Identification of Serum Metabolites Associated With Risk of Type 2 Diabetes Using a Targeted Metabolomic Approach

Anna Floegel,[1] Norbert Stefan,[2] Zhonghao Yu,[3] Kristin Mühlenbruch,[4] Dagmar Drogan,[1] Hans-Georg Joost,[5] Andreas Fritsche,[2] Hans-Ulrich Häring,[2] Martin Hrabě de Angelis,[6] Annette Peters,[7] Michael Roden,[8,9] Cornelia Prehn,[6] Rui Wang-Sattler,[3] Thomas Illig,[3,10] Matthias B. Schulze,[4] Jerzy Adamski,[6] Heiner Boeing,[1] and Tobias Pischon[1,11]

Metabolomic discovery of biomarkers of type 2 diabetes (T2D) risk may reveal etiological pathways and help to identify individuals at risk for disease. We prospectively investigated the association between serum metabolites measured by targeted metabolomics and risk of T2D in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam (27,548 adults) among all incident cases of T2D ($n = 800$, mean follow-up 7 years) and a randomly drawn subcohort ($n = 2,282$). Flow injection analysis tandem mass spectrometry was used to quantify 163 metabolites, including acylcarnitines, amino acids, hexose, and phospholipids, in baseline serum samples. Serum hexose; phenylalanine; and diacyl-phosphatidylcholines C32:1, C36:1, C38:3, and C40:5 were independently associated with increased risk of T2D and serum glycine; sphingomyelin C16:1; acyl-alkyl-phosphatidylcholines C34:3, C40:6, C42:5, C44:4, and C44:5; and lysophosphatidylcholine C18:2 with decreased risk. Variance of the metabolites was largely explained by two metabolite factors with opposing risk associations (factor 1 relative risk in extreme quintiles 0.31 [95% CI 0.21–0.44], factor 2 3.82 [2.64–5.52]). The metabolites significantly improved T2D prediction compared with established risk factors. They were further linked to insulin sensitivity and secretion in the Tübingen Family study and were partly replicated in the independent KORA (Cooperative Health Research in the Region of Augsburg) cohort. The data indicate that metabolic alterations, including sugar metabolites, amino acids, and choline-containing phospholipids, are associated early on with a higher risk of T2D. *Diabetes* 62:639–648, 2013

From the [1]Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; the [2]Department of Internal Medicine IV, Divisions of Endocrinology, Diabetology, Nephrology, Vascular Disease, and Clinical Chemistry, University of Tübingen, Tübingen, Germany; the [3]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; the [4]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; the [5]Department of Pharmacology, German Institute of Human Nutrition Potsdam–Rehbruecke, Nuthetal, Germany; the [6]Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; the [7]Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; the [8]Institute of Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University, Düsseldorf, Germany; the [9]Department of Metabolic Diseases, University Clinics, Düsseldorf, Germany; the [10]Hannover Unified Biobank, Hannover Medical School, Hannover, Germany; and the [11]Molecular Epidemiology Group, Max Delbrück Center for Molecular Medicine (MDC) Berlin-Buch, Berlin, Germany.
Corresponding author: Anna Floegel, anna.floegel@dife.de.

Type 2 diabetes (T2D) is characterized by impaired insulin sensitivity of several tissues and inadequate insulin secretion from β-cells (1). A detailed understanding of the pathophysiology of T2D is a prerequisite for the development of preventive strategies. In particular, the identification of early metabolic alterations is promising in the study of etiological pathways and may further help to identify high-risk individuals. A number of biomarkers have been proposed as indicators for the estimation of T2D risk, such as fasting plasma glucose and glycated hemoglobin A_{1c} (HbA$_{1c}$) (2), triglycerides (3), HDL cholesterol (4), inflammatory markers (5), adiponectin (5,6), liver enzymes (7), and fetuin-A (8). However, most biomarkers fail to grasp the complexity of T2D etiology (3). Design and advancement of high-throughput analytical techniques determined the emergence of metabolomics, which is the simultaneous study of numerous low-molecular weight compounds, namely metabolites. Metabolites represent intermediates and end products of metabolic pathways that reflect more rapidly physiological dysfunctions than current biomarkers and, thus, may mirror earlier stages of T2D (9). Cross-sectional studies have linked alterations in metabolic profiles with obesity (10), glucose tolerance (11), and prevalent diabetes (12–14). The most prominent metabolic shifts involved blood acylcarnitines and branched-chain amino acids (BCAAs). Recently, observations from a prospective study found that a set of five amino acids was predictive for T2D, representing pioneering work in the emerging field of systems epidemiology (15).

In the current study, we investigated whether a targeted metabolomic approach involving a broader spectrum of metabolites and a larger number of study participants may help to identify metabolites associated with the risk of T2D and the mechanisms involved. Therefore, we profiled 163 serum metabolites in originally healthy individuals who were consecutively followed up for incident T2D in two large-scale prospective cohort studies in Germany and studied cross-sectional relationships of the identified metabolites with insulin sensitivity and secretion in precisely phenotyped participants. In addition, we evaluated the usefulness of the metabolites for T2D risk prediction compared with the German Diabetes Risk Score (DRS) (16) and established biomarkers.

RESEARCH DESIGN AND METHODS

European Prospective Investigation into Cancer and Nutrition-Potsdam study. The European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam is part of the ongoing multicenter EPIC study and comprises 27,548 participants from the general population in the area of Potsdam in eastern Germany who were mainly 35–65 years of age at time of recruitment between 1994 and 1998 (17). At baseline, participants underwent an examination by qualified staff, including medical history, blood pressure measurement, and anthropometry (18). Participants also completed sociodemographic and lifestyle questionnaires and a validated food frequency questionnaire. In addition, 30 mL blood were drawn (random sampling) and immediately processed (19). Only participants with morning appointments were asked to fast overnight. Blood was fractionated into serum, plasma, buffy coat, and erythrocytes; aliquotted into straws of 0.5 mL each; and stored in tanks of liquid nitrogen at −196°C until analysis. Besides metabolomic profiling, other biomarkers have been measured in baseline blood samples as described previously (7,20,21). Every 2–3 years, follow-up questionnaires were sent to participants to identify incident cases of T2D, with response rates of ∼95% (22,23). Once a participant was identified as a potential case, disease status was further verified with medical records, including the correct diagnosis (International Classification of Diseases, 10th revision, E11, non–insulin-dependent diabetes), the date of the diagnosis, and the means of diagnosis confirmation. This verification was achieved by sending a standard inquiry form to the treating physician. Consent was obtained from all study participants a priori, and the study was approved by the ethics committee of the Medical Society of the State of Brandenburg.

We constructed a case-cohort study within EPIC-Potsdam, including all incident cases of T2D of the full cohort identified up to 31 August 2005 (n = 849, mean follow-up 7 years), and a subcohort (n = 2,500) randomly drawn from the EPIC-Potsdam study population. Because the subcohort was representative of the full cohort, it included 2,415 noncases and 85 cases of incident T2D (i.e., internal cases). The remaining 764 cases did not belong to the subcohort (external cases). By randomly selecting the subcohort and using the appropriate statistic for this study design, the biomarkers only needed to be measured in the case-cohort sample; however, the results are expected to be generalizable to the full cohort (7). The case-cohort design was previously chosen based on its advantages, including a reduced chance of selection bias for the control group (24).

For the present analysis, we further excluded participants with prevalent T2D at baseline (n = 110), with missing or nonverified data on incident or prevalent T2D (n = 13), with missing blood samples or biomarker measurements (n = 64), and with missing covariate information (n = 80). Thus, the analytical sample included 2,282 individuals of the subcohort and 800 individuals with incident T2D.

