High-Coverage Targeted Lipidomics Reveals Novel Serum Lipid Predictors and Lipid Pathway Dysregulation Antecedent to Type 2 Diabetes Onset in Normoglycemic Chinese Adults

Jieli Lu,[1] Sin Man Lam,[2] Qin Wan,[3] Lixin Shi,[4] Yanan Huo,[5] Lulu Chen,[6] Xulei Tang,[7] Bowen Li,[2] Xueyan Wu,[1] Kui Peng,[1] Mian Li,[1] Shuangyuan Wang,[1] Yu Xu,[1] Min Xu,[1] Yufang Bi,[1] Guang Ning,[1] Guanghou Shui,[2] and Weiqing Wang[1]

Diabetes Care 2019;42:2117–2126 | https://doi.org/10.2337/dc19-0100

OBJECTIVE

Comprehensive assessment of serum lipidomic aberrations before type 2 diabetes mellitus (T2DM) onset has remained lacking in Han Chinese. We evaluated changes in lipid coregulation antecedent to T2DM and identified novel lipid predictors for T2DM in individuals with normal glucose regulation (NGR).

RESEARCH DESIGN AND METHODS

In the discovery study, we tested 667 baseline serum lipids in subjects with incident diabetes and propensity score–matched control subjects (*n* = 200) from a prospective cohort comprising 3,821 Chinese adults with NGR. In the validation study, we tested 250 lipids in subjects with incident diabetes and matched control subjects (*n* = 724) from a pooled validation cohort of 14,651 individuals with NGR covering five geographical regions across China. Differential correlation network analyses revealed perturbed lipid coregulation antecedent to diabetes. The predictive value of a serum lipid panel independent of serum triglycerides and 2-h postload glucose was also evaluated.

RESULTS

At the level of false-discovery rate <0.05, 38 lipids, including triacylglycerols (TAGs), lyso-phosphatidylinositols, phosphatidylcholines, polyunsaturated fatty acid (PUFA)–plasmalogen phosphatidylethanolamines (PUFA-PEps), and cholesteryl esters, were significantly associated with T2DM risk in the discovery and validation cohorts. A preliminary study found most of the lipid predictors were also significantly associated with the risk of prediabetes. Differential correlation network analysis revealed that perturbations in intraclass (i.e., non–PUFA-TAG and PUFA-TAGs) and interclass (i.e., TAGs and PUFA-PEps) lipid coregulation preexisted before diabetes onset. Our lipid panel further improved prediction of incident diabetes over conventional clinical indices.

CONCLUSIONS

These findings revealed novel changes in lipid coregulation existing before diabetes onset and expanded the current panel of serum lipid predictors for T2DM in normoglycemic Chinese individuals.

[1]Shanghai National Clinical Research Center for Endocrine and Metabolic Diseases, Key Laboratory for Endocrine and Metabolic Diseases of the National Health Commision of the People's Republic of China, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao-Tong University School of Medicine, Shanghai, China
[2]State Key Laboratory of Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China
[3]Affiliated Hospital of Southwest Medical University, Luzhou, China
[4]Affiliated Hospital of Guiyang Medical College, Guiyang, China
[5]Jiangxi Provincial People's Hospital Affiliated to Nanchang University, Nanchang, China
[6]Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[7]The First Hospital of Lanzhou University, Lanzhou, China

Corresponding authors: Weiqing Wang, wqingw61@163.com, and Guanghou Shui, ghshui@genetics.ac.cn

Received 15 January 2019 and accepted 29 July 2019

This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc19-0100/-/DC1.

J.L. and S.M.L. contributed equally to this work.

Diabetes has become a major cause of death and disability worldwide (1). The prevalence of diabetes in China has more than quadrupled in recent decades (2,3). An estimated 10.9% of Chinese adults had diabetes and 35.7% had prediabetes in 2013 (3). Identification of biomarkers or novel pathway dysregulation predictive of diabetes denotes an area of intense research interest.

The advent of omics in recent decades offers clinicians an additional avenue to examine disease-relevant metabolite changes with an unprecedented resolution and coverage (4). Metabolomics captures changes of both endogenous and exogenous origins (5) and thus confers further insights to the intricate pathophysiology of diseases such as diabetes, for which phenotypic manifestations integrate both genetic and environmental inputs (4).

Lipidomics has emerged as an independent subfield of metabolomics, and a growing number of studies have investigated the relationship of dysregulation in lipid metabolism and pathogenesis of type 2 diabetes mellitus (T2DM) (6–16). A lipid biomarker panel comprising triacylglycerols (TAGs) of lower total carbon number and carbon double bonds was found to associate with an increased risk of diabetes in the Framingham Heart Study (FHS) cohort (9). Odd-chain TAGs, after adjustment for total TAGs, were shown to inversely correlate with diabetes risk in the Prevención con Dieta Mediterránea (PREDIMED) trial cohort (17). Recent studies have also linked lyso-phosphatidylcholine (LPC), phosphatidylcholine (PC), phosphatidylethanolamine (PE), and diacylglycerol (DAG) with increased diabetes risk (10,14,17), whereas sphingomyelins, a major class of mammalian sphingolipids, were associated with decreased risk of diabetes (11,12,17). In addition, changes in acylcarnitines (18,19) and fatty acids (13) have also been reported to associate with diabetes risk.

Most previous studies on human blood (serum/plasma) did not achieve a satisfactory lipidomic coverage to render investigation of lipid pathway dysregulation (9,11,20). A recent omics study used only one internal standard to measure 207 plasma lipids, which could considerably compromise quantitative accuracy (17). Given the immense complexities of the human serum lipidome, limited lipidomic coverage may hinder an unbiased evaluation of the subtle lipid pathway perturbations essentially underlying T2DM onset. In the current study, we used a high-coverage targeted lipidomics approach constructed principally on high-performance lipid chromatography coupled to multiple reaction monitoring (HPLC-MRM), which simultaneously confers accurate quantitation and extensive coverage of essential lipid classes central to the homeostasis of endogenous lipid metabolism (21).

Moreover, preceding lipidomics studies were predominantly conducted using a mixture of individuals with normal glucose regulation (NGR) or impaired glucose regulation (IGR) at baseline (17), making it difficult to dissect whether these identified lipid changes are harbingers of early dysglycemia or dysglycemia preceded these lipid changes (15). A comprehensive evaluation of lipidomic changes in individuals with NGR could therefore identify new markers and pathways that define early T2DM pathogenesis and possibly improve prediction of incident T2DM beyond clinical risk factors.

Herein, we extensively investigated serum lipidome changes in two prospective Chinese cohorts using a nested case-control design. Our aims were to *1*) evaluate incipient lipidome patterns of diabetes existing before dysglycemia, *2*) uncover lipid markers that could improve prediction of T2DM beyond clinical risk factors, and *3*) systematically investigate lipid pathway dysregulation in the prodromal stage of diabetes.

