

Increased lipogenesis and impaired β -oxidation predict type 2 diabetic kidney disease progression in American Indians

Farsad Afshinnia, ... , Matthias Kretzler, Subramaniam Pennathur

JCI Insight. 2019;4(21):e130317. <https://doi.org/10.1172/jci.insight.130317>.

Clinical Medicine

Metabolism

Nephrology

Graphical abstract

METHODS

Liquid Chromatography-mass spectrometry based lipidomic analysis of serum at baseline in progressors and non-progressors of type-2 diabetic kidney disease (DKD) with $\text{GFR} \geq 90$ mL/min at baseline

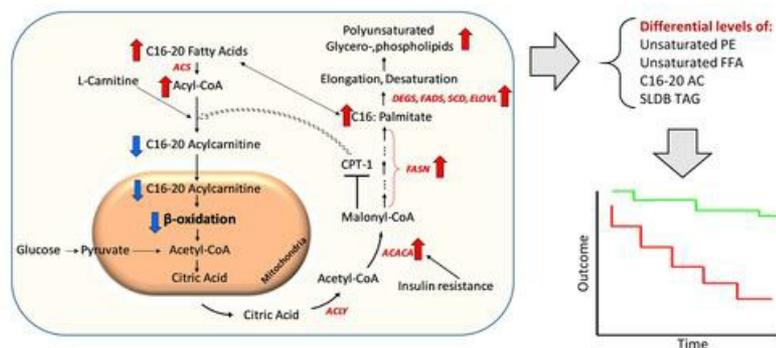


Non-progressors
n = 60

Progressors (40% decline in GFR)
n = 32

N = 92

OUTCOME: 40% sustained decline in longitudinally measured GFR



CONCLUSION: Markers of increased lipogenesis and impaired mitochondrial β -oxidation predict progression of DKD in type 2 diabetes with preserved GFR. Insulin resistance induced renal acetyl-CoA carboxylase activation may be the underlying mechanism linking lipid abnormalities to DKD progression.

Find the latest version:

<https://jci.me/130317/pdf>



Increased lipogenesis and impaired β -oxidation predict type 2 diabetic kidney disease progression in American Indians

Farsad Afshinnia,¹ Viji Nair,¹ Jiahe Lin,² Thekkelnaycke M. Rajendiran,^{3,4} Tanu Soni,³ Jaeman Byun,¹ Kumar Sharma,⁵ Patrice E. Fort,⁶ Thomas W. Gardner,⁶ Helen C. Looker,⁷ Robert G. Nelson,⁷ Frank C. Brosius,⁸ Eva L. Feldman,⁹ George Michailidis,^{10,11} Matthias Kretzler,¹ Subramaniam Pennathur^{1,3,12}

¹Division of Nephrology, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA. ²Department of Statistics, University of Michigan, Ann Arbor, Michigan, USA. ³Michigan Regional Comprehensive Metabolomics Resource Core and ⁴Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan, USA. ⁵Division of Nephrology, Department of Internal Medicine, University of Texas Health at San Antonio, San Antonio, Texas, USA. ⁶Department of Ophthalmology and Visual Sciences, University of Michigan Medical School, Ann Arbor, Michigan, USA. ⁷Chronic Kidney Disease Section, National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, Arizona, USA. ⁸Division of Nephrology, Department of Medicine, University of Arizona College of Medicine, Tucson, Arizona, USA. ⁹Department of Neurology, University of Michigan Medical School, Ann Arbor, Michigan, USA. ¹⁰Department of Statistics and ¹¹Informatics Institute, University of Florida, Gainesville, Florida, USA. ¹²Department of Molecular and Integrative Physiology, University of Michigan Medical School, Ann Arbor, Michigan, USA.

BACKGROUND. In this study, we identified the lipidomic predictors of early type 2 diabetic kidney disease (DKD) progression, which are currently undefined.

METHODS. This longitudinal study included 92 American Indians with type 2 diabetes. Serum lipids (406 from 18 classes) were quantified using mass spectrometry from baseline samples when iothalamate-based glomerular filtration rate (GFR) was at least 90 mL/min. Affymetrix GeneChip Array was used to measure renal transcript expression. DKD progression was defined as at least 40% decline in GFR during follow-up.

RESULTS. Participants had a mean age of 45 ± 9 years and median urine albumin/creatinine ratio of 43 (interquartile range 11–144). The 32 progressors had significantly higher relative abundance of polyunsaturated triacylglycerols (TAGs) and a lower abundance of C16–C20 acylcarnitines (ACs) ($P < 0.001$). In a Cox regression model, the main effect terms of unsaturated free fatty acids and phosphatidylethanolamines and the interaction terms of C16–C20 ACs and short-low-double-bond TAGs by categories of albuminuria independently predicted DKD progression. Renal expression of acetyl-CoA carboxylase-encoding gene (*ACACA*) correlated with serum diacylglycerols in the glomerular compartment ($r = 0.36$, and $P = 0.006$) and with low-double-bond TAGs in the tubulointerstitial compartment ($r = 0.52$, and $P < 0.001$).

CONCLUSION. Collectively, the findings reveal a previously unrecognized link between lipid markers of impaired mitochondrial β -oxidation and enhanced lipogenesis and DKD progression in individuals with preserved GFR. Renal acetyl-CoA carboxylase activation accompanies these lipidomic changes and suggests that it may be the underlying mechanism linking lipid abnormalities to DKD progression.

TRIAL REGISTRATION. ClinicalTrials.gov, NCT00340678.

FUNDING. NIH R24DK082841, K08DK106523, R03DK121941, P30DK089503, P30DK081943, and P30DK020572.

Conflict of interest: KS reports receipt of consulting fees from Boehringer Ingelheim. TWG reports grants from Zebra Biologics. MK reports grants from JDRF, Astra-Zeneca, NovoNordisc, Eli Lilly, Gilead, Goldfinch Bio, Merck, Janssen, and Boehringer-Ingelheim and has a patent titled “Biomarkers for CKD progression” (encompassing urinary EGF as biomarker of CKD progression) issued.

Role of funding source: Financial support of the project.