Cooperative Health Research in the Region of Augsburg study. The Cooperative Health Research in the Region of Augsburg (KORA) study consists of population-based surveys and follow-up periods in the area of Augsburg in southern Germany. A total of 4,261 individuals between 25 and 74 years of age participated in the S4 survey between 1999 and 2001 (25). In the KORA cohort, blood was drawn into serum gel tubes after a fasting period of at least 8 h. Blood samples were rested for coagulation for 30 min at room temperature; serum was obtained by centrifugation at 2,750g at 15°C for 10 min and stored in a freezer at −80°C. A total of 3,080 individuals took part in the F4 follow-up survey during the years 2006–2008 (26). The identification of incident T2D was based on an oral glucose tolerance test (OGTT) or a validated physician diagnosis (27). All KORA participants gave written informed consent, and the KORA study was approved by the ethics committee of the Bavarian medical association.

A subcohort of 876 S4 participants 55–74 years of age without T2D at baseline and with metabolomics data available was included in the current study. Of them, 91 developed incident T2D during the 7-year follow-up.

Tübingen Family study for T2D. The Tübingen Family (TüF) study is in an ongoing investigation of the pathophysiology of T2D in southern Germany (28). Individuals meeting at least one of the following criteria were included in the study: a family history of T2D, a BMI >27 kg/m², and previous impaired glucose tolerance or gestational diabetes mellitus. They were considered healthy according to a physical examination and routine laboratory tests. Written informed consent was obtained from all participants, and the medical ethics committee of the University of Tübingen approved the protocol.

All individuals underwent a 75-g OGTT. Venous plasma samples were drawn at 0, 30, 60, 90, and 120 min for plasma glucose, insulin, C-peptide, and metabolomic analyses (minute 0). Insulin sensitivity was calculated from the OGTT (29). The plasma glucose and C-peptide areas under the curve (AUCs) during the OGTT were calculated by applying the trapezoid method. Insulin secretion was calculated from $AUC_{C-peptide}/AUC_{glucose}$. The present analysis included 76 Caucasians from the TüF study who had measurements of insulin sensitivity and insulin secretion as well as metabolomics data available.

Serum metabolite concentrations. Serum concentrations of metabolites were determined with the Absolute*IDQ* p150 and p180 Kits (Biocrates Life Sciences AG, Innsbruck, Austria) using the flow injection analysis tandem mass spectrometry (FIA-MS/MS) technique (30). The metabolomic method simultaneously quantified 163 metabolites, including 41 acylcarnitines (Cx:y), 14 amino acids, 1 hexose (sum of six-carbon monosaccharides without distinction of isomers), 92 glycerophospholipids (lyso-, diacyl-, and acyl-alkyl-phosphatidylcholines), and 15 sphingomyelins. To ensure valid measurements, metabolites below the limit of detection (n = 30) and those with very high analytical variance (n = 6) in our samples were excluded, leaving 127 metabolites for the present analysis.

Metabolomic measurements were performed in the Genome Analysis Center at the Helmholtz Zentum München. Sample preparation was done according to the manufacturer's protocol (Biocrates user's manual UM-P150) and has been described previously (30). In brief, after centrifugation, 10 μL serum were inserted into a filter on a 96-well sandwich plate, which already contained stable isotope-labeled internal standards. Amino acids were derivated with 5% phenylisothiocyanate reagent. Metabolites and internal standards were extracted with 5 mmol/L ammonium acetate in methanol. The solution was then centrifuged through a filter membrane and diluted with mass spectrometry running solvent. Final extracts were analyzed by FIA-MS/MS, and metabolites were quantified in μmol/L by appropriate internal standards. The method has been validated, and analytical specifications were provided in the Biocrates manual AS-P150. The manufacturer selected the metabolites based on the robustness of their measurements. The uncertainty of the measurements was <10% for most of the metabolites. Regarding accuracy, all included metabolites were found in the range of 80–115% of their theoretical values. The median analytical variance of EPIC-Potsdam samples was a 7.3% within-plate coefficient of variation and a 11.3% between-plates coefficient of variation (31). To account for run-order effects, serum samples were randomly analyzed together, regardless of the case status. We have shown previously that most of the metabolites had moderate to high intraclass correlation coefficients measured in participants over a 4-month period, indicating reasonable reliability of the measurements (31).

Statistical analysis

Step 1: Identification of metabolites associated with T2D risk in EPIC-Potsdam. Cox proportional hazards regression with weighting as suggested by Prentice (32) and robust sandwich covariance estimates to account for the case-cohort design were used to calculate multivariable-adjusted hazard ratios as a measure of relative risk (RR) and 95% CI, with age as the underlying time scale from recruitment to study exit (T2D diagnosis or censoring) of each participant. We considered z score–standardized metabolite concentrations (mean 0 [SD 1]) as the exposure variable and calculated a multivariable-adjusted model to select metabolites associated with T2D risk. This model was adjusted for age, sex, alcohol intake from beverages (nonconsumers; women >0–6, 6–12, and >12 g/day; and men >0–12, 12–24, and >24 g/day), smoking (never, former, current ≤20 cigarettes/day, current >20 cigarettes/day), cycling and sports (h/week), education (no degree/vocational training, trade/technical school, university degree), coffee intake (cups/day), red meat intake (g/day), whole-grain bread intake (g/day), prevalent hypertension (yes/no), BMI (kg/m²), and waist circumference (cm). Because the metabolomic approach is exploratory, the P values from Cox regression were corrected for multiple testing (n = 127) using the Bonferroni-Holm procedure (33), and a corrected $P < 0.05$ (two-sided testing) was considered significant to select metabolites.

We next calculated a model that included the covariates and all the identified metabolites and used stepwise Cox regression to select the independent predictors. To also account for the intercorrelation of some of the metabolites, we conducted a principal component analysis (PCA). In brief, the PCA aggregates the individual metabolites based on their degree of correlation with one another to a smaller number of metabolite factors (principal components). These metabolite factors are extracted in a way that explains the major fraction of the variance of individual metabolites. We included all metabolites associated with T2D risk in the PCA, based the PCA on the correlation matrix of metabolites, and used an orthogonal rotation procedure with the varimax method. We retained two metabolite factors according to the scree test and because they accounted for most of the observed variance. Thus, the proportion of explained variance of factor 1 and factor 2 were 34.2 and 16.1%, respectively. To investigate the association between metabolite factors and T2D risk, we estimated RR and 95% CI across quintiles of metabolite factors, particularly choosing quintiles of metabolite factors to facilitate the interpretation. We also investigated a possible effect modification of sex or fasting status on the association between metabolite factors and T2D risk by including multiplicative interaction terms into the models. We then repeated the PCA to include only fasting blood samples in order to evaluate whether the metabolite factors were different from those obtained from random blood samples. Finally, we calculated hazard rates of T2D during different periods of follow-up and tested whether they were different with a test of heterogeneity (34).

Step 2: Additional analyses, risk prediction, and replication in KORA.
For significant metabolites found in step 1, we calculated multivariable-adjusted models with additional adjustment for blood glucose, HbA_{1c}, HDL cholesterol, and triglycerides. We also adjusted for the amino acids phenylalanine, tyrosine, and isoleucine, which have been found to be associated with T2D risk (35). Using data of the EPIC-Potsdam subcohort, we calculated Spearman partial correlation coefficients between identified metabolites and established T2D biomarkers. Data of the TüF study were used to calculate Spearman partial correlation coefficients between identified metabolites and measures of insulin sensitivity and secretion. We calculated measures of discrimination and calibration in different multivariable-adjusted logistic regression models using the DRS (16) as the reference model and adding established T2D biomarkers and the identified metabolites. Receiver operating characteristic (ROC) AUCs were compared using the DeLong test (36).

The results were replicated in the prospective KORA study, and metabolite factors were recalculated in KORA using the linear factor equations retrieved from the PCA in the EPIC-Potsdam sample. The risk estimates from EPIC-Potsdam and KORA were combined using a meta-analytical approach (37). Power calculation suggested a detectable RR per SD of 1.26 (38). Additionally, multivariable-adjusted RRs of T2D were calculated for the amino acids that were recently identified by Wang et al. (35) and that were also measured in the EPIC-Potsdam study. The statistical analyses were conducted with SAS version 9.2 (SAS Institute, Inc, Cary, NC) and in the R statistical environment (www.r-project.org).