RESEARCH DESIGN AND METHODS

Study Population

The discovery cohort is a prospective, community-based cohort initiated in Jiading district, Shanghai, China, in 2010. Cohort design and characteristics were previously reported in detail elsewhere (22–23). Briefly, 10,569 subjects aged ≥40 years were invited by telephone or door-to-door visits, and 10,375 were enrolled at the baseline survey. Participants with diabetes or IGR at baseline were excluded from the analysis. After a mean follow-up of 4.4 years, among 3,821 eligible subjects with NGR at baseline, 189 developed diabetes. Owing to funding constraints, our pilot study cohort comprised 100 randomly selected case subjects with diabetes and 100 sex-matched NGR control subjects chosen using propensity score matching (PSM) (24) with a logistic model that includes age, BMI, and fasting plasma glucose (FPG).

The validation cohort comprising 55,062 individuals was randomly selected from 5 of 25 centers of the REACTION (Risk Evaluation of cAncers in Chinese diabetic Individuals: a lONgitudinal) study, a nationwide population-based prospective cohort of 259,657 individuals (aged ≥40 years) enrolled between 2011 and 2012 (25–28). The five centers included two cohorts (Jiangxi and Hubei) from the central region, and three cohorts (Guizhou, Gansu, and Sichuan) from the western region of China. Participants with diabetes or IGR at baseline were excluded from the analysis. During a follow-up of 3.8 years, among 14,651 subjects defined as NGR based on an oral glucose tolerance test (75 g) at baseline, 364 developed diabetes. Two serum samples were not available at baseline, thus 362 individuals with diabetes and 362 NGR individuals were included into the validation study using the same selection and matching methods as described above for the discovery cohort (Supplementary Fig. 1).

Furthermore, in a preliminary study to test whether the lipid predictors for diabetes were also significantly associated with the risk of developing IGR from NGR, 87 participants who developed IGR and 87 who remained NGR during follow-up were randomly selected from the Shanghai cohort.

Ethical Approval

The study protocol was approved by the Institutional Review Board of Ruijin Hospital affiliated to the Shanghai Jiao-Tong University School of Medicine. All participants provided written informed consent.

Definition of Diabetes and IGR

In both discovery and validation studies at baseline and follow-up visits, all participants underwent a 75-g oral glucose tolerance test, and plasma glucose was obtained at 0 h and 2 h during the test. Blood specimens were processed within 2 h of blood collection at the field center, and sera were shipped by air on dry ice to

the central laboratory located at Shanghai Institute of Endocrine and Metabolic Diseases, which is certified by the College of American Pathologists (26).

Incident diabetes was defined as FPG ≥126 mg/dL, or 2-h postload plasma glucose (2hPG) ≥200 mg/dL, or self-reported previous diagnosis of diabetes by physicians and the use of antidiabetic medications. IGR was defined as FPG levels between 100 and 125 mg/dL, and 2hPG levels between 140 and 199 mg/dL in participants without prior diabetes diagnosis.

Lipidomics Analyses

Serum lipid profiles were measured by a high-coverage targeted lipidomics approach constructed principally on HPLC-MRM. Lipids were extracted from serum (20 μL) using a modified Bligh and Dyer extraction procedure (double rounds of extraction) and dried in the SpeedVac under OH mode. All lipidomic analyses were performed on an Exion LC-system coupled with a QTRAP 6500 PLUS system (Sciex), and individual lipids from various classes were quantitated relative to their respective internal standards, as described previously (29–31). Additional details, including quantitative accuracy, calibration, and annotation, are provided in section 2 of the Supplementary Data.

Statistical Analyses

Concentrations for individual lipids (mol/L) were log-transformed and standardized to *z* scores. Because association of lipids with diabetes risk can differ based on acyl chain length and unsaturation degree (9,17), lipids were grouped and further analyzed based on carbon atom and double bond numbers (Supplementary Data sections 1.3 and 1.4).

To identify lipids predictors, odds ratios (ORs) of developing T2DM per log-transformed SD increase in each lipid species were calculated by conditional logistic regression models, after adjusting for age, BMI, smoking and drinking status, education, family history of diabetes, physical activity, systolic blood pressure (SBP), and FPG. The *P* value was corrected for multiple testing via false-discovery rate (FDR) using the Benjamini-Hochberg method. Lipids that showed significant associations with incident diabetes (*P* value <0.05 and FDR <0.05) in both discovery and validation cohorts were further tested in a preliminary IGR study.

To identify independent predictors, serum triglycerides (TGs) and 2hPG, which showed significant differences between case subjects and control subjects at baseline, were further adjusted in conditional logistic regression models. The predictive values for incident diabetes of the identified lipids panel were evaluated in discovery cohort and validation cohort. Model performance was presented as receiver operating characteristic (ROC) curves. The Delong test was used for comparing areas under the ROC curves (Supplementary Data section 1.5).

R package MEGENA was used to build correlation networks from differentially correlated lipid pairs. Differential correlation was calculated using R package DGCA. Only lipid pairs with differential correlation (empirical $P < 0.05$) were included for analyses (see Supplementary Data section 1.6 for details).

All statistical analyses were performed using SAS 9.3 (SAS Institute) and R 3.4.2 software. Two-sided $P < 0.05$ was considered as statistically significant.

RESULTS

Baseline Characteristics of the Discovery and Validation Cohorts

In the discovery cohort, in addition to sex, age, BMI, and FPG matched under PSM, baseline fasting HDL cholesterol, LDL cholesterol, insulin level, SBP, family history of diabetes, and lifestyle factors including smoking, drinking, physical activity, and education status in case subjects and control subjects were also well-matched. In the discovery cohort, case subjects showed higher levels of 2hPG and fasting TG (Table 1). In the validation cohort, case subjects had a significantly higher SBP and fasting TGs than control subjects (Table 1).

Lipid Profiling

A total of 667 lipid species (of a targeted library screening >800 lipids) spanning 24 individual lipid classes was identified and quantitated in the serum lipidome of the discovery study (n = 200), after removing peaks without satisfactory signal-to-noise ratios (<3). Statistical analyses revealed 122 candidate lipids for segregating subjects with incident diabetes from healthy control subjects (Supplementary Data section 1.2). Based on results from the discovery cohort, we used a streamlined method comprising a subset of 250 lipids of outstanding interest to analyze the serum lipidomes of the validation cohort (n = 724) (Supplementary Fig. 2).

Lipid Profiles and Diabetes

In multivariable logistic regression analysis, a panel of 38 lipids (FDR <0.05 and $P < 0.05$) emerged significant between case subjects and control subjects in the discovery and validation cohorts after adjusting for age, sex, BMI, FPG, smoking status, drinking status, education, physical activity, family history of diabetes, and SBP (Fig. 1). The panel contained 34 TAGs spanning across a comprehensive range of total carbon atom numbers (C44–C58) and unsaturation degree (n [C = C] = 0–8) that were consistently associated with an increased risk of T2DM. In addition, short-chain cholesteryl ester (CE14:0), diunsaturated PCs (PC34:3), and lyso-phosphatidylinositol (LPI)16:1 were positively associated with risk of T2DM (Fig. 1). However, polyunsaturated plasmalogen PE (i.e., PE38:4p [18:0p/20:4]) was negatively associated with risk of incident diabetes (Fig. 1).