Copyright: © 2019, American Society for Clinical Investigation.

Submitted: May 17, 2019

Accepted: September 25, 2019

Published: November 1, 2019.

Reference information: *JCI Insight*. 2019;4(21):e130317.
<https://doi.org/10.1172/jci.insight.130317>.

Introduction

Diabetes is the leading cause of end-stage kidney disease in the United States and many other parts of the world (1). According to the 2018 United States Renal Data System report, the prevalence of end-stage kidney disease from diabetes continues to increase (2) and is expected to increase by 44% by 2030 (3). Efforts aimed at risk modification and halting diabetic kidney disease (DKD) progression are focused on the early stages of DKD, when treatment to slow disease development may be most successful. Clinical care of patients with DKD presently relies on traditional biomarkers, such as estimated glomerular filtration rate (eGFR) and urine albumin/creatinine ratio (ACR), which have limited precision and prognostic value early in the disease. Accordingly, there is a need for biomarkers that predict DKD progression early in its course for proper risk stratification and for rational targeted intervention.

Lipid studies in human clinical research are often limited to measuring a traditional panel of lipids, including total cholesterol, lipoproteins, and total triacylglycerols (TAGs) (4–8). However, lipids are the most abundant and diverse class of molecules in the circulation, with numerous physiological and metabolic functions (9, 10). Mass spectrometry-based (MS-based) technological advances for identifying and quantifying lipids in biological samples have revealed the diversity of lipids and also uncovered their clinical relevance in a number of diseases (9–17). In a cross-sectional study of 214 patients from the Clinical Phenotyping and Resource Biobank Core (CPROBE) cohort with chronic kidney disease (CKD), we found progressive differences in the human plasma lipidome from CKD stage 2 to stage 5 (18). In particular, we demonstrated an increase in the abundance of C16–C20 saturated free fatty acids (FFAs) and polyunsaturated longer chain complex lipids in the later stages of CKD. In more advanced CKD, this lipid phenotype was associated with a lower long (C16–C20) to intermediate chain (C6–C14) acylcarnitine (AC) ratio, a marker of impaired β -oxidation of long chain (C16–C20) fatty acids. This leads to an increased abundance of C16–C20 carbon FFAs, rather than their acylation with carnitine to generate ACs and longer polyunsaturated complex lipids (18). In the lipidomic analysis of the Chronic Renal Insufficiency Cohort (CRIC), we found a higher abundance of phosphatidylethanolamines (PEs) in progressors compared with nonprogressors (19). While there have been several efforts to identify progressive renal decline in diabetic patients with impaired renal function (19–24), less is known about the mechanisms, determinants, and markers of early decline in patients with preserved renal function. Early decline was studied recently by the Joslin group, and markers of tubular damage were identified as predictors of progression (23, 25). The central hypothesis of our present study was that the differential lipid alterations observed at more advanced DKD stages are a continuum of a process that starts before the onset of clinical DKD. Additionally, lipid alterations identified earlier may serve as markers of early DKD progression when traditional biomarkers, such as eGFR and urine ACR, are still normal. DKD has been characterized extensively over many years in American Indians from the Gila River Indian Community (26–31). In the present study, we examined the lipidomic predictors of DKD progression in this cohort. We hypothesized, based on our earlier observations (18, 19), that a subset of ACs, TAGs, FFAs, and PEs would predict DKD progression at an early stage in patients with type 2 diabetes with preserved GFR (>90 mL/min) and normal urinary albumin excretion (ACR <30 mg/g). The presence of extensive clinical phenotyping and longitudinal follow-up, and the availability of matched serum and kidney biopsy specimens when renal function is well preserved, make this cohort ideal for performing systems integration of lipidomic and renal lipid gene expression profiling. We believe our studies reveal a previously unrecognized lipidomic signature of early DKD progression.

Results

Patient characteristics. American Indians from the Gila River Indian Community participated in a longitudinal study of diabetes and its complications. Between 1996 and 2001, 169 of the study participants were recruited for a randomized, double-blind, placebo-controlled clinical trial to assess the efficacy of the angiotensin receptor blocker losartan on onset and progression of diabetic nephropathy in type 2 diabetes (ClinicalTrials.gov, NCT00340678). Iothalamate-based GFR was measured annually (32, 33). Of 169 participants in the clinical trial, 111 underwent a protocol kidney biopsy at the end of 6 years of treatment. We selected 92 participants who underwent a kidney biopsy and had both a GFR at least 90 mL/min and serum samples available that had been collected during the GFR measurement taken closest to the kidney biopsy. This examination was considered the baseline (Figure 1). The primary outcome was defined as a sustained decline in GFR by 40% in follow-up examinations, an outcome accepted as a surrogate marker

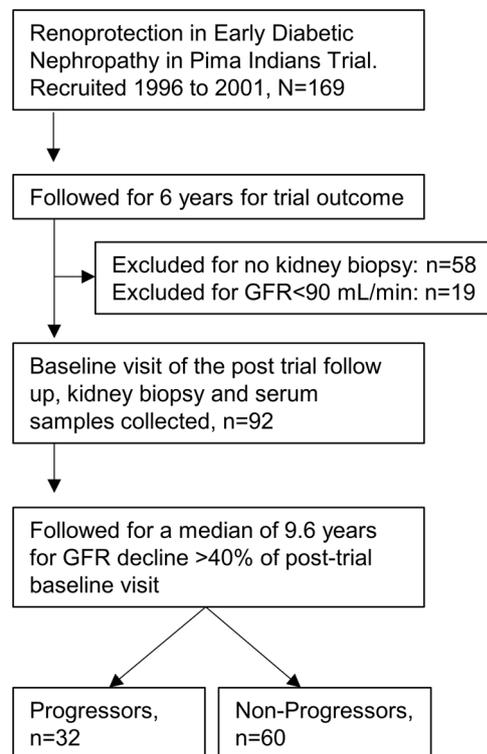


Figure 1. Flow diagram of analytical strategy and post-trial follow-up of participants of the Renoprotection in Early Diabetic Nephropathy in Pima Indians trial analyzed in this study.

of CKD progression (34). Time to event was defined as the interval between the date of baseline GFR measurement and the date of the primary outcome. GFRs less than 90, less than 60, and less than 30 mL/min were defined as secondary outcomes. The 92 participants, including 67 women and 25 men, were followed for a maximum of 13.6 years (median 9.6 years and interquartile range 7.3–11.0 years). Their mean age was 45 ± 9 years, and body mass index (BMI) was 36.1 ± 8.3 kg/m² in the female participants and 34.2 ± 7.9 kg/m² in the male participants ($P = 0.33$). Thirty-two of the participants reached the primary renal outcome of 40% decline in measured GFR during follow-up.