RESULTS

Baseline characteristics of the EPIC-Potsdam sample are presented in Table 1. Of all the metabolites measured using a targeted metabolomic approach, 34 were significantly associated with T2D risk in the EPIC-Potsdam study after correction for multiple testing (Supplementary Table 1). These relations were also independent of relevant dietary and lifestyle factors as well as anthropometry and hypertension. Among these 34 metabolites, 14 were identified to be significantly associated with T2D risk independent of

the others (Table 2). Specifically, hexose, phenylalanine, and diacyl-phosphatidylcholines C32:1, C36:1, C38:3, and C40:5 were significantly positively associated with T2D risk, whereas glycine, sphingomyelin C16:1, lysophosphatidylcholine C18:2, and acyl-alkyl-phosphatidylcholines C34:3, C40:6, C42:5, C44:4, and C44:5 were significantly inversely related to T2D risk. Using PCA, we identified two metabolite factors that included multiple metabolites and explained most of their variation (Fig. 1). These metabolite factors showed significant and opposing associations with T2D risk. When comparing extreme quintiles, metabolite factor 1, which mainly contains acyl-alkyl-phosphatidylcholines, sphingomyelins, and lysophosphatidylcholines, was associated with a significant 69% reduced risk of T2D (RR 0.31 [95% CI 0.21–0.44]) (Table 3), whereas metabolite factor 2, consistent of diacyl-phosphatidylcholines, BCAA and aromatic amino acids, propionylcarnitine, and hexose, was associated with a significant 3.82-fold increased risk of T2D (3.82 [2.64–5.52]). When we restricted the PCA to fasting samples ($n = 429$), very similar metabolite factors could be generated with only minor differences in factor loadings (Supplementary Table 2). We also estimated the joint effects of both metabolite factors by summing them (factor 1 received a negative sign because it was inversely associated with T2D risk) and calculating RR of T2D across quintiles of combined factors. The RR (95% CI) of T2D from quintile 1 to quintile 5 of summed metabolite factors was as follows: 1.0, 1.47 (0.96–2.24), 2.29 (1.54–3.38), 2.67 (1.79–4.0), and 6.69 (4.50–9.96) (*P* for trend <0.0001).

Adjustment for established T2D biomarkers only marginally affected the magnitude of risk association for most

TABLE 1
Baseline characteristics of the EPIC-Potsdam case-cohort sample (1994–1998)

	Subcohort ($n = 2,282$)	Incident T2D cases ($n = 800$)	*P* for difference†
Age (years)‡	49.5 (8.9)	54.7 (7.3)	<0.0001
Women (%)‡	62.0	42.2	<0.0001
BMI (kg/m²)	26.1 (0.09)	30.1 (0.15)	<0.0001
Waist circumference, men (cm)§	93.7 (0.34)	103.6 (0.46)	<0.0001
Waist circumference, women (cm)§	80.6 (0.30)	93.4 (0.62)	<0.0001
Prevalent hypertension	49.5	70.8	<0.0001
Education			
No degree/vocational training	37.1	45.6	<0.0001
Trade/technical school	24.0	25.4	0.412
University degree	39.0	29.0	<0.0001
Smoking status			
Never	46.9	36.2	<0.0001
Former	33.0	42.3	<0.0001
Current	20.1	21.5	0.472
Among smokers, cigarettes/day	12.6 (0.43)	16.0 (0.74)	<0.0001
Physical activity (h/week)¶	2.8 (0.07)	2.2 (0.13)	<0.0001
Alcohol intake from beverages (g/day)	14.8 (0.41)	14.5 (0.71)	0.761
Coffee consumption (cups/day)	2.8 (0.04)	2.7 (0.08)	0.148
Whole-grain bread intake (g/day)	45.9 (1.11)	38.2 (1.91)	0.0003
Red meat intake (g/day)	43.3 (0.61)	48.9 (1.06)	<0.0001
Biomarkers			
Glucose (mg/dL)	88.1 (0.53)	107.0 (0.92)	<0.0001
HbA_{1c} (%)‖	5.42 (0.01)	6.30 (0.03)	<0.0001
HDL cholesterol (mg/dL)	47.5 (0.25)	40.9 (0.43)	<0.0001
Triglycerides (mg/dL)	114.8 (2.12)	177.2 (3.65)	<0.0001

Data are age- and sex-adjusted mean (SE) for continuous variables or % for categorical variables. †*P* for difference comparing incident cases of T2D with noncases of the subcohort. ‡Unadjusted mean (SD) or %. §Age-adjusted mean (SE). ¶Average of cycling and sports during summer and winter seasons. ‖Data were only available for $n = 2,900$.

TABLE 2
Mean serum concentrations of identified metabolites and their association with risk of T2D in EPIC-Potsdam

	Subcohort (n = 2,282) (μmol/L)	Cases (n = 800) (μmol/L)	RR per SD (95% CI)					
			Basic model[†]	+ Glucose[‡]	+ HbA$_{1c}$[‡]	+ HDL chol[‡]	+ Triglycerides[‡]	+ Glucose + HbA$_{1c}$ + HDL chol + Triglycerides[‡]
Hexose	4,698 (984)	5,783 (2,167)	2.36 (2.06–2.71)	1.62 (1.37–1.91)	2.11 (1.82–2.44)	2.43 (2.11–2.79)	2.34 (2.04–2.69)	1.51 (1.28–1.80)
Phenylalanine	56.4 (12.0)	61.0 (12.9)	1.35 (1.22–1.49)	1.24 (1.12–1.37)	1.37 (1.23–1.52)	1.34 (1.21–1.49)	1.30 (1.16–1.44)	1.25 (1.12–1.39)
Glycine	256 (77.2)	227 (59.6)	0.73 (0.64–0.83)	0.81 (0.71–0.92)	0.81 (0.71–0.93)	0.74 (0.65–0.84)	0.74 (0.65–0.84)	0.89 (0.78–1.01)
SM C16:1	17.6 (3.86)	17.4 (4.02)	0.80 (0.72–0.88)	0.82 (0.73–0.92)	0.74 (0.67–0.83)	0.87 (0.78–0.97)	0.79 (0.71–0.88)	0.91 (0.80–1.03)
PC ae C34:3	8.51 (2.48)	7.01 (2.02)	0.64 (0.56–0.72)	0.70 (0.61–0.80)	0.67 (0.58–0.77)	0.72 (0.63–0.82)	0.66 (0.58–0.74)	0.92 (0.78–1.07)
PC ae C40:6	5.52 (1.43)	5.03 (1.34)	0.77 (0.69–0.86)	0.81 (0.72–0.92)	0.83 (0.74–0.94)	0.82 (0.73–0.92)	0.76 (0.67–0.85)	0.94 (0.82–1.08)
PC ae C42:5	2.35 (0.54)	2.06 (0.46)	0.71 (0.63–0.79)	0.76 (0.67–0.86)	0.82 (0.72–0.92)	0.76 (0.68–0.86)	0.73 (0.64–0.82)	0.96 (0.83–1.12)
PC ae C44:4	0.38 (0.10)	0.33 (0.09)	0.76 (0.67–0.85)	0.77 (0.68–0.87)	0.84 (0.75–0.96)	0.82 (0.72–0.93)	0.77 (0.68–0.87)	0.97 (0.84–1.12)
PC ae C44:5	1.77 (0.48)	1.53 (0.43)	0.71 (0.63–0.80)	0.74 (0.65–0.84)	0.81 (0.71–0.91)	0.77 (0.68–0.87)	0.74 (0.65–0.83)	0.95 (0.82–1.09)
PC aa C32:1	16.9 (9.76)	20.4 (11.3)	1.18 (1.08–1.29)	1.12 (1.02–1.24)	1.12 (1.02–1.23)	1.23 (1.11–1.36)	1.12 (1.02–1.22)	1.15 (1.03–1.29)
PC aa C36:1	57.3 (15.6)	62.1 (17.2)	1.18 (1.08–1.30)	1.18 (1.07–1.30)	1.09 (0.98–1.21)	1.26 (1.14–1.39)	1.07 (0.96–1.19)	1.25 (1.10–1.41)
PC aa C38:3	55.4 (14.8)	65.3 (18.2)	1.37 (1.24–1.51)	1.35 (1.22–1.50)	1.31 (1.18–1.45)	1.38 (1.25–1.52)	1.24 (1.11–1.38)	1.38 (1.22–1.55)
PC aa C40:5	11.4 (3.49)	12.8 (4.17)	1.18 (1.08–1.29)	1.21 (1.10–1.33)	1.12 (1.02–1.24)	1.21 (1.10–1.32)	1.07 (0.96–1.18)	1.19 (1.06–1.32)
LysoPC a C18:2	35.2 (13.7)	29.6 (11.1)	0.74 (0.66–0.84)	0.72 (0.64–0.81)	0.77 (0.68–0.88)	0.79 (0.70–0.90)	0.72 (0.63–0.81)	0.84 (0.73–0.96)