We systematically looked for lipid patterns associated with diabetes risk that are dependent on carbon atom numbers and double bonds from the various lipid classes investigated. Statistically significant changes ($P < 0.05$) based on carbon atom numbers and unsaturation degree were observed only for TAGs. The discovery cohort was chosen for this aim due to its comprehensive coverage of the serum lipidome in both the carbon atom numbers and double bonds (Supplementary Fig. 2). We found that all serum TAGs examined, regardless of carbon atom numbers (C44–C60) and double bonds (n [C = C] = 0–9), were positively associated ($P < 0.05$) with diabetes risk (Fig. 2*A*, filled red shapes). After adjusting for total TAGs, however, only TAGs with low carbon atom numbers (C48–50) and low double bond numbers (n [C = C] = 2–3) were significantly ($P < 0.05$) associated with elevated diabetes risk (filled red shapes), while TAGs of higher carbon atom numbers (C54–60) or higher double bond numbers (n [C = C] ≥ 4) became inversely associated ($P < 0.05$) with diabetes risk (Fig. 2*B*, filled green shapes). In accordance with Razquin et al. (17), odd-chain TAGs (C53, C55) were

Table 1—Baseline characteristics of the nested case-control subjects selected from the discovery and validation cohorts

	Discovery			Validation		
	Case subjects (n = 100)	Control subjects (n = 100)	P value	Case subjects (n = 362)	Control subjects (n = 362)	P value
Age, years	59.8 ± 8.0	58.8 ± 8.4	0.38	57.9 ± 8.9	57.2 ± 8.9	0.29
Male sex	36 (36.0)	36 (36.0)	1.00	125 (34.5)	125 (34.5)	1.00
BMI, kg/m^2	25.2 ± 3.3	25.0 ± 2.9	0.70	24.0 ± 3.7	23.8 ± 3.4	0.30
High school or above education	28 (28.0)	18 (18.0)	0.13	132 (36.5)	124 (34.3)	0.38
Current cigarette smoking	19 (19.0)	20 (20.0)	1.00	65 (18.0)	61 (16.9)	0.96
Current alcohol drinking	8 (8.0)	9 (9.0)	0.60	39 (10.8)	34 (9.4)	0.93
Physically active during leisure time	11 (11.0)	17 (17.0)	0.31	62 (17.1)	54 (14.9)	0.73
Family history of diabetes	14 (14.0)	5 (5.0)	0.05	50 (13.8)	33 (9.1)	0.11
SBP, mmHg	142.1 ± 18.9	139.5 ± 18.9	0.34	129.6 ± 19.4	126.2 ± 19.3	0.02
DBP, mmHg	83.9 ± 9.6	82.7 ± 9.7	0.36	76.8 ± 11.0	76.3 ± 10.7	0.53
FPG, mg/dL	92.5 ± 8.0	92.7 ± 6.9	0.84	96.8 ± 7.6	96.1 ± 7.6	0.21
2hPG, mg/dL	113.1 ± 20.9	102.4 ± 20.8	0.0004	114.3 ± 19.1	112.2 ± 18.6	0.13
Fasting HDL cholesterol, mg/dL	50.6 ± 12.6	52.5 ± 12.5	0.27	50.8 ± 17.0	52.7 ± 15.0	0.11
Fasting LDL cholesterol, mg/dL	129.4 ± 33.2	124.0 ± 27.5	0.21	101.7 ± 31.9	103.6 ± 30.0	0.40
Fasting TGs, mg/dL	107.1 (83.8–150.4)	92.9 (68.0–128.2)	0.01	100.8 (69.9–145.9)	88.7 (64.7–127.1)	0.0004

Data are mean ± SD or median (interquartile) for the continuous variables, or number (%) for categorical variables. DBP, diastolic blood pressure.

inversely associated ($P < 0.05$) with diabetes (filled green shapes) risk after adjusting for total TAGs.

Lipid Profiles and IGR

Next, we investigated the association of the lipid panel in Fig. 1 with the development of IGR in a preliminary study of 87 individuals who developed IGR and 87 individuals who remained NGR during the 4.4-year follow-up, randomly selected from the discovery cohort (Supplementary Fig. 1). We found that 26 of 38 identified lipids showed significant associations ($P < 0.05$) with the risk of IGR (Supplementary Fig. 3). Most identified TAGs, as well as CE14:0 and PC34:3, were also positively associated with risk of IGR, whereas PE38:4p (18:0p/20:4) was negatively associated with IGR (Supplementary Fig. 3).

Independent Lipid Markers and Predictive Value

To identify independent lipid predictors for diabetes, lipids were included one-by-one into the conditional logistic regression model with further adjustment for TG and 2hPG. Of 38 lipids, 6 reached nominal significance ($P < 0.05$): LPI16:1, PC34:3, PE38:4p (18:0p/20:4), TAG50:2 (16:2), TAG51:0 (17:0), and TAG54:7 (22:6) (Fig. 3*A*).

We then assessed the predictive value of these six identified lipids using ROC curve analyses. Matching factors were removed from the reference model. In the discovery cohort, combining our lipid panel with the reference (Ref) model increased the C statistic from 0.664 (Ref) to 0.764 (Ref + lipids) ($P = 0.004$) and from 0.710 (Ref + 2hPG) to 0.781 (Ref + 2hPG + lipids) ($P = 0.016$) (Fig. 3*B* and Supplementary Table 1). In the validation cohorts, when the lipids were added to the established risk prediction models of diabetes, discrimination was significantly improved. The addition of these six lipids increased the C statistic from 0.693 (Ref) to 0.717 (Ref + lipids) ($P = 0.037$) and from 0.698 (Ref + 2hPG) to 0.722 (Ref + 2hPG + lipids) ($P = 0.028$) (Fig. 3*B* and Supplementary Table 1). Models using penalized logistic regression yielded similar results (Supplementary Fig. 4 and Supplementary Table 2).