Baseline characteristics. The baseline characteristics of progressors and nonprogressors are shown in Table 1. Most characteristics were similar in the 2 groups, except for statistically higher mean diastolic blood pressure, fasting plasma glucose (FPG), HbA1c, GFR, and urine ACR in the progressors. Baseline and last measured GFRs for both groups are shown in Figure 2.

Baseline lipidomic profile of progressors and nonprogressors. We measured 406 lipids from 18 classes, including TAGs, diacylglycerols (DAGs), monoacylglycerols (MAGs), phosphatidylcholines (PCs), PEs, plasmalogen-PCs (pPCs), plasmalogen-PEs (pPEs), lyso-PCs (LPCs), lyso-PEs, FFAs, cholesteryl-esters (CEs), cardiolipins, phosphatidic acids (PAs), phosphatidylinositols (PIs), phosphatidylglycerols (PGs), phosphatidylserines (PSs), sphingomyelins (SMs), and ceramide phosphates in positive and negative ionization modes (Supplemental Table 1; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.130317DS1>). After combining the different adducts of the same feature and eliminating the classes consisting of 2 or fewer lipid molecules (PA, PG, PS, pPC, ceramide phosphate, and MAG), 236 unique lipids, including 16 FFAs (6.8%), 76 glycerolipids (32.2%), 83 phospholipids (35.1%), 12 CEs (5.1%), 20 SMs (8.5%), and 29 ACs (12.3%), were included in the analysis.

In progressors ($\geq 40\%$ decline in GFR), there was a higher relative abundance of longer TAGs with more double bonds ($P < 0.0001$) (Figure 3A). Conversely, in nonprogressors there was a lower relative abundance of longer TAGs with more double bonds at baseline ($P = 0.022$) (Figure 3A). The alterations in lipid abundance by carbon number or saturation status were independent of baseline GFR, ACR, weight, FPG, and HbA1c. Similar alterations in the relative abundance of TAGs were observed when all participants were divided based on whether they reached a sustained GFR less than 90 mL/min, less than 60 mL/min, or less than 30 mL/min with stronger trends with increased CKD severity ($P < 0.0001$) (Figure 3B). There were no significant alterations between progressors and nonprogressors in the levels of other complex lipids by differences in the number of carbons or double bonds.

Table 1. Comparing baseline characteristics and follow-up of progressors and nonprogressors

	Nonprogressors n = 60	Progressors n = 32	P value
Duration of follow-ups, years	8.8 ± 3.0	8.9 ± 3.0	0.932
Number of follow-up visits	10 ± 4	10 ± 3	0.827
Age, years	46 ± 9	43 ± 9	0.120
Male sex, n (%)	20 (33.3)	5 (15.6)	0.069
BMI, kg/m ²	36.2 ± 8.3	34.5 ± 8.0	0.355
Systolic blood pressure, mmHg	122 ± 15	127 ± 18	0.193
Diastolic blood pressure, mmHg	76 ± 9	81 ± 9	0.010
Fasting plasma glucose, mg/dL	207.2 ± 90.5	262.3 ± 108.7	0.011
HbA1c, %	9.0 ± 2.3	10.2 ± 1.9	0.012
Total cholesterol, mg/dL	167.0 ± 35.6	184.9 ± 47.4	0.071
Triglyceride, mg/dL	180.4 ± 136.6	273.6 ± 283.3	0.137
Total serum protein, g/dL	6.8 ± 0.4	6.8 ± 0.5	0.851
Serum albumin, g/dL	3.5 ± 0.4	3.4 ± 0.4	0.166
AST, IU/L	24.9 ± 20.3	23.6 ± 22.4	0.802
Serum alkaline phosphatase, IU/L	107.5 ± 32.8	118.6 ± 35.3	0.160
Total bilirubin, mg/dL	0.5 ± 0.2	0.4 ± 0.2	0.067
GFR, mL/min	149 ± 46	170 ± 51	0.038
Urine ACR, median [IQR], mg/g	19 [9–66]	77 [50–396]	<0.001
Intervention arm, n (%)	36 (60.0)	15 (46.9)	0.228
Medications			
Antihypertensive, n (%)	24 (40.0)	13 (40.6)	0.954
Insulin, n (%)	23 (38.3)	16 (50.0)	0.281
Oral hypoglycemic, n (%)	48 (80.0)	25 (78.1)	0.832
Statins, n (%)	10 (16.7)	9 (28.1)	0.196
Other lipid lowering, n (%)	12 (20.0)	10 (31.3)	0.228

Means with SDs are given for continuous variables. ACR, albumin/creatinine ratio; AST, aspartate aminotransferase; GFR, glomerular filtration rate; HbA1c, hemoglobin A1c.

Baseline AC profiles in progressors and nonprogressors. For the primary outcome of ≥40% GFR decline (Figure 3C), there was a significantly lower abundance of serum ACs with a longer chain length (C16–C20) in progressors ($P < 0.001$). Conversely, the nonprogressors exhibited a higher abundance of serum ACs with longer chain length ($P = 0.026$). These changes were inverse to the alterations observed in TAGs in progressors and nonprogressors. Similar changes in the abundance of longer chain ACs (C16–C20) were observed when all participants were divided based on whether they reached a sustained GFR less than 90 mL/min, less than 60 mL/min, or less than 30 mL/min (Figure 3D). Overall, the abundance of C16–C20 ACs was significantly lower in participants who reached the GFR threshold compared with those who did not ($P \leq 0.015$), suggesting impaired β -oxidation.