Data are mean (SD) metabolite concentrations. a, acyl; aa, diacyl; ae, acyl-alkyl; chol, cholesterol; PC, phosphatidylcholine; SM, sphingomyelin. †Metabolites were selected by stepwise Cox proportional hazard regression modified by the Prentice method after standardizing metabolite concentrations to a mean of 0 (SD 1) and including all metabolites associated with T2D risk as well as relevant covariates. RRs for the basic model were then calculated in a continuous model adjusted for age, sex, alcohol intake from beverages (nonconsumers; women >0–6, 6–12, and >12 g/day; men >0–12, 12–24, and >24 g/day), smoking (never, former, current ≤20 cigarettes/day, current >20 cigarettes/day), physical activity (cycling and sports in h/week), education (low, medium, high), coffee intake (cups/day), red meat intake (g/day), whole-grain bread intake (g/day), prevalent hypertension (yes/no), BMI (kg/m^2), and waist circumference (cm). ‡Adjustment according to the basic model with additional adjustment for specified biomarkers.
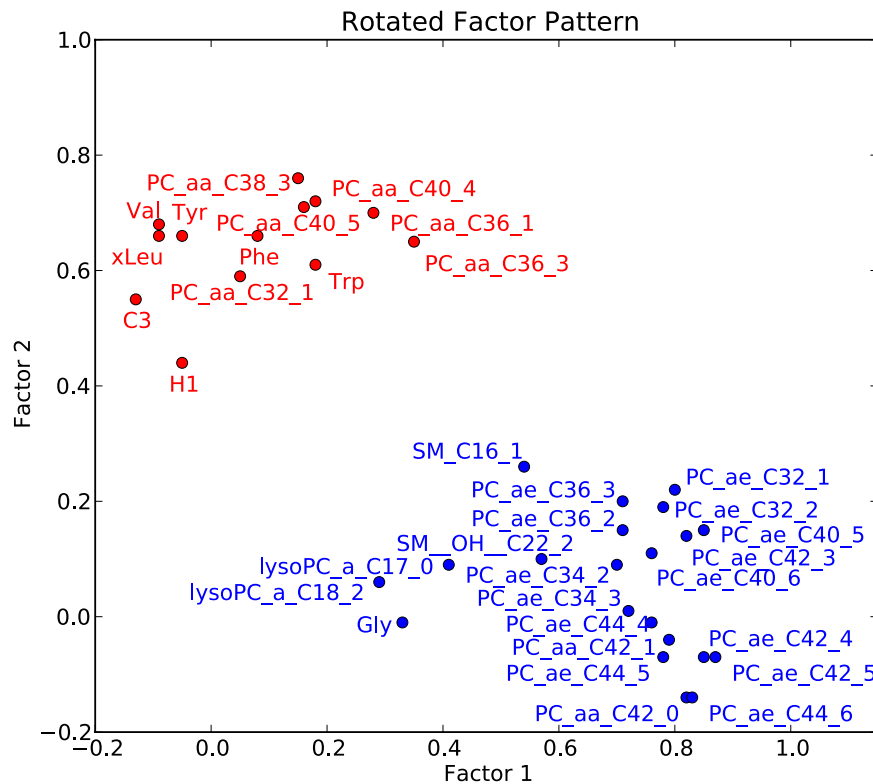
FIG. 1. Two metabolite factors associated with risk of T2D. Presented is a two-dimensional factor loading plot obtained from PCA. For simple interpretation, metabolites that cluster together in the plot are related to one another. Metabolites presented in blue are associated with decreased risk of T2D, whereas metabolites presented in red are associated with increased risk of T2D. More specifically, the factor loadings represent the correlation coefficients of individual metabolites with corresponding metabolite factors and may range from −1 to 1. They were identified by PCA based on the correlation matrix of all metabolites significantly associated with risk of T2D in the EPIC-Potsdam study. An orthogonal varimax rotation was used, and two factors were retained because they accounted for >50% of the observed variance. a, acyl; aa, diacyl; ae, acyl-alkyl; C3, propionylcarnitine; Gly, glycine; H1, hexose; PC, phosphatidylcholine; Phe, phenylalanine; SM, sphingomyelin; Trp, tryptophan; Tyr, tyrosine; Val; valine; xLeu, isoleucine.

of the metabolites (Table 2). An exception was that the inverse association between acyl-alkyl-phosphatidylcholines and sphingomyelin C16:1 and T2D risk was attenuated and no longer significant after adjustment for blood glucose, HbA_{1c}, HDL cholesterol, and triglycerides. The positive association between hexose and T2D risk was considerably weakened but remained significant after adjustment for plasma glucose. When we also adjusted for phenylalanine, tyrosine, and isoleucine, the associations for metabolite factor 1 were unchanged. The associations for metabolite factor 2, which included these three amino acids, were weakened but remained significant (data not shown).

We further observed that metabolites were linked to established T2D biomarkers (Supplementary Table 3). Specifically, metabolite factor 1, which was inversely associated with T2D risk, was negatively correlated to plasma glucose, HbA_{1c}, and triglycerides and positively related to HDL cholesterol and adiponectin. Metabolite factor 2, which was positively associated with T2D risk, was positively correlated with triglycerides and liver enzymes. Data from the TüF study revealed that acyl-alkyl-phosphatidylcholines, lysophosphatidylcholine C18:2, and glycine were positively associated with insulin sensitivity, whereas hexose and diacyl-phosphatidylcholines were inversely related to insulin sensitivity (Table 4). Furthermore, phenylalanine was positively associated with insulin secretion, whereas hexose, sphingomyelin C16:1, and acyl-alkyl-phosphatidylcholines were inversely related to insulin secretion. The

potential of identified metabolites to discriminate between T2D cases and noncases was comparable to that of the DRS (16) (ROC AUC 0.849 and 0.847, respectively, P for difference = 0.838) (Fig. 2). When the metabolites were added to established risk prediction models of T2D, discrimination was slightly but significantly improved up to a ROC AUC of 0.912, and these models were well calibrated (Fig. 2 and Supplementary Table 4). Replication in KORA revealed significant associations with T2D risk for metabolite factor 2 and hexose (Table 3 and Supplementary Table 5). In KORA, similar trends as in EPIC-Potsdam were seen for metabolite factor 1, acyl-alkyl-phosphatidylcholines, glycine, lysophosphatidylcholine C18:2, and sphingomyelin C16:1, with borderline significance. Although the risk estimates for diacyl-phosphatidylcholines were considerably lower in KORA than in EPIC-Potsdam, there was no significant heterogeneity between studies. We also calculated the RR of T2D for the BCAA and aromatic amino acids, which have recently been reported to be associated with T2D risk in the Framingham Offspring cohort and the Malmö Diet and Cancer study, to facilitate the comparison (35). In EPIC-Potsdam, isoleucine, valine, tyrosine, and phenylalanine were positively associated with T2D risk (RR per SD 1.30 [95% CI 1.17–1.43], 1.27 [1.16–1.40], 1.31 [1.18–1.45], 1.35 [1.22–1.49], respectively); leucine was not measured in EPIC-Potsdam. When combining isoleucine, tyrosine, and phenylalanine, the RR of T2D from lowest to highest quartile was 1.0, 1.13 (0.82–1.54), 1.45 (1.07–1.98), and 2.18 (1.62–2.95), respectively (P for trend < 0.0001).