Differential Correlation Network Analysis

To systematically evaluate the perturbed lipid coregulation underlying diabetes development, we performed multiscale embedded correlation analysis based on the validation cohorts. Multiscale embedded correlation networks illustrate differential correlation between various lipids in subjects with incident diabetes relative to control subjects (Fig. 4). Three notable modules (Fig. 4, lower panel) were identified from the global networks via multiscale clustering analysis as part of MEGENA (Supplementary Data section 1.6). Module I illustrates that polyunsaturated fatty acid (PUFA)–plasmalogen phosphatidylethanolamine (PUFA-PEp; i.e., PE40:5p [22:5]) (Fig. 4, purple hub, bolded) became negatively correlated with several TAGs (orange-red) in incident diabetes relative to control (pink lines, 0/–), indicating that PUFA-PEp became negatively coregulated with TAGs specifically in incident diabetes. Module II comprises predominantly non–PUFA-TAGs (n [C = C] $\leq$ 3) radiating from TAG48:1 (16:1) and TAG44:1 (16:0) as central hubs, whereas module III constitutes PUFA-TAGs (n [C = C] $\geq$ 3) with TAG56:5 (18:2) as the central hub. Each TAG pair was connected by bright green lines (+/++), indicating that individual non–PUFA-TAGs in module II and PUFA-TAGs in module III became increasingly coregulated within the respective modules in subjects with incident diabetes relative to control subjects. Module II and module III therefore cumulatively showed that intraclass coregulation of serum TAGs became increasingly segregated on the basis of fatty acyl unsaturation in subjects with incident diabetes relative to control subjects. Taken together, differential correlation networks revealed that even before diabetes onset, perturbations in intraclass (i.e., non–PUFA-TAG and PUFA-TAGs) and interclass

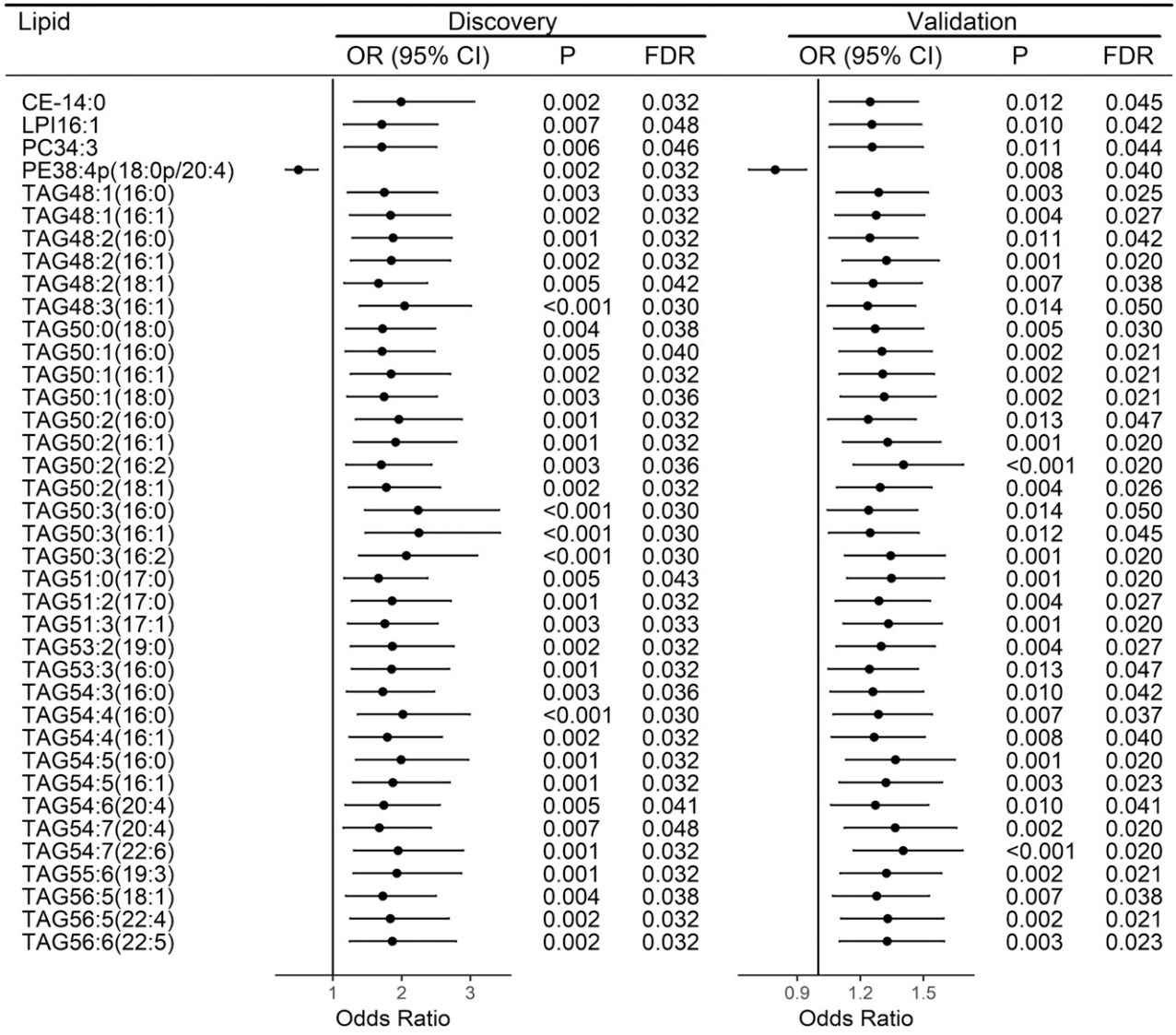

Figure 1—Lipid profiling and risk of incident diabetes. Multivariable-adjusted ORs per one SD increment and 95% CI of lipid species that emerged significant (FDR <0.05) in the discovery (left panel, $n = 200$) and validation (right panel, $n = 724$) cohorts. The multivariate model is adjusted for age, sex, BMI, smoking status, drinking status, education, physical activity, family history of diabetes, FPG, and SBP. After adjustment for multiple testing, 38 lipid species were consistently found to associate with incidence of diabetes.

(i.e., TAGs and PUFA-PEps) lipid coregulation already existed.

CONCLUSIONS

In this prospective study, using a high-coverage targeted HPLC-MRM lipidomics approach, we presented serum lipid predictors for T2DM in a discovery cohort and replicated the findings in separate prospective cohorts of normoglycemic Chinese adults. We showed that these lipids substantially improved T2DM prediction beyond conventional clinical risk factors. Moreover, our results uncovered that intraclass coregulation among TAGs of differing carbon atom numbers and unsaturation degree and the correlation between PUFA-PEps and TAGs were altered in antecedent diabetes before the initiation of glucose perturbance. To our knowledge, this is the largest and most comprehensive lipidomics study investigating the association between lipid profiles and risk of developing diabetes in individuals with NGR at baseline.