Baseline long chain/intermediate chain AC ratio correlation with chain length and saturation of other complex lipids. The data in Figure 3, A–D, suggest that the abundance of longer chain ACs correlates directly with shorter TAGs and inversely with longer TAGs. We also examined the correlation of the long/intermediate chain ACs with different principal components (Supplemental Table 2) characterized by their length and saturation status. We found a significant direct correlation between long/intermediate chain AC ratio and the short complex lipids with fewer double bonds in the TAG, pPE, PE, and DAG lipid classes (Figure 4). We also observed a significant inverse correlation between long/intermediate chain ACs and the longer chain complex lipids with more double bonds in the CE classes (Figure 4).

Alteration of ACs by categories of urinary ACR. The abundance of ACs in progressors and nonprogressors by change in number of carbons and double bonds was also compared when stratified by the baseline ACR category (Figure 5). The heatmap suggests a trend toward lower levels of long chain ACs (C16–C20) from nonprogressors with ACR less than 30 mg/g to progressors with ACR more than 300 mg/g. Within each ACR category, the change in AC abundance by increased carbon number reached statistical significance and increased in nonprogressors with ACR 30–299 mg/g ($P = 0.006$) and decreased in progressors with

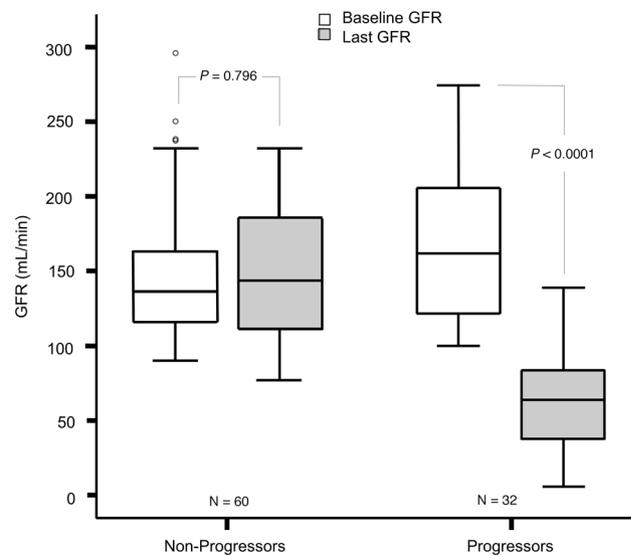


Figure 2. Baseline and last measured GFRs in 32 progressors and 60 non-progressors. The boxes represent median and IQR, and bars represent 1.5-fold \times the IQR below the 25th percentile and above the 75th percentile. Outliers outside the mean + 2 SD are shown with dots. Paired 2-tailed *t* test was used.

ACR more than 300 mg/g ($P < 0.0001$). Overall, the mean level of C16–C20 ACs in progressors with ACR more than 300 mg/g was significantly lower than in any other subgroup ($P \leq 0.004$), except nonprogressors with ACR more than 300 mg/g, where the difference did not reach statistical significance.

Risk prediction of DKD progression from lipid changes. We used the lipid principal components (Supplemental Table 2) to construct Cox regression models for risk prediction of DKD progression. We applied different models with varying levels of adjustment. First, we built the base model, which included the study principal components without adjusting for the clinical variables (Table 2, model 1: ACR + GFR). In this model, each 1 SD increase in unsaturated PEs was associated with 1.78-fold higher risk of progression (95% CI: 1.24–2.57, and $P = 0.002$), and each 1 SD increase in unsaturated FFAs was associated with 0.66-fold lower risk of progression (95% CI: 0.46–0.95, and $P = 0.026$) independent of other lipid factors. In the next step, we identified the principal components with a significant interaction term by ACR category and then added them to the base model (Table 2, model 2: lipids). By doing so, we observed strengthening of the significance of unsaturated PEs and unsaturated FFAs. However, we also noted that each 1 SD increase in C16–C20 ACs in participants with ACR less than 30 mg/g, compared with ACR at least 300 mg/g, was associated with a 3.36-fold higher risk of DKD progression (95% CI: 1.00–11.28, and $P = 0.05$). Similarly, each 1 SD increase in short-low-double-bond TAGs was associated with significantly lower risk of progression in participants with ACR less than 30 mg/g or ACR 30–299 mg/g compared with ACR at least 300 mg/g (Table 2, model 2: lipids). Further adjusting using baseline GFR, ACR, FPG, and HbA1c did not alter the results, and similar estimations were obtained in the fully adjusted model (Table 2, model 3: ACR + GFR + lipids).

Next, we compared the probabilistic risk scores of progressors and nonprogressors obtained from these 3 models. Model 3 (ACR + GFR + lipids) showed the largest difference between the risk score of progressors and nonprogressors, on average, compared with the other 2 models (Figure 6). When percentage with progression was compared by the risk score quartiles, the first quartile of the full model was more likely to consist of nonprogressors, and its fourth quartile was more likely to consist of progressors compared with the other 2 models (Table 3).

Differential network analysis. To understand changes in the potential interactions between lipids, we performed differential network analysis using sparse correlation networks in the progressor and nonprogressor subgroups. The differential network analysis revealed a significant loss of inter- and intraclass correlation of lipids in progressors compared with nonprogressors. This was observed most frequently in the TAG, PE, and AC subclasses and, to a lesser extent, in other subclasses (Figure 7). Overall, out of 55,460 possible permutations of bivariate correlations in 236 lipids, nonprogressors had 1028 significant edges while progressors had 547 ($P < 0.0001$). Among classes with at least 5 significant correlations with other lipids, there were 519 significant edges in nonprogressors and 287 edges in progressors, a 50% decline ($P < 0.0001$).