TABLE 3
RR of T2D by quintiles of metabolite factors

	EPIC-Potsdam		Replication in KORA	
	Cases (N)	RR (95% CI)*	Cases (N)	RR (95% CI)*
Factor 1†				
Quintile 1	296	1.00	37	1.00
Quintile 2	194	0.61 (0.46–0.80)	15	0.59 (0.31–1.33)
Quintile 3	148	0.50 (0.37–0.67)	20	0.60 (0.31–1.14)
Quintile 4	95	0.37 (0.27–0.51)	12	0.56 (0.27–1.21)
Quintile 5	67	0.31 (0.21–0.44)	7	0.49 (0.17–1.25)
P for trend		2.95×10^{-13}		1.00×10^{-1}
Factor 2†				
Quintile 1	51	1.00	6	1.00
Quintile 2	102	1.23 (0.82–1.85)	9	1.38 (0.45–4.25)
Quintile 3	149	1.72 (1.17–2.51)	15	1.50 (0.52–4.32)
Quintile 4	191	1.80 (1.24–2.63)	18	1.81 (0.62–5.23)
Quintile 5	307	3.82 (2.64–5.52)	43	4.95 (1.96–12.48)
P for trend		6.64×10^{-18}		1.10×10^{-5}

a, acyl; aa, diacyl; ae, acyl-alkyl; PC, phosphatidylcholine; SM, sphingomyelin. *Relative risks were calculated with multivariate Cox regression (using the Prentice method to account for the case-cohort design in EPIC-Potsdam) across quintiles of metabolite factors after standardizing metabolite concentrations to a mean of 0 (SD 1). The model was adjusted for age, sex, alcohol intake from beverages (nonconsumers; women >0–6, 6–12, and >12 g/day; men >0–12, 12–24, and >24 g/day), smoking (never, former, current ≤20 cigarettes/day, current >20 cigarettes/day), physical activity (cycling and sports in h/week), education (low, medium, high), coffee intake (cups/day), red meat intake (g/day), whole-grain bread intake (g/day), prevalent hypertension (yes/no), BMI (kg/m²), and waist circumference (cm). †By conducting a PCA with orthogonal varimax rotation of all metabolites associated with T2D risk in EPIC-Potsdam, two metabolite factors could be identified that explained >50% of the variation of the metabolites. The corresponding linear factor equations (which equal the summed products of each metabolite's standardized concentration and corresponding factor loading) were as follows: Factor 1 = (0.80 × PC ae C32:1) + (0.78 × PC ae C32:2) + (0.70 × PC ae C34:2) + (0.72 × PC ae C34:3) + (0.71 × PC ae C36:2) + (0.71 × PC ae C36:3) + (0.85 × PC ae C40:5) + (0.76 × PC ae C40:6) + (0.82 × PC ae C42:3) + (0.85 × PC ae C42:4) + (0.87 × PC ae C42:5) + (0.76 × PC ae C44:4) + (0.78 × PC ae C44:5) + (0.83 × PC ae C44:6) + (0.82 × PC aa C42:0) + (0.79 × PC aa C42:1) + (0.54 × SM C16:1) + (0.57 × SM OH C22:2) + (0.41 × lysoPC a C17:0). Factor 2 = (0.55 × propionylcarnitine) + (0.66 × phenylalanine) + (0.61 × tryptophan) + (0.66 × tyrosine) + (0.68 × valine) + (0.66 × isoleucine) + (0.59 × PC aa C32:1) + (0.70 × PC aa C36:1) + (0.65 × PC aa C36:3) + (0.76 × PC aa C38:3) + (0.72 × PC aa C40:4) + (0.71 × PC aa C40:5) + (0.44 × hexose).

We conducted several sensitivity analyses. In EPIC-Potsdam, a small proportion of the participants (14.3%) had fasted. The proportion of incident T2D cases was equally distributed among fasting and nonfasting participants (26.8% and 26.7%, respectively). Additional adjustment for fasting status did not change the results. Further, we did not observe an effect modification of fasting status on the association between metabolite factors 1 and 2 and T2D risk (P for interaction = 0.115 and 0.688, respectively). We observed no interaction with sex (P = 0.407 and 0.441, respectively), and in both men and women, the risk associations were similar. Hazard rates of T2D in different periods of follow-up were not different (factor 1 P = 0.126, factor 2 P = 0.994), indicating that follow-up time did not affect the association between metabolite factors and risk of T2D. To ensure that the metabolite changes preceded the onset of T2D and were not attributed to prediabetic conditions, we repeated the analysis, excluding all cases of T2D that occurred shortly after the baseline examination

TABLE 4
Correlation between metabolites associated with T2D risk and measures of insulin sensitivity and secretion in the TüF study

	Insulin sensitivity	Insulin secretion
Hexose	−0.52	−0.39
Phenylalanine	−0.08	0.24
Glycine	0.34	0.08
SM C16:1	0.01	−0.13
PC ae C34:3	0.15	−0.24
PC ae C40:6	0.25	−0.08
PC ae C42:5	0.11	−0.24
PC ae C44:4	0.08	−0.21
PC ae C44:5	0.12	−0.30
PC aa C32:1	−0.15	−0.02
PC aa C36:1	−0.05	−0.09
PC aa C38:3	−0.24	0.01
PC aa C40:5	0.07	0.12
LysoPC a C18:2	0.36	−0.14
Factor 1†	0.23	−0.27
Factor 2‡	−0.15	0.00

Data are partial Spearman correlation coefficients. Insulin sensitivity was adjusted for age and sex; insulin secretion was adjusted for age, sex, and insulin sensitivity. Insulin sensitivity and insulin secretion were estimated from OGTTs in the TüF study (n = 76). a, acyl; aa, diacyl; ae, acyl-alkyl; PC, phosphatidylcholine; SM, sphingomyelin. †Factor 1 = (0.82 × PC aa C42:0) + (0.79 × PC aa C42:1) + (0.80 × PC ae C32:1) + (0.78 × PC ae C32:2) + (0.70 × PC ae C34:2) + (0.72 × PC ae C34:3) + (0.71 × PC ae C36:2) + (0.71 × PC ae C36:3) + (0.85 × PC ae C40:5) + (0.76 × PC ae C40:6) + (0.82 × PC ae C42:3) + (0.85 × PC ae C42:4) + (0.87 × PC ae C42:5) + (0.76 × PC ae C44:4) + (0.78 × PC ae C44:5) + (0.83 × PC ae C44:6) + (0.54 × SM C16:1) + (0.57 × SM OH C22:2) + (0.41 × lysoPC a C17:0). ‡Factor 2 = (0.55 × propionylcarnitine) + (0.66 × phenylalanine) + (0.61 × tryptophan) + (0.66 × tyrosine) + (0.68 × valine) + (0.66 × isoleucine) + (0.59 × PC aa C32:1) + (0.70 × PC aa C36:1) + (0.65 × PC aa C36:3) + (0.76 × PC aa C38:3) + (0.72 × PC aa C40:4) + (0.71 × PC aa C40:5) + (0.44 × hexose).

during the first 2 years of follow-up (n = 208). The risk associations were slightly lower, but not markedly different.