Our study contributed a systematic evaluation of serum lipid profile changes predictive of incident diabetes in individuals with NGR. Among the 38 lipids illustrated, 34 species belong to the class of TAGs. We identified far more TAGs associated with increased risk of T2DM than other studies, including FHS (9), European Prospective Investigation into Cancer and Nutrition (EPIC) (10), and PREDIMED (17) (Supplementary Table 3), which may be due to the differences in population characteristics, analytical platform used, and variables adjusted in statistical models. Remarkably, all TAGs identified were associated with increased risk of diabetes. Our results partially corroborated the previous observations that TAGs of lower carbon atom numbers and fewer double bonds were associated with increased risk of diabetes. TAGs of higher carbon atom numbers and more double bonds were, however, found to associate with decreased risk of diabetes in the FHS cohort (9). This discrepancy could be partially explained by the differences in the study design, including only NGR at baseline in

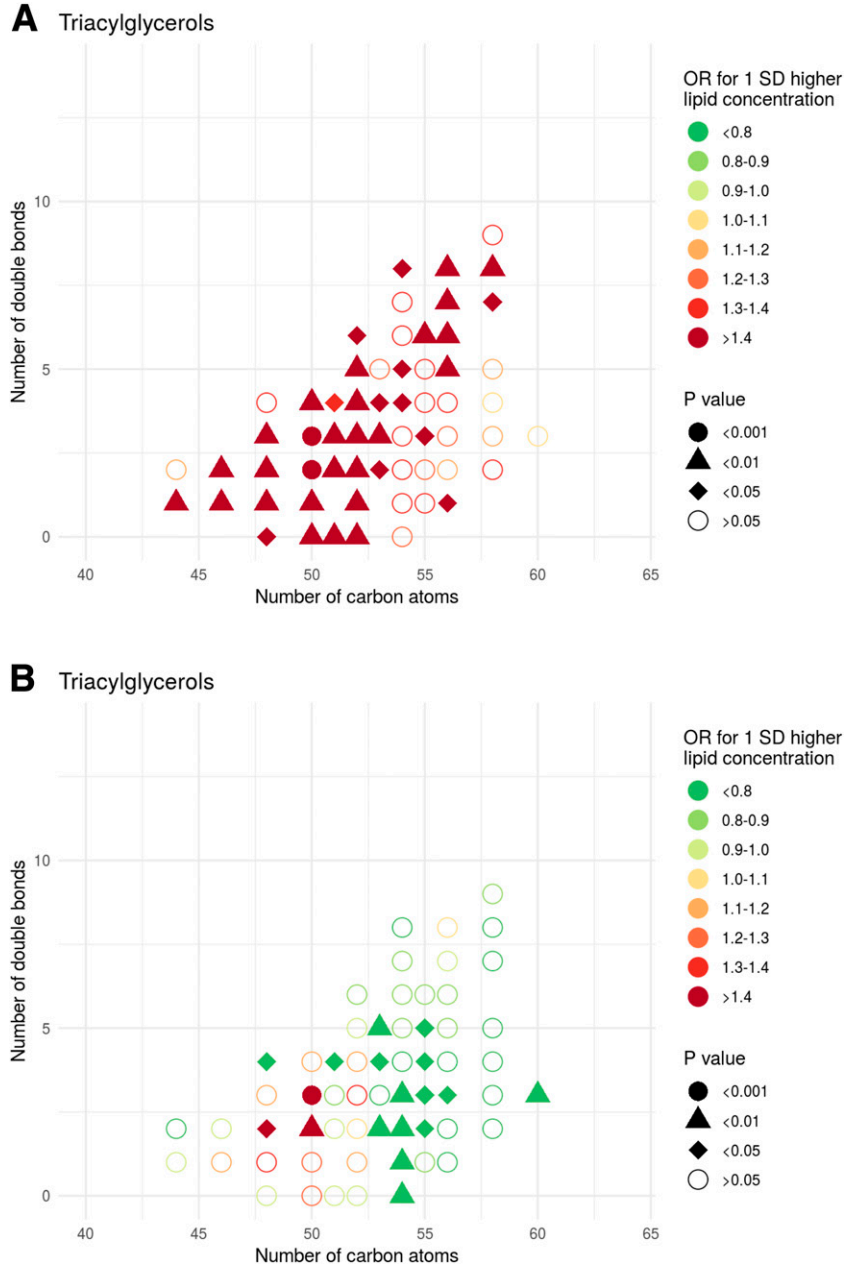

Figure 2—Relationship between diabetes risk and total number of carbon atoms and degree of acyl chain saturation of various lipid classes. Multivariable-adjusted ORs per one SD increment and *P* values of TAG species grouped by carbon atom numbers and double bond numbers in the discovery cohort. The multivariate model is adjusted for age, sex, BMI, smoking status, drinking status, education, physical activity, family history of diabetes, FPG, and SBP. *A*: TAGs in each group were summed, log-transformed, and standardized before conditional logistic regression analysis. *B*: Sum of each TAG group was further normalized to the total TAG content.

the current study versus both NGR and IGR in FHS, and other cohort epidemiological features, including ethnicity-related genetic and lifestyle variations. Interestingly, after adjustment for total measured TAGs, we observed a comparable pattern to Rhee et al. (9) that long-chain PUFA-TAGs were inversely correlated with risk of diabetes, while shorter and more saturated TAGs were associated with increased risk. These observations imply that given a constant total TAG load, compositional changes in specific subtypes of TAGs are associated with differential risk of diabetes. In particular, preceding studies (9,17) and our work have consistently reported that in both Caucasian and Chinese cohorts, long-chain polyunsaturated and/or odd-chain TAGs are inversely associated with diabetes risk, whereas short-chain and saturated TAGs are associated with elevated diabetes risk.

Our preliminary IGR study revealed that the panel of 38 lipids displayed associations with the risk of diabetes and IGR in the same direction, suggesting that our identified lipid predictors are effective in revealing incipient diabetes even before the onset of dysglycemia. It is worth mentioning that early metabolic disturbances might be already present in individuals classified as normal in terms of glycemia. Compared with preceding studies that included IGR at baseline, our reported predictors already have added pathophysiological value in diabetes prediction relative to matched normoglycemic control subjects.

A novel and possibly pathologically important observation from our correlation network analyses is the association between serum PUFA-PEps and the perturbed coregulation between PUFA-TAGs versus non–PUFA-TAGs, in subjects with incident diabetes relative to control subjects, which confer useful lipid-centric insights on the etiology of diabetes. Plasmalogen phospholipids denote essential components of mammalian cellular membranes by virtue of their distinct biophysical and biochemical properties, and previous works had shown that the levels of diacyl PEs are closely adjusted to counter fluctuations in cellular contents of PEps to maintain an overall steady level of total PEs (32,33). Synthesized in the peroxisomes, cells deficient in PEps were reported to exhibit aberrations in peroxisome assembly that drastically reduced the number of peroxisomes, albeit a causal link has yet to be elucidated (32). Peroxisomes are exceptionally dynamic organelles that adapt their number, morphology, and activity in response to physiological needs and nutritional status, which play crucial roles in fatty acid oxidation by their unique ability to consume very-long-chain fatty acids (VLCFAs) and branched-chain fatty acids (BRCFAs) (equivalent to odd-chain fatty acids in mammals) that cannot be β-oxidized by the mitochondria (34). Our observations therefore point to a possibility of attenuated peroxisomal fatty acid oxidation in incident diabetes that may be associated with reduction in PUFA-PEps, which could subsequently skew the compositional profiles of serum TAGs toward species comprising VLCFAs