Systems integration of renal transcriptomics and serum lipidomics. Because a protocol kidney biopsy taken near the time of baseline serum collection was available, we had the unique opportunity to perform an integrative transcriptomic analysis using dissected glomerular and tubulointerstitial kidney tissue and serum lipids. The transcriptomic-lipidomic analysis in the glomerular compartment revealed significant correlations

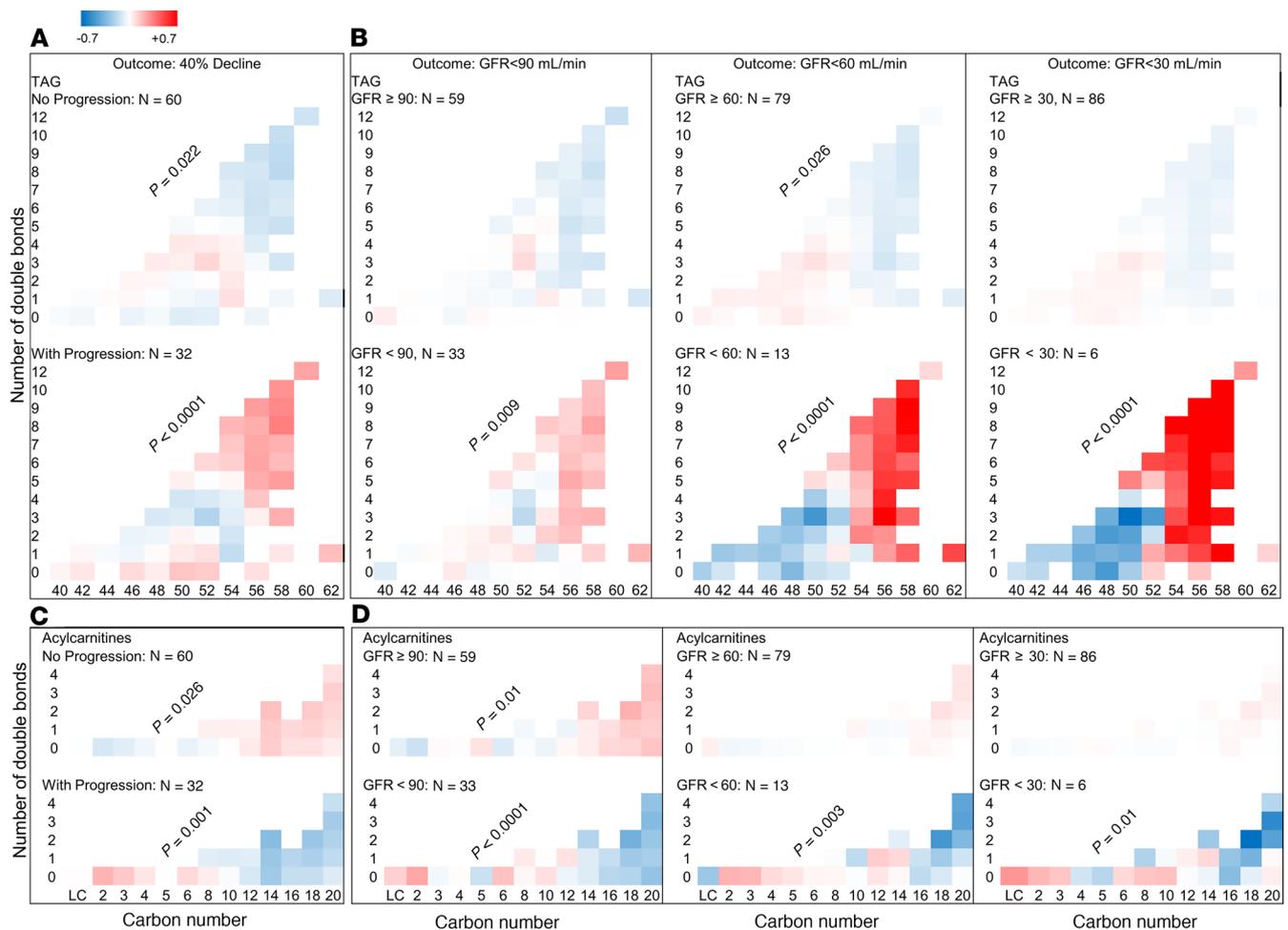


Figure 3. Differences in carbon chain length and number of double bonds in complex lipids and ACs between progressors and nonprogressors. Total $N = 92$ in all panels. Abundance of TAGs and ACs were measured at the baseline visit of this study and compared in progressors and nonprogressors. **(A)** In serum from progressors ($n = 32$), there was a greater relative abundance of longer TAGs with more double bonds. An opposite trend was observed in nonprogressors ($n = 60$). **(B)** A similar pattern was found when all participants (progressors + nonprogressors) were grouped based on whether they achieved a sustained GFR < 90 ($n = 33$), < 60 ($n = 13$), or < 30 mL/min ($n = 6$). Bonferroni's threshold for multiple comparisons was set to 0.0063 (0.05 divided by 8 panels/class). **(C)** In progressors ($n = 32$), there was a lower relative abundance of longer ACs with more double bonds. An opposite trend was observed in nonprogressors ($n = 60$). **(D)** A similar trend was noted in abundance of ACs by carbon number, when all participants were grouped based on whether they achieved a sustained GFR < 90 ($n = 33$), < 60 ($n = 13$), or < 30 mL/min ($n = 6$). Bonferroni's threshold for multiple comparisons was set to 0.0063 (0.05 divided by 8 panels/class). P values are products of testing abundance of lipid by "carbon number \times double bond" interaction term in progressors versus nonprogressors using mixed-linear models.

between lipid factors, including FAs, glycerolipids, and ACs, with the genes regulating their synthesis and metabolism (Supplemental Table 4). Specifically, a higher abundance of longer polyunsaturated DAGs was associated with higher gene expression of the acetyl-CoA carboxylase gene (*ACACA*; $P = 0.0062$, and FDR = 0.060). Conversely, the abundance of long chain ACs was inversely correlated with expression of fatty acid metabolism–encoding genes ($P \leq 0.0052$, and FDR ≤ 0.052). In the tubulointerstitial compartment, there were significant correlations between fatty acids, glycerolipids, phospholipids, and ACs and the expression of genes involved in the synthesis and metabolism of fatty acids, phospholipids, glycerolipids, and sterol lipids. There were a larger number of significant correlations between the genes responsible for de novo synthesis, desaturation, elongation, and β -oxidation of various serum fatty acids, glycerolipids, and long chain ACs (Supplemental Table 5). Specifically, a higher abundance of intermediate chain low-double-bond TAGs was associated with a higher *ACACA* expression ($P = 0.0042$, FDR = 0.009). Additionally, in the tubulointerstitial compartment, we noticed a significant direct correlation between fatty acid abundance in serum and the expression of genes responsible for de novo fatty acid biosynthesis (*ACSL5* and *LYPLA3* or *PLA2G15*) and their β -oxidation (*ACAD10*). We also observed a direct correlation between short-low-double-bond DAGs and unsaturated FFAs

