabolomic analysis reveals progression axes for glucose intolerance and insulin resistance statuses. *J Proteome Res* 2009;8:5188–5195
- 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;18:1695–1700
- Fiehn O, Garvey WT, Newman JW, Lok KH, Hoppel CL, Adams SH. Plasma metabolomic profiles reflective of glucose homeostasis in non-diabetic and type 2 diabetic obese African-American women. *PLoS One* 2010;5:e15234
- Zhou Y, Qiu L, Xiao Q, et al. Obesity and diabetes related plasma amino acid alterations. *Clin Biochem* 2013;46:1447–1452
- Menni C, Fauman E, Erte I, et al. Biomarkers for type 2 diabetes and impaired fasting glucose using a nontargeted metabolomics approach. *Diabetes* 2013;62:4270–4276
- Xu F, Tavintharan S, Sum CF, Woon K, Lim SC, Ong CN. Metabolic signature shift in type 2 diabetes mellitus revealed by mass spectrometry-based metabolomics. *J Clin Endocrinol Metab* 2013;98:E1060–E1065
- Meikle PJ, Wong G, Barlow CK, et al. Plasma lipid profiling shows similar associations with prediabetes and type 2 diabetes. *PLoS One* 2013;8:e74341
- Thalacker-Mercer AE, Ingram KH, Guo F, Ilkayeva O, Newgard CB, Garvey WT. BMI, RQ, diabetes, and sex affect the relationships between amino acids and clamp measures of insulin action in humans. *Diabetes* 2014;63:791–800
- Messana I, Forni F, Ferrari F, Rossi C, Giardina B, Zuppi C. Proton nuclear magnetic resonance spectral profiles of urine in type II diabetic patients. *Clin Chem* 1998;44:1529–1534
- Wang C, Kong H, Guan Y, et al. Plasma phospholipid metabolic profiling and biomarkers of type 2 diabetes mellitus based on high-performance liquid chromatography/electrospray mass spectrometry and multivariate statistical analysis. *Anal Chem* 2005;77:4108–4116
- Yuan K, Kong H, Guan Y, Yang J, Xu G. A GC-based metabolomics investigation of type 2 diabetes by organic acids metabolic profile. *J Chromatogr B Anal Technol Biomed Life Sci* 2007;850:236–240
- Li X, Xu Z, Lu X, et al. Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabolomics: Biomarker discovery for diabetes mellitus. *Anal Chim Acta* 2009;633:257–262
- 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;139:1073–1081
- Zhang J, Yan L, Chen W, et al. Metabolomics research of diabetic nephropathy and type 2 diabetes mellitus based on UPLC-oeTOF-MS system. *Anal Chim Acta* 2009;650:16–22
- Suhre K, Meisinger C, Döring A, et al. Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. *PLoS One* 2010;5:e13953
- Han L-D, Xia J-F, Liang Q-L, et al. Plasma esterified and non-esterified fatty acids metabolic profiling using gas chromatography-mass spectrometry and its application in the study of diabetic mellitus and diabetic nephropathy. *Anal Chim Acta* 2011;689:85–91
- Ha CY, Kim JY, Paik JK, et al. The association of specific metabolites of lipid metabolism with markers of oxidative stress, inflammation and arterial stiffness in men with newly diagnosed

- type 2 diabetes. *Clin Endocrinol (Oxf)* 2012;76:674–682
43. Kaur P, Rizk N, Ibrahim S, et al. Quantitative metabolomic and lipidomic profiling reveals aberrant amino acid metabolism in type 2 diabetes. *Mol Biosyst* 2013;9:307–317
44. Zhang AH, Sun H, Yan GL, Yuan Y, Han Y, Wang XJ. Metabolomics study of type 2 diabetes using ultra-performance LC-ESI/quadrupole-TOF high-definition MS coupled with pattern recognition methods. *J Physiol Biochem* 2014;70:117–128
45. Wang-Sattler R, Yu Z, Herder C, et al. Novel biomarkers for pre-diabetes identified by metabolomics. *Mol Syst Biol* 2012;8:615
46. Cheng S, Rhee EP, Larson MG, et al. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation* 2012;125:2222–2231
47. Mahendran Y, Cederberg H, Vangipurapu J, et al. Glycerol and fatty acids in serum predict the development of hyperglycemia and type 2 diabetes in Finnish men. *Diabetes Care* 2013;36:3732–3738
48. Mahendran Y, Vangipurapu J, Cederberg H, et al. Association of ketone body levels with hyperglycemia and type 2 diabetes in 9,398 Finnish men. *Diabetes* 2013;62:3618–3626
49. Wang TJ, Ngo D, Psychogios N, et al. 2-Aminoadipic acid is a biomarker for diabetes risk. *J Clin Invest* 2013;123:4309–4317
50. 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–E468
51. 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;61:487–497
52. Zhao J, Zhu Y, Hyun N, et al. Novel metabolic markers for the risk of diabetes development in American Indians. *Diabetes Care* 2015;38:220–227
53. Fizeleva M, Miilunpohja M, Kangas AJ, et al. Associations of multiple lipoprotein and apolipoprotein measures with worsening of glycemia and incident type 2 diabetes in 6607 non-diabetic Finnish men. *Atherosclerosis* 2015;240:272–277
54. 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;58:968–979
55. Shah SH, Crosslin DR, Haynes CS, et al. Branched-chain amino acid levels are associated with improvement in insulin resistance with weight loss. *Diabetologia* 2012;55:321–330
56. Würtz P, Soinen P, Kangas AJ, et al. Branched-chain and aromatic amino acids are predictors of insulin resistance in young adults. *Diabetes Care* 2013;36:648–655
57. Cobb J, Gall W, Adam K-P, et al. A novel fasting blood test for insulin resistance and prediabetes. *J Diabetes Sci Technol* 2013;7:100–110
58. Krebs M, Krssak M, Bernroider E, et al. Mechanism of amino acid-induced skeletal muscle insulin resistance in humans. *Diabetes* 2002;51:599–605
59. Yano M, Watanabe K, Yamamoto T, et al. Mitochondrial dysfunction and increased reactive oxygen species impair insulin secretion in sphingomyelin synthase 1-null mice. *J Biol Chem* 2011;286:3992–4002
60. Horie S, Ishii H, Suga T. Changes in peroxisomal fatty acid oxidation in the diabetic rat liver. *J Biochem* 1981;90:1691–1696
61. Kahn BB. Type 2 diabetes: when insulin secretion fails to compensate for insulin resistance. *Cell* 1998;92:593–596
62. McGill JB, Cole TG, Nowatzke W, et al.; U.S. trial of the GlycoMark assay. Circulating 1,5-anhydroglucitol levels in adult patients with diabetes reflect longitudinal changes of glycemia: a U.S. trial of the GlycoMark assay. *Diabetes Care* 2004;27:1859–1865
63. Ley SH, Hamdy O, Mohan V, Hu FB. Prevention and management of type 2 diabetes: dietary components and nutritional strategies. *Lancet* 2014;383:1999–2007
64. Cornelis MC, Hu FB. Systems epidemiology: a new direction in nutrition and metabolic disease research. *Curr Nutr Rep* 2013;2
65. Tzoulaki I, Ebbels TMD, Valdes A, Elliott P, Ioannidis JPA. Design and analysis of metabolomics studies in epidemiologic research: a primer on -omic technologies. *Am J Epidemiol* 2014;180:129–139