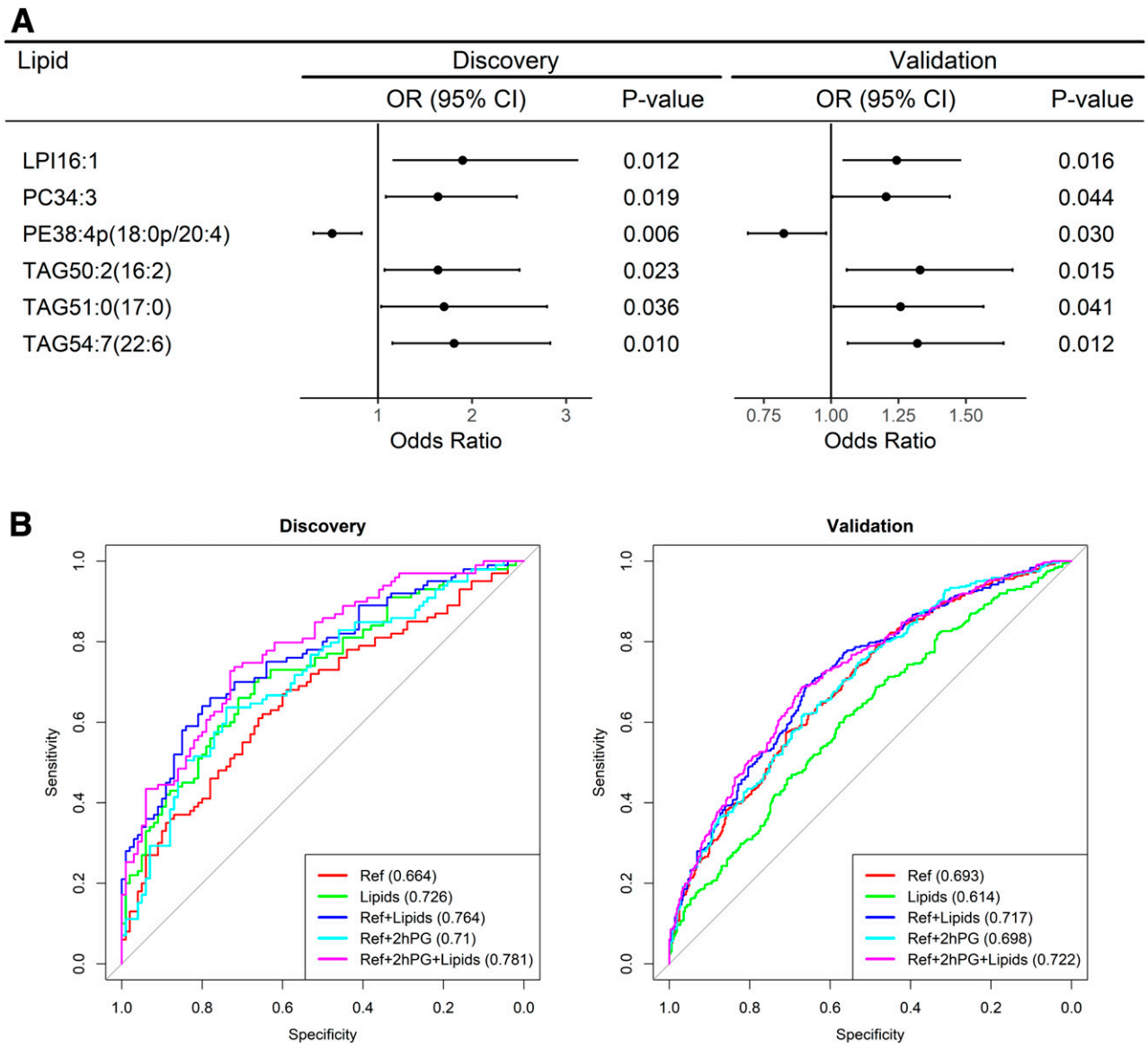

Figure 3—Multivariate logistic regression models of selected lipid panel for diabetes risk prediction in discovery and validation cohorts. *A*: Plot of ORs per one SD increment and 95% CIs of lipid species that emerged significant ($P < 0.05$) in the discovery (left panel, $n = 200$) and validation (right panel, $n = 724$) cohorts. The lipids included LPI16:1, PC34:3, PE38:4p (18:0p/20:4), TAG50:2 (16:2), TAG51:0 (17:0), and TAG54:7 (22:6). The multivariate model is adjusted for age, sex, BMI, smoking status, drinking status, education, physical activity, family history of diabetes, FPG, SBP, serum TGs, and 2hPG. *B*: Plots of area under the curves in the discovery cohort and validation cohort. The reference model included smoking status, drinking status, education, physical activity, family history of diabetes, SBP, and serum TGs. Subsequent models included combinations of the basic clinical variables, plus 2hPG, as well as identified lipid predictors as indicated.

and BRCFAs (i.e., long-chain TAGs and odd-chain TAGs) that may account for their inverse correlation with diabetes risk compared with the remaining TAGs. Nonetheless, the precise causal links between these observations await further mechanistic elucidation, and whether reduction in PUFA-PEps and altered composition of the TAG pool denote cumulating or compensatory events of diabetes onset remains an interesting open question.

Another main clinical finding of the current study is that a selected panel of six lipids substantially improved T2DM prediction beyond that achieved using conventional risk factors in individuals classified as NGR based on an oral glucose tolerance test. We wish to emphasize the pathological relevance of our identified lipids with regard to incipient diabetes, because prediction was performed relative to normoglycemic individuals at baseline. These metabolite predictors may be particularly useful because traditional risk factors, such as blood glucose, may not serve as effective predictors of incident diabetes in such healthy individuals compared with those with IGR. Indeed, a recent study based on the FHS Offspring cohort conducted on individuals with normal fasting glucose at baseline also revealed that a metabolite panel appreciably improved T2DM risk prediction over traditional clinical factors (15).