Low-double bond shorter lipids:

High-double bond longer lipids:

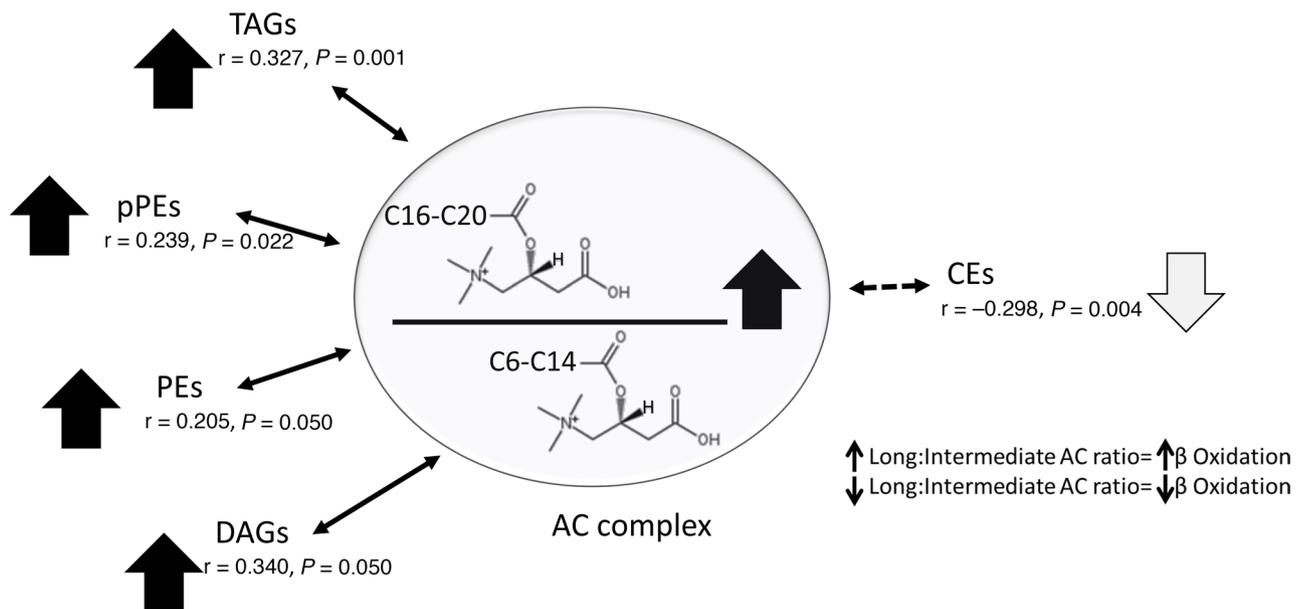


Figure 4. Adjusted correlation of long (C16-C20) to intermediate chain (C6-C14) AC ratio with complex lipids of various chain length and double bonds. Long/intermediate AC ratio is directly correlated with shorter complex lipids with fewer double bonds but is inversely correlated with longer complex lipids with more double bonds in the CE class. P values are products of testing partial correlation coefficients in multiple linear regression models adjusting for sex and ACR.

and genes responsible for elongation (*ELOVL1* and *ELOVL5*) and desaturation (*DEGSI*). There were inverse correlations between high-double-bond TAGs and saturated FFAs and genes coding elongation (*ELOVL1* and *ELOVL5*) and between short-low-double-bond DAGs and TAGs, unsaturated FFAs, and long chain ACs and genes regulating β -oxidation (Supplemental Table 5). Ingenuity Pathway Analysis based on the top lipid transcripts revealed enrichment of G protein-coupled signaling pathways as regulators of de novo lipogenesis in both glomerular (Figure 8A) and tubulointerstitial compartments (Figure 8B). The genes involved in lipid metabolism downstream of the nuclear hormone-activated receptors peroxisome proliferator-activated receptor γ (*PPARG*) and peroxisome proliferator-activated receptor α (*PPARA*) were expressed with a high degree of similarity in both glomerular (Figure 8C) and tubulointerstitial compartments (Figure 8D), suggesting similar lipogenic processes are activated in both compartments by these upstream regulators.

Discussion

We identified potentially novel, previously unrecognized lipid predictors of progressive DKD in diabetic American Indians with preserved kidney function, as indicated by GFR at least 90 mL/min. Lipid factors representative of unsaturated FFAs and PEs, short-low-double-bond TAGs, and long chain ACs predicted DKD progression in these patients. A transcriptomic-lipidomic integrative analysis revealed a significant correlation between serum FFAs and genes regulating de novo fatty acid synthesis in the glomerular compartment. Similarly, the integrative analysis showed a significant correlation between serum FFAs, glycerolipids, and ACs and genes regulating de novo fatty acid synthesis, desaturation, elongation, and β -oxidation. Importantly, the gene *ACACA* was directly associated with long polyunsaturated DAGs in the glomerular compartment and with low-double-bond intermediate-length TAGs in the tubulointerstitial compartment. The differential networks identified in the present study reflect a differential interactome among various lipids by status of DKD progression.