DISCUSSION

In this prospective investigation using a targeted metabolomic approach at population level, we found increased concentrations of hexose; phenylalanine; and diacyl-phosphatidylcholines C32:1, C36:1, C38:3, and C40:5 and reduced concentrations of glycine; sphingomyelin C16:1; acyl-alkyl-phosphatidylcholines C34:3, C40:6, C42:5, C44:4, and C44:5; and lysophosphatidylcholine C18:2 to be independently predictive of T2D in EPIC-Potsdam. The results agree with data from cross-sectional studies showing that patients with T2D had increased concentrations of sugar metabolites (13), acylcarnitines (14), and BCAA (13) and reduced concentrations of glycine (12). We were able to further replicate the results of Wang et al. (35), who recently reported that BCAA and aromatic amino acids predicted T2D in the prospective Framingham Offspring cohort and the Malmö Diet and Cancer study. In agreement with Wang et al. (35), we found higher concentrations of phenylalanine, isoleucine, tyrosine, and valine to be associated with increased risk of T2D and glycine to be associated with reduced risk of T2D. However, in the present study, BCAA and aromatic amino acids were linked to each other, and only phenylalanine was independently associated with T2D risk when accounting for the other metabolites. BCAAs may serve as substrates
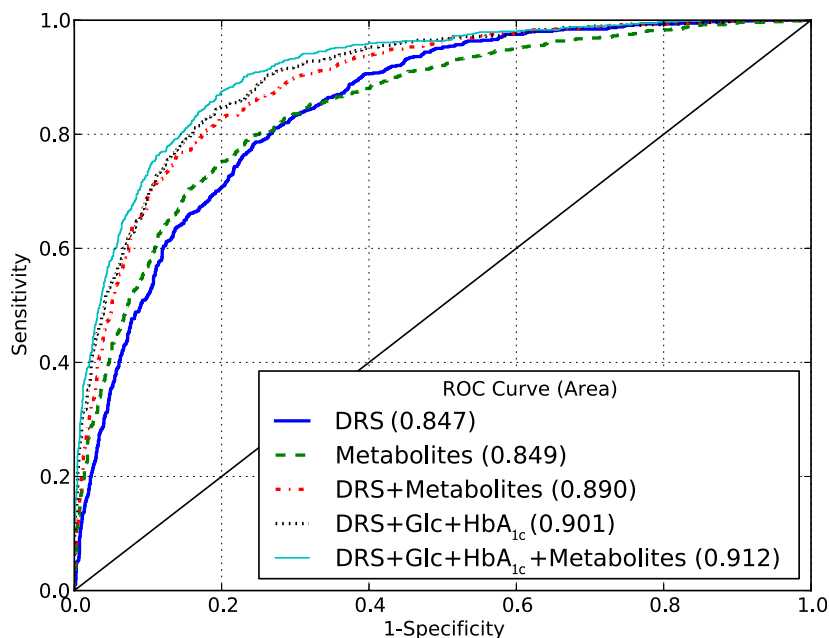
FIG. 2. Relative contribution of metabolites to predict T2D in EPIC-Potsdam. Presented are ROC curves comparing different multivariable-adjusted models to predict T2D, including the DRS, the identified metabolites, glucose (Glc), and HbA$_{1c}$. The DRS (16) combines information on several diabetes risk factors, such as diet, lifestyle, and anthropometry, to estimate risk of developing T2D. The DRS is computed according to the following formula: DRS = (7.4 × waist circumference [cm]) − (2.4 × height [cm]) + (4.3 × age [years]) + (46 × hypertension [self-report]) + (49 × red meat [each 150 g/day]) − (9 × whole-grain bread [each 50 g/day]) − (4 × coffee [each 150 g/day]) − (20 × moderate alcohol [between 10 and 40 g/day]) − (2 × physical activity [h/week]) + (24 × former smoker) + (64 × current heavy smoker [≥ 20 cigarettes/day]). Metabolites are hexose; phenylalanine; glycine; sphingomyelin C16:1; diacyl-phosphatidylcholines C32:1, C36:1, C38:3, and C40:5; acyl-alkyl-phosphatidylcholines C34:3, C40:6, C42:5, C44:4, and C44:5; and lysophosphatidylcholine C18:2.

for the glucose-alanine cycle in skeletal muscle. Through alanine aminotransferase–catalyzed transamination reactions, this may result in increased substrate availability for hepatic gluconeogenesis, thereby increasing hepatic glucose production (39). Conversely, glycine is a gluconeogenic amino acid; therefore, reduced serum glycine may also reflect increased gluconeogenesis. Alternative theories suggest that glycine depletion may reflect glutathione consumption driven by oxidative stress (40) or abundance of incompletely oxidized fuels that are excreted as urinary acylglycine conjugates (41–43). The frequently observed increase of BCAA in subjects with insulin resistance is also believed to be the result of reduced activities of key BCAA catabolic enzymes in liver and adipose tissue (44). Furthermore, amino acids may directly cause muscular insulin resistance by disrupting insulin signaling (45).

The positive association between hexose and T2D risk remained significant after adjustment for glucose. This observation could be an artifact from the different methods used to measure hexose and glucose. However, it has to be noted that hexose represented not only glucose but also the sum of all six-carbon monosaccharides. Previous studies have shown that in addition to glucose, fructose levels were elevated in individuals with T2D (12) and that intake of fructose was positively associated with risk of insulin resistance and T2D (46). In insulin-resistant conditions, the body aims to compensate for decreased glucose uptake of peripheral tissues through increased pancreatic insulin secretion (1). However, at the stage of overt insulin resistance, this system will eventually be exhausted, as caused by β-cell dysfunction and, subsequently, insulin secretion decreases (1). Phenylalanine, which was positively correlated to insulin secretion in the present study, may be involved in pathways to compensate early stages of insulin resistance through stimulation of insulin secretion. In contrast, increased hexose concentrations may indicate manifest insulin resistance and defect of β-cells.