The major strengths in our current study lie in our adopted analytical approaches as well as the well-characterized study cohorts. Our targeted lipidomics approach constructed upon HPLC-MRM, with specific quantitative internal standards catered to each lipid class, allows unambiguous

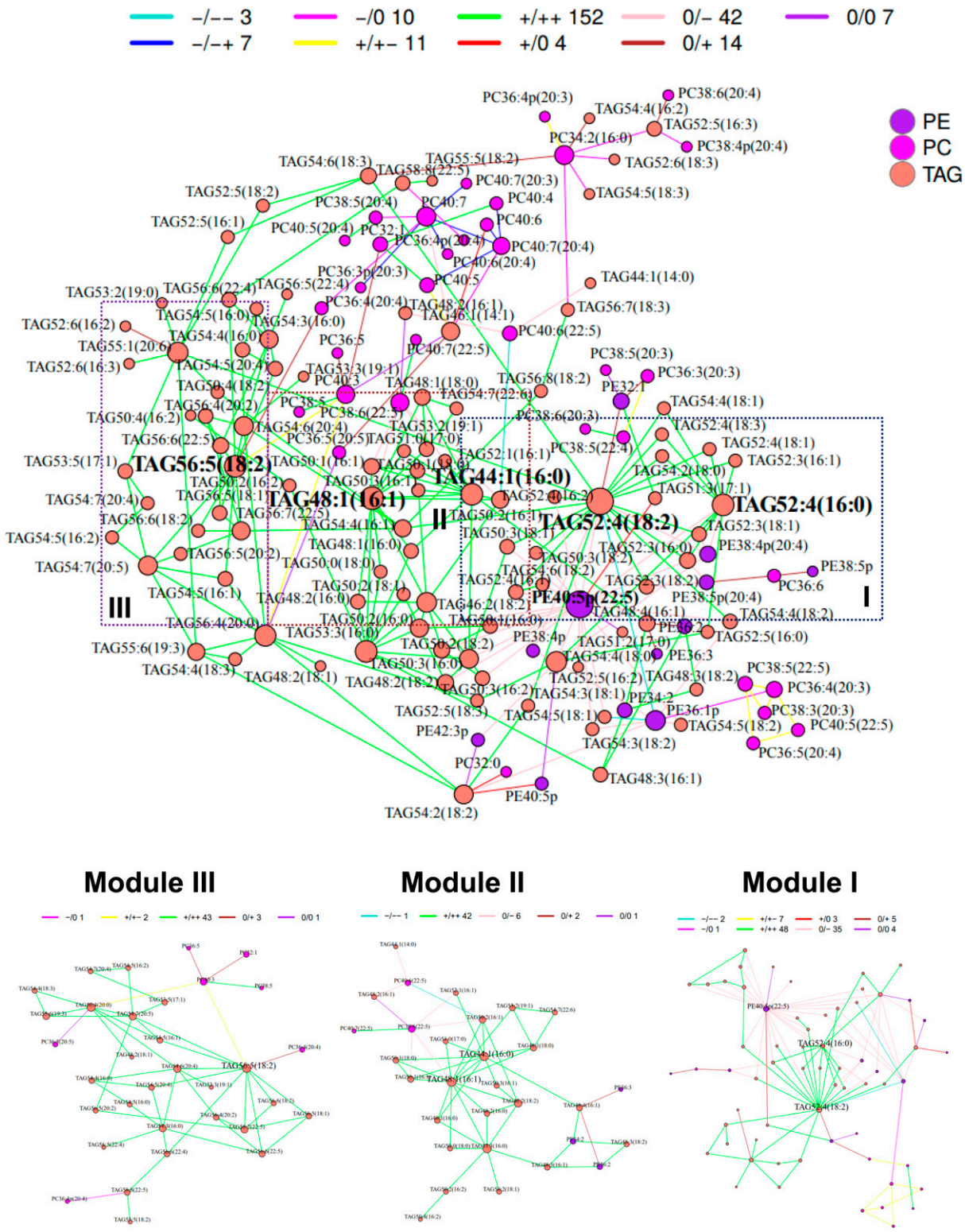

Figure 4—Multiscale embedded correlation network analysis illustrates the differential correlation between various PCs, PEs, and TAGs in control subjects and subjects with incident diabetes. Only lipid pairs with significant differential correlations (empirical $P < 0.05$) are included. Sign/sign indicates the direction and strength of the correlation in control/incident diabetes, and the number that follows indicates the number of lipid pairs in the global networks exhibiting this pattern of change. For instance, the bright green line +/++ 152 in the upper legend of the global networks indicates that correlation between two connected lipid pairs was positive (+) in control subjects, and the correlation became even more strongly positive (++) in those with incident diabetes. A total of 152 lipid pairs connected by bright green lines in the global network displayed this pattern of change (+/++). Three modules identified from multiscale clustering analysis—(I) PUFA-PEps and TAGs, (II) PUFA-PEps and non–PUFA-TAGs, and (III) PUFA-PEps and PUFA-TAGs—are separately illustrated in the lower panel.

identification and accurate quantitation of serum lipids to unveil subtle changes underlying prodromal disease stage. Our targeted approach is also high coverage, which extensively covers several key lipid classes essentially controlling the overall homeostatic balance of endogenous lipid metabolism, rendering possible the systematic evaluation of lipid pathway coregulation via multiscale embedded correlation analyses. In terms of the study cohorts, our discovery and multicenter-derived validation cohorts spanned six different regions of China, which are expected to better represent the general population and increase the credibility and clinical significance of the findings. It is noteworthy that in addition to our novel lipid predictors and pathway dysregulation underlying antecedent diabetes put forth in this study, numerous of our identified lipid predictors also coincided with those reported (Supplementary Table 3) in the overseas Singapore Chinese (LPI16:1) (14), the FHS (TAG44:1 [16:0], TAG44:1 [16:1], TAG48:1 [16:0], TAG48:1 [16:1], and TAG50:0 [18:0]) (9), and the PREDIMED trial (TAG) (17). Acylcarnitines did not emerge as relevant lipids associated with incident diabetes as reported in a previous study (18), which might be related to our study population being normoglycemic at baseline.

Our study, however, also has several limitations. First, the duration of follow-up is only 4 years, which limits the predictive potential of our identified lipid panel to a relatively narrow time window. Second, food consumption habits may influence the baseline levels of circulating lipids and were not controlled in the current model because diet information from participants is lacking. Third, because BMI and age were PSM for the case subjects and control subjects, discrimination by the reference model must therefore be driven by the other included risk factors (e.g., family history, TGs), and it is thus likely that the predictive value may have been affected by PSM. Fourth, we recognize that insulin resistance and insulin secretion are not evaluated in these cohorts. Fifth, all participants in this study were Chinese, and further work is needed to determine whether our findings can be extrapolated to other races and ethnicities.

In conclusion, our foregoing data underscore the potential importance of lipid metabolism early in the pathogenesis of diabetes and suggest that lipid profiles could effectively improve diabetes risk assessment in normoglycemic Chinese populations beyond that achieved by conventional clinic indices. Furthermore, we put forth a conceptual framework that perturbed peroxisomal oxidation of VLCFAs and BRCFAs associated with a reduction in PUFA-PEps may possibly give rise to an altered composition of TAGs essentially underlying the pathogenesis of dyslipidemia that precedes the development of overt diabetes.

Acknowledgments. The authors thank the staff and participants of the REACTION Study for their important contributions.

Funding. This study was supported by the Ministry of Science and Technology of China under Award Numbers 2016YFC1305600, 2016YFC1305202, 2016YFC1304904, 2017YFC1310700, 2018YFC1311705, and 2018YFC1311800; the National Natural Science Foundation of China under Award Numbers 81621061, 81670795, 81622011, and 81561128019; the National Key R&D Program of China under Award Numbers 2018YFA0506900 and 2018YFA0506902; the Shanghai Science and Technology Commission under Award Number YDZX20173100004881; and the Strategic Priority Research Program of the Chinese Academy of Sciences under Award Numbers XDB13000000 and XDA12030200.