The results of this study carry significant clinical implications. While tubular markers have been proposed as indicators of early renal function decline, there are no known lipid predictors of DKD progression at a stage when GFR is preserved and other patient characteristics are noninformative. The proposed panel accurately predicts DKD onset and progression and, therefore, may have clinical applicability for risk stratification in patients with diabetes. Second, the lipid markers unravel a potentially novel mechanism underlying

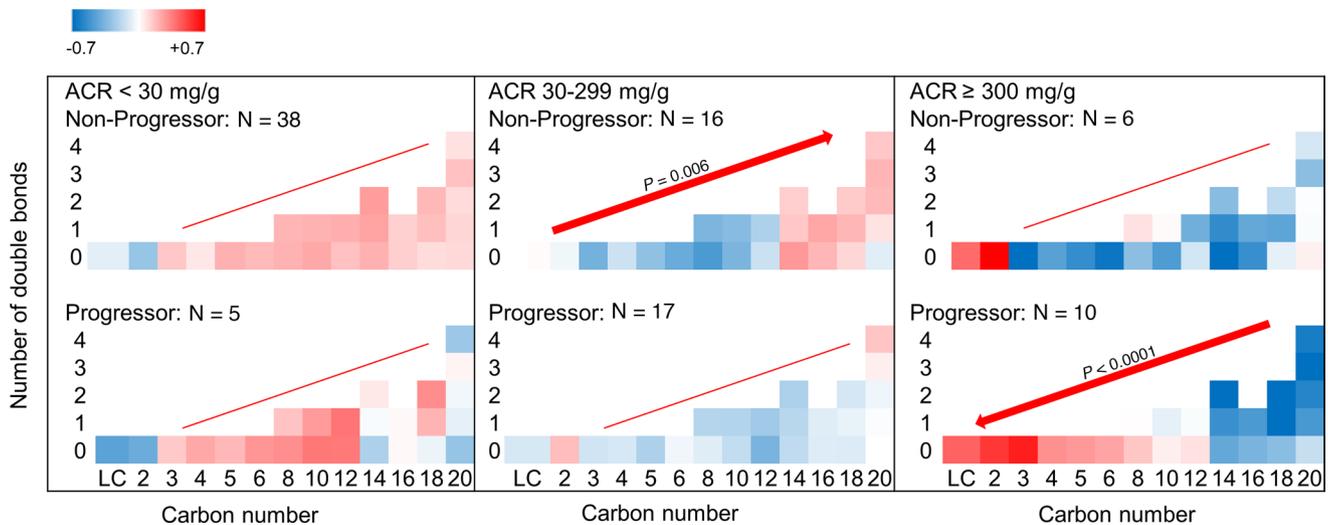


Figure 5. AC alterations by categories of baseline ACR in progressors and nonprogressors. ACs of various chain length and double bonds were quantified in serum at the baseline visit of progressors and nonprogressors. Distribution of various ACs by ACR category revealed lower levels of C16–C20 ACs from nonprogressors with normoalbuminuria (upper left) to progressors with ACR more than 300 mg/g (lower right). Within each ACR category, long chain AC abundance (C16–C20) increased in nonprogressors with ACR 30–299 mg/g ($P = 0.006$) and decreased in progressors with ACR more than 300 mg/g ($P < 0.0001$). Sample size 43 in ACR less than 30, 33 in ACR 30–299, and 16 in ACR at least 300 mg/g. Bonferroni's threshold for multiple comparisons was set to 0.0083 (0.05 divided by 6 panels). P values are products of testing abundance of lipids by “carbon number \times double bond” interaction term in progressors versus nonprogressors using mixed-linear models.

DKD progression: impaired mitochondrial β -oxidation and altered complex lipid remodeling. Importantly, the ability to perform integrative renal gene expression with serum lipids has uncovered acetyl-CoA carboxylase (ACC) activation as a critical determinant. Consequently, ACC inhibition might be a therapeutic strategy to prevent DKD or slow its progression. Not all patients with diabetes develop DKD; therefore, the possible contribution of ACC polymorphisms to renal outcomes needs to be investigated. Furthermore, in an earlier study of this cohort, progressors were shown to have a higher mesangial fractional volume, percentage of global glomerular sclerosis, nonpodocyte cell number per glomerulus, glomerular basement membrane width, mean glomerular volume, and podocyte foot process width. However, they had lower glomerular filtration surface density, with fewer endothelial fenestrations after adjustment for baseline age, sex, duration of diabetes, HbA1c, GFR, and treatment assignment during the clinical trial (26). Although such early histological alterations provided a readout of future decline in renal function, lipid alterations provide a noninvasive readout of similar long-term outcomes and demonstrate potential for early-stage risk stratification in clinical practice.

There are a number of similarities in the results of this study and our previous lipidomic study in the CRIC (19). However, a few differences between these 2 studies are worth noting. At baseline examination, participants of the CRIC had CKD stage 3A or 3B, whereas in this study all were at stage 1 (GFR > 90 mL/min). Only 50% of the CRIC participants had type 2 diabetes, whereas every participant in this study had type 2 diabetes. FFAs and ACs were not measured in the CRIC, nor were kidney biopsies routinely obtained as part of research data collection. Despite these differences, we noted a higher risk of DKD progression at higher levels of unsaturated PEs in American Indians with diabetes, which aligns with enrichment of PE as a class and its higher abundance in CKD progressors in the CRIC study (19). In a metabolomic investigation in participants of the African American Study of Kidney Disease and Hypertension and the Modification of Diet in Renal Disease, a significant enrichment of PEs was noted among the top metabolites associated with proteinuria in patients with CKD (35). Aligned with these observations, a diabetic mouse model showed accumulation of the Amadori-PE species in tubular cells (36). The mechanistic link between PE and DKD progression has yet to be identified. However, PEs have diverse cellular functions, including oxidative phosphorylation, mitochondrial biogenesis, and autophagy. PEs are PC precursors, are substrates for posttranslational modifications, influence membrane topology, and promote cell and organelle membrane fusion (37). Cell culture experiments also suggest that exposure to PE alters the plasma membrane bilayer with total loss of asymmetrical aminophospholipids and promotes apoptosis (38) and autophagy (39).

Table 2. Hazard ratios of lipid predictors of DKD progression

Predictors	Hazard rate ratio	95% CI	P value	FDR
Model 1 (base model)				
Unsaturated PEs (1 SD)	1.78	1.24–2.57	0.002	0.025
Unsaturated FFAs (1 SD)	0.66	0.46–0.95	0.026	0.050
Model 2 (model 1 + AC, TAG interaction with ACR)				
Unsaturated PEs (1 SD)	2.36	1.56–3.58	<0.001	0.01
Unsaturated FFAs (1 SD)	0.59	0.41–0.84	0.004	0.03
C16–C20 AC (1 SD change in ACR <30)	3.36	1.00–11.28	0.05	0.05
C16–C20 AC (1 SD change in ACR 30–299)	2.39	0.84–6.49	0.103	0.06
SLDB TAG (1 SD change in ACR <30)	0.25	0.09–0.66	0.005	0.04
SLDB TAG (1 SD change in ACR 30–299)	0.19	0.07–0.52	0.001	0.02
Model 3 (model 2 + GFR + ACR)				
Unsaturated PEs (1 SD)	2.57	1.66–3.98	<0.001	0.008
Unsaturated FFAs (1 SD)	0.54	0.36–0.79	0.002	0.023
C16–C20 AC (1 SD change in ACR <30)	4.07	1.03–16.06	0.045	0.053
C16–C20 AC (1 SD change in ACR 30–299)	2.44	0.78–7.62	0.125	0.060
SLDB TAG (1 SD change in ACR <30)	0.16	0.05–0.52	0.002	0.025
SLDB TAG (1 SD change in ACR 30–299)	0.14	0.05–0.41	<0.001	0.015
GFR (mL/min)	1.009	1.002–1.017	0.015	0.045
ACR (mg/g)	1.001	1.000–1.001	0.006	0.038

Decline in GFR: 40%. Model 1, or base model, shows independent lipids without adjustment with baseline ACR and GFR. Model 2 is model 1 plus an interaction term of C16–20 AC and SLDB TAGs by ACR category with ACR more than 300 being the reference category. Model 3 is model 2 plus baseline GFR and ACR independent of HbA1c and fasting plasma glucose. GFR, iothalamate-based glomerular filtration rate; SLDB TAG, short-low-double-bond triacylglycerol.

Another finding in our study was a lower risk of DKD progression with higher unsaturated FFA abundance. This finding is aligned with earlier studies that illuminated salutary effects of longer unsaturated FFAs, such as oleate and eicosapentaenoic acid, on inflammation, endoplasmic reticulum stress, and eventually better podocyte and pancreatic cell survival compared with the effect of saturated FFA palmitate (40–43). In a cross-sectional analysis of the CPROBE study, we showed that at more advanced CKD stages, there was a significantly greater abundance of longer polyunsaturated TAGs and a lower abundance of C16–C20 ACs (18). Patients at CKD stage 5 thus had the highest abundance of longer polyunsaturated TAGs and the lowest abundance of C16–C20 ACs. Replication of this pattern was seen in this study in DKD progressors, but at a much earlier stage, when GFR was still above 90 mL/min. In the CRIC study, although DAGs as a class were enriched in CKD progressors, TAGs were not significantly different (19), a finding inconsistent with CPROBE and the current study. A likely explanation is the narrow range of baseline GFRs in CRIC participants compared with the CPROBE study, whose patients were at all stages, and the marked heterogeneity in CRIC enrollees compared to the homogeneity of the American Indian cohort. TAGs differ in their renal toxicities by their acyl length and number of double bonds, and in part by the characteristics of their acyl constituents, especially when they are nonesterified (before use in the construction of TAGs or after their lipolysis). Overall, the main renal toxicities are determined by saturated fatty acids, such as palmitate, through mechanisms that involve activation of AMP-activated protein kinase (AMPK) and mammalian target of rapamycin complex-1 (mTOR1) signaling pathways. Furthermore, they can promote insulin resistance, mitochondrial superoxide generation, and endoplasmic reticulum stress. Other mechanisms of toxicity include impairment of the podocyte actin cytoskeleton, induction of autophagy, and eventual triggering of apoptosis and cell death (40, 42, 44–49). Within the tubulointerstitial compartment, palmitate upregulates the monocyte chemoattractant protein-1, leading to intracellular activation of DAG followed by protein kinase C (PKC), which further promotes renal tubular cell injury (43). Abundance of saturated fatty acids (specifically palmitate) upregulates the elongation and desaturation of shorter FFAs and their incorporation in the construct of TAGs. The longer polyunsaturated TAGs have a higher melting temperature with greater fluidity and thus are less toxic than the nonesterified saturated fatty acids. Therefore, although they served as a surrogate marker of DKD progression, their higher abundance in

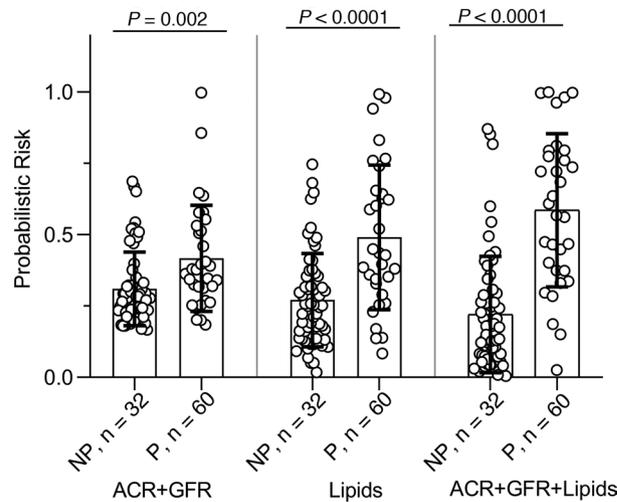


Figure 6. Predicting DKD progression with probabilistic risk scores. Probabilistic risk scores derived from odds of progression by 3 models were compared. Model 1 incorporated baseline ACR and GFR (ACR + GFR), model 2 incorporated independent lipid factors predicting progression (lipids), and model 3 consisted of the lipids plus baseline ACR and GFR (ACR + GFR + lipids). Progressors ($n = 32$) had a higher probabilistic risk scores compared with nonprogressors ($n = 60$) in all models, and the largest score was noted in model 3, when lipids were included with baseline ACR and GFR. Bar graphs show mean and error bars are 1 SD above and below the mean. A2-tailed t test for independent variables was used.

progressors likely signifies upregulation of