It is noteworthy that we observed significant associations between choline-containing phospholipids (i.e., diacyl-, acyl-alkyl-, and lysophosphatidylcholines and sphingomyelins) and T2D risk. Diacyl-phosphatidylcholines consist of glycerol linked to phosphocholine and two fatty acid residues, and removal of one fatty acid produces lysophosphatidylcholines. The corresponding acyl-alkyl-phosphatidylcholines comprise an ether linkage to one alkyl chain and one polyunsaturated fatty acid (47). Sphingomyelins are built of a ceramide core linked to one fatty acid and a phosphocholine or phosphoethanolamine (Fig. 3). Together, these phospholipids make up the main constituent of cellular membranes and may be involved in cellular signal transduction (48). In addition, they represent a major fraction of the human plasma lipidome because they are most abundant in all lipoproteins (49). Diacyl-phosphatidylcholines are particularly essential for hepatic secretion of triglyceride-rich VLDL particles and HDL (48), whereas acyl-alkyl-phosphatidylcholines may act as serum antioxidants to prevent lipoprotein oxidation (50). Their hepatic synthesis requires dietary choline (48). It was previously shown that choline-deficient mice on a high-fat diet showed reduced phosphatidylcholine biosynthesis and accumulated hepatic fat, but at the same time, they had reduced fasting insulin and improved glucose tolerance (51). In addition, impaired hepatic phosphatidylcholine biosynthesis led to reduced levels of plasma triglycerides and HDL cholesterol in vivo (52).
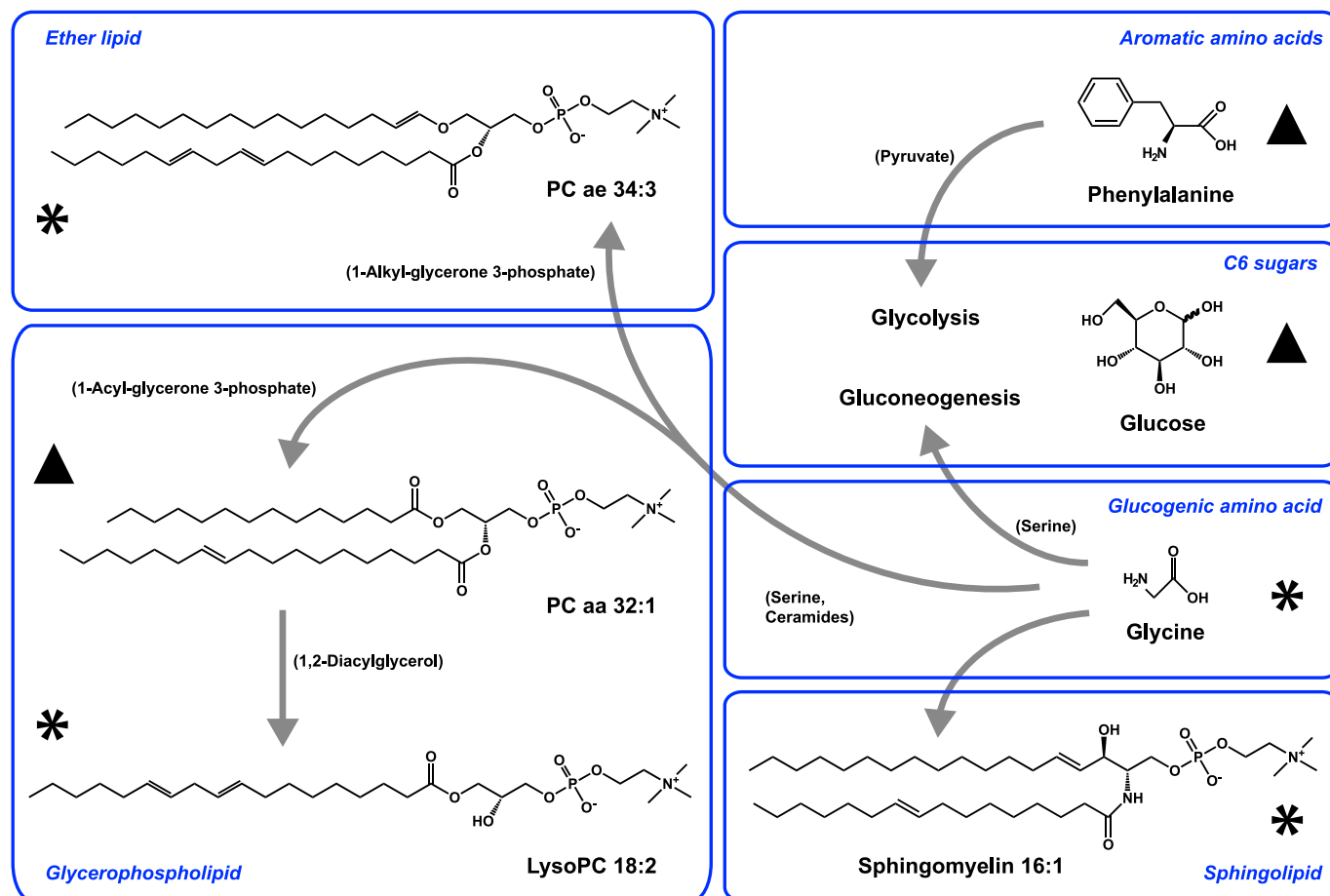
FIG. 3. Examples of metabolites associated with risk of T2D. ▲Metabolites with an increased risk (hexose, phenylalanine, and diacyl-phosphatidylcholines [PCs] C32:1, C36:1, C38:3, and C40:5). *Metabolites with a decreased risk (glycine; sphingomyelin C16:1; acyl-alkyl-PCs C34:3, C40:6, C42:5, C44:4, and C44:5; and lysophosphatidylcholine C18:2). Note that the mass spectrometric assay used does not distinguish molecular lipids and sugar types among hexoses. Therefore, formulas are given for a molecule corresponding to molecular mass and composition. Positions of double bonds and chain length may vary if more than one acid residue is present. Arrows represent many reactions, and key intermediates are given in the brackets. aa, diacyl; ae, acyl-alkyl. (A high-quality color representation of this figure is available in the online issue.)

Accordingly, phosphatidylcholines and sphingomyelins were positively related to plasma HDL cholesterol in the present study. Furthermore, acyl-alkyl-phosphatidylcholines were inversely correlated to plasma triglycerides, opposite to diacyl-phosphatidylcholines, and higher levels of acyl-alkyl-phosphatidylcholines but not diacyl-phosphatidylcholines were linked to improved insulin sensitivity and reduced insulin secretion. Previous studies reported that acyl-alkyl-phosphatidylcholine levels were lower in obese subjects and subjects with insulin resistance (50,53). These mechanisms may contribute to the antithetical association between two phosphatidylcholine subclasses and T2D risk found in the present study and may indicate a key role of the type of linkage between phospholipid core and fatty acid residue. Furthermore, those phosphatidylcholines containing fatty acids with a lower number of carbons and double bonds were positively associated with T2D risk, contrary to those with a higher number of carbons and double bonds. Similar observations have recently been reported for fatty acid compositions of erythrocyte membrane phospholipids (54) and triglycerides (55), suggesting that lipids with a shorter chain length and saturated fatty acid residues may trigger development of T2D, whereas those containing longer chains and unsaturated fatty acids may offer protection.

In summary, the present data suggest that the identified metabolites could be part of different pathways involved in the early genesis of T2D. Therefore, these novel candidates could be useful in clinical practice to identify high-risk individuals earlier in order to delay or prevent disease onset. Furthermore, serum metabolites in this study predicted risk of T2D in a similar manner to a combination of classic risk factors; thus, their measurement may be a useful approach to predict T2D risk for individuals in the future. The metabolites could also serve as markers for specific metabolic pathways that are deranged and, thereby, allow the implementation of individualized preventive and therapeutic strategies. However, future investigations are warranted to calculate in detail the individual risks and to better understand the metabolic effects of these biomarkers and their biological mechanisms.

The primary strength of this study is that, to our knowledge, we were among the first to adopt a targeted metabolomic approach at population level and included a large sample from three independent, well-described study populations. Furthermore, our targeted metabolomic platform covered a wide variety of metabolites with known identity and quantitative measurements. Because we used a prospective design with consecutive follow-up,

we were able to investigate time-dependent exposure-disease associations.

The study, however, had several limitations. First, because we used independent study populations, the conditions of biosample collection, storage, and preparation were not necessarily the same, which may be a source of variation. In the KORA and TüF studies, fasting blood samples were collected from all participants, whereas in EPIC-Potsdam only a small proportion of participants provided fasting blood samples. Nevertheless, we did not observe an effect modification of fasting status on the association between metabolite factors and T2D risk, and we could reproduce very similar metabolite factors comparing fasting to nonfasting samples. Furthermore, the metabolomic analyses were based on serum samples in the EPIC-Potsdam and KORA studies but on plasma samples in the TüF study. As previously reported (56), the correlation between these serum and plasma metabolites was high; however, the absolute metabolite concentrations were higher in serum, which could lead to systematic changes. Second, the analytical method detected most of the metabolites with high specificity; however, it may not have detected all possible interferences among metabolites. Of the metabolites that we identified, diacyl-phosphatidylcholine C38:3 and sphingomyelin C16:1 may be interfering compounds for sphingomyelin C24:1 and diacyl-phosphatidylcholine C30:2, respectively. Third, we only had a limited number of incident T2D cases available from KORA. We may not have had sufficient statistical power, and the replication results have to be interpreted with caution. Fourth, there is a chance that reverse causation may explain the results, implying that overt diabetic conditions that were undiagnosed may have caused these metabolite changes. When we accounted for this issue, the results remained robust. Last, because this was an observational study, we cannot prove causality but only show associations. However, the identified metabolites were also correlated to established T2D biomarkers as well as to measures of insulin sensitivity and secretion in a different population, which underlines the biological plausibility of the results.

In conclusion, this prospective investigation using metabolomics data of independent study populations identified sugar metabolites, amino acids, and choline-containing phospholipids to be independently associated with risk of T2D. Beyond the classic pathways, these candidates point toward a novel role of phospholipid and lipoprotein metabolism in T2D pathophysiology. Future studies should further elucidate the biological mechanisms.

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;58:773–795
2. Peters AL, Davidson MB, Schriger DL, Hasselblad V; Meta-analysis Research Group on the Diagnosis of Diabetes Using Glycated Hemoglobin Levels. A clinical approach for the diagnosis of diabetes mellitus: an analysis using glycosylated hemoglobin levels. JAMA 1996;276:1246–1252
3. Herder C, Baumert J, Zierer A, et al. Immunological and cardiometabolic risk factors in the prediction of type 2 diabetes and coronary events: MONICA/KORA Augsburg case-cohort study. PLoS ONE 2011;6:e19852
4. Abu-Qamar M, Wilson A. Evidence-based decision-making: the case for diabetes care. Int J Evid Based Healthc 2007;5:254–260
5. Swellam M, Sayed Mahmoud And M, Abdel-Fatah Ali A. Clinical implications of adiponectin and inflammatory biomarkers in type 2 diabetes mellitus. Dis Markers 2009;27:269–278
6. Li S, Shin HJ, Ding EL, van Dam RM. Adiponectin levels and risk of type 2 diabetes: a systematic review and meta-analysis. JAMA 2009;302:179–188
7. Ford ES, Schulze MB, Bergmann MM, Thamer C, Joost HG, Boeing H. Liver enzymes and incident diabetes: findings from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. Diabetes Care 2008;31:1138–1143
8. Stefan N, Fritsche A, Weikert C, et al. Plasma fetuin-A levels and the risk of type 2 diabetes. Diabetes 2008;57:2762–2767
9. Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB. Metabolomics applied to diabetes research: moving from information to knowledge. Diabetes 2009;58:2429–2443
10. 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
11. Wopereis S, Rubingh CM, van Erk MJ, et al. Metabolic profiling of the response to an oral glucose tolerance test detects subtle metabolic changes. PLoS ONE 2009;4:e4525
12. 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
13. 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
14. 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
15. Hu FB. Metabolic profiling of diabetes: from black-box epidemiology to systems epidemiology. Clin Chem 2011;57:1224–1226
16. Schulze MB, Hoffmann K, Boeing H, et al. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. Diabetes Care 2007;30:510–515
17. Boeing H, Korfmann A, Bergmann MM. Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. Ann Nutr Metab 1999;43:205–215
18. Kroke A, Bergmann MM, Lotze G, Jeckel A, Klipstein-Grobusch K, Boeing H. Measures of quality control in the German component of the EPIC study. European Prospective Investigation into Cancer and Nutrition. Ann Nutr Metab 1999;43:216–224

19. Boeing H, Wahrendorf J, Becker N. EPIC-Germany—a source for studies into diet and risk of chronic diseases. European Investigation into Cancer and Nutrition. Ann Nutr Metab 1999;43:195–204

20. Weikert C, Stefan N, Schulze MB, et al. Plasma fetuin-A levels and the risk of myocardial infarction and ischemic stroke. Circulation 2008;118:2555–2562

21. Montonen J, Drogan D, Joost HG, et al. Estimation of the contribution of biomarkers of different metabolic pathways to risk of type 2 diabetes. Eur J Epidemiol 2011;26:29–38

22. Bergmann MM, Bussas U, Boeing H. Follow-up procedures in EPIC-Germany—data quality aspects. European Prospective Investigation into Cancer and Nutrition. Ann Nutr Metab 1999;43:225–234

23. Schienkiewitz A, Schulze MB, Hoffmann K, Kroke A, Boeing H. Body mass index history and risk of type 2 diabetes: results from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. Am J Clin Nutr 2006;84:427–433

24. Rothman KJ, Greenland S, Lash TL. Case-control studies. In *Modern epidemiology*. 3rd ed. Rothman KJ, Greenland S, Eds. Philadelphia, Lippincott-Williams & Wilkins, 2008, p. 111–127

25. Meisinger C, Strassburger K, Heier M, et al. Prevalence of undiagnosed diabetes and impaired glucose regulation in 35-59-year-old individuals in Southern Germany: the KORA F4 Study. Diabet Med 2010;27:360–362

26. Rathmann W, Strassburger K, Heier M, et al. Incidence of type 2 diabetes in the elderly German population and the effect of clinical and lifestyle risk factors: KORA S4/F4 cohort study. Diabet Med 2009;26:1212–1219

27. Kowall B, Rathmann W, Strassburger K, Meisinger C, Holle R, Mielck A. Socioeconomic status is not associated with type 2 diabetes incidence in an elderly population in Germany: KORA S4/F4 cohort study. J Epidemiol Community Health 2011;65:606–612

28. Stefan N, Kantartzis K, Machann J, et al. Identification and characterization of metabolically benign obesity in humans. Arch Intern Med 2008;168:1609–1616

29. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22:1462–1470

30. Römisch-Margl W, Prehn C, Bogumil R, Röhring C, Suhre K, Adamski J. Procedure for tissue sample preparation and metabolite extraction for high throughput targeted metabolomics. Metabolomics 11 March 2011 [Epub ahead of print]

31. Floegel A, Drogan D, Wang-Sattler R, et al. Reliability of serum metabolite concentrations over a 4-month period using a targeted metabolomic approach. PLoS ONE 2011;6:e21103

32. Prentice RL. Design issues in cohort studies. Stat Methods Med Res 1995;4:273–292

33. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979;6:65–70

34. Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. Stat Med 1998;17:841–856

35. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. Nat Med 2011;17:448–453

36. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845

37. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–560

38. Hsieh FY, Lavori PW. Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates. Control Clin Trials 2000;21:552–560

39. Ruderman NB. Muscle amino acid metabolism and gluconeogenesis. Annu Rev Med 1975;26:245–258

40. Sekhar RV, McKay SV, Patel SG, et al. Glutathione synthesis is diminished in patients with uncontrolled diabetes and restored by dietary supplementation with cysteine and glycine. Diabetes Care 2011;34:162–167

41. Koves TR, Ussher JR, Noland RC, et al. Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance. Cell Metab 2008;7:45–56

42. Muoio DM, Neufer PD. Lipid-induced mitochondrial stress and insulin action in muscle. Cell Metab 2012;15:595–605

43. Vianey-Liaud C, Divry P, Gregersen N, Mathieu M. The inborn errors of mitochondrial fatty acid oxidation. J Inherit Metab Dis 1987;10(Suppl 1):159–200

44. She P, Van Horn C, Reid T, Hutson SM, Cooney RN, Lynch CJ. Obesity-related elevations in plasma leucine are associated with alterations in enzymes involved in branched-chain amino acid metabolism. Am J Physiol Endocrinol Metab 2007;293:E1552–E1563

45. Tremblay F, Brûlé S, Hee Um S, 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;104:14056–14061

46. Montonen J, Järvinen R, Knekt P, Heliövaara M, Reunanen A. Consumption of sweetened beverages and intakes of fructose and glucose predict type 2 diabetes occurrence. J Nutr 2007;137:1447–1454

47. Magnusson CD, Haraldsson GG. Ether lipids. Chem Phys Lipids 2011;164:315–340

48. Cole LK, Vance JE, Vance DE. Phosphatidylcholine biosynthesis and lipoprotein metabolism. Biochim Biophys Acta 2012;182:754–761

49. Quehenberger O, Dennis EA. The human plasma lipidome. N Engl J Med 2011;365:1812–1823

50. Wallner S, Schmitz G. Plasmalogens the neglected regulatory and scavenging lipid species. Chem Phys Lipids 2011;164:573–589

51. Raubenheimer PJ, Nyirenda MJ, Walker BR. A choline-deficient diet exacerbates fatty liver but attenuates insulin resistance and glucose intolerance in mice fed a high-fat diet. Diabetes 2006;55:2015–2020

52. Jacobs RL, Devlin C, Tabas I, Vance DE. Targeted deletion of hepatic CTP: phosphocholine cytidylyltransferase alpha in mice decreases plasma high density and very low density lipoproteins. J Biol Chem 2004;279:47402–47410

53. Pietiläinen KH, Sysi-Aho M, Rissanen A, et al. Acquired obesity is associated with changes in the serum lipidomic profile independent of genetic effects—a monozygotic twin study. PLoS ONE 2007;2:e218

54. Kröger J, Zietemann V, Enzenbach C, et al. Erythrocyte membrane phospholipid fatty acids, desaturase activity, and dietary fatty acids in relation to risk of type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. Am J Clin Nutr 2011;93:127–142

55. 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

56. Yu Z, Kastenmüller G, He Y, et al. Differences between human plasma and serum metabolite profiles. PLoS ONE 2011;6:e21230