The sponsors had no role in the design or conduct of the study.

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

Author Contributions. J.L. and S.M.L. contributed to interpreting the data and writing the manuscript. J.L., S.M.L., Y.B., G.N., and G.S. conceived and designed the study. Q.W., L.S., Y.H., L.C., and X.T. contributed to data acquisition. B.L., X.W., and K.P. performed data processing and analysis. M.L., S.W., Y.X., M.X., Y.B., and W.W. revised the manuscript. G.N., G.S., and W.W. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. GBD 2015 Child Mortality Collaborators. Global, regional, national, and selected subnational levels of stillbirths, neonatal, infant, and under-5 mortality, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015 [published correction appears in Lancet 2017;389:e1]. Lancet 2016; 388:1725–1774

2. Xu Y, Wang L, He J, et al.; 2010 China Noncommunicable Disease Surveillance Group. Prevalence and control of diabetes in Chinese adults. JAMA 2013;310:948–959

3. Wang L, Gao P, Zhang M, et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. JAMA 2017;317:2515–2523

4. Sas KM, Karnovsky A, Michailidis G, Pennathur S. Metabolomics and diabetes: analytical and computational approaches. Diabetes 2015;64: 718–732

5. Lam SM, Wang Y, Li B, Du J, Shui G. Metabolomics through the lens of precision cardiovascular medicine. J Genet Genomics 2017;44:127–138

6. Guasch-Ferré M, Hruby A, Toledo E, et al. Metabolomics in prediabetes and diabetes: a systematic review and meta-analysis. Diabetes Care 2016;39:833–846

7. Chien K, Cai T, Hsu H, et al. A prediction model for type 2 diabetes risk among Chinese people. Diabetologia 2009;52:443–450

8. Schulze MB, Weikert C, Pischon T, et al. Use of multiple metabolic and genetic markers to improve the prediction of type 2 diabetes: the EPIC-Potsdam Study. Diabetes Care 2009; 32:2116–2119

9. 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

10. 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

11. 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

12. Fall T, Salihovic S, Brandmaier S, et al. Nontargeted metabolomics combined with genetic analyses identifies bile acid synthesis and phospholipid metabolism as being associated with incident type 2 diabetes. Diabetologia 2016;59: 2114–2124

13. Würtz P, Tiainen M, Mäkinen VP, et al. Circulating metabolite predictors of glycemia in middle-aged men and women. Diabetes Care 2012;35:1749–1756

14. Lu Y, Wang Y, Zou L, et al. Serum lipids in association with type 2 diabetes risk and prevalence in a Chinese population. J Clin Endocrinol Metab 2018;103:671–680

15. Merino J, Leong A, Liu CT, et al. Metabolomics insights into early type 2 diabetes pathogenesis and detection in individuals with normal fasting glucose. Diabetologia 2018;61:1315–1324

16. Wang-Sattler R, Yu Z, Herder C, et al. Novel biomarkers for pre-diabetes identified by metabolomics. Mol Syst Biol 2012;8:615

17. Razquin C, Toledo E, Clish CB, et al. Plasma lipidomic profiling and risk of type 2 diabetes in the PREDIMED trial. Diabetes Care 2018;41: 2617–2624

18. Sun L, Liang L, Gao X, et al. Early prediction of developing type 2 diabetes by plasma acylcarnitines: a population-based study. Diabetes Care 2016;39:1563–1570

19. Qiu G, Zheng Y, Wang H, et al. Plasma metabolomics identified novel metabolites associated with risk of type 2 diabetes in two prospective cohorts of Chinese adults. Int J Epidemiol 2016;45:1507–1516

20. 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

21. Lam SM, Tian H, Shui G. Lipidomics, en route to accurate quantitation. Biochim Biophys Acta Mol Cell Biol Lipids 2017;1862:752–761

22. Chen Y, Lu J, Huang Y, et al. Association of previous schistosome infection with diabetes and metabolic syndrome: a cross-sectional study in rural China. J Clin Endocrinol Metab 2013;98:E283–E287

23. Peng K, Lin L, Wang Z, et al. Short sleep duration and longer daytime napping are associated with non-alcoholic fatty liver disease in Chinese adults. J Diabetes 2017;9:827–836

24. Parsons L. Performing a 1:N case-control match on propensity score. In *Proceedings of the 29th Annual SAS Users Group International Conference, Montreal, Canada, 2004*. Cary, NC, SAS Institute. Available from: https://support.sas.com/resources/papers/proceedings/proceedings/sugi29/165-29.pdf. Accessed 15 January 2019

25. Lin L, Lu J, Chen L, et al.; REACTION Study Group. Glycemic status and chronic kidney disease in Chinese adults: Findings from the REACTION study. J Diabetes 2017;9:837–845

26. Bi Y, Lu J, Wang W, et al. Cohort profile: risk evaluation of cancers in Chinese diabetic individuals: a longitudinal (REACTION) study. J Diabetes 2014;6:147–157

27. Lu J, Bi Y, Wang T, et al. The relationship between insulin-sensitive obesity and cardiovascular diseases in a Chinese population: results of the REACTION study. Int J Cardiol 2014;172:388–394

28. Lu J, He J, Li M, et al. Predictive value of fasting glucose, postload glucose, and hemoglobin A_{1c} on risk of diabetes and complications in Chinese adults. Diabetes Care 2019;42:1539–1548

29. Lam SM, Tong L, Duan X, Petznick A, Wenk MR, Shui G. Extensive characterization of human tear fluid collected using different techniques unravels the presence of novel lipid amphiphiles. J Lipid Res 2014;55:289–298

30. Shui G, Guan XL, Low CP, et al. Toward one step analysis of cellular lipidomes using liquid chromatography coupled with mass spectrometry: application to *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* lipidomics. Mol Biosyst 2010;6:1008–1017

31. Shui G, Cheong WF, Jappar IA, et al. Derivatization-independent cholesterol analysis in crude lipid extracts by liquid chromatography/mass spectrometry: applications to a rabbit model for atherosclerosis. J Chromatogr A 2011;1218:4357–4365

32. Zoeller RA, Raetz CR. Isolation of animal cell mutants deficient in plasmalogen biosynthesis and peroxisome assembly. Proc Natl Acad Sci U S A 1986;83:5170–5174

33. Dorninger F, Brodde A, Braverman NE, et al. Homeostasis of phospholipids - the level of phosphatidylethanolamine tightly adapts to changes in ethanolamine plasmalogens. Biochim Biophys Acta 2015;1851:117–128

34. Fransen M, Lismont C, Walton P. The peroxisome-mitochondria connection: how and why? Int J Mol Sci 2017;18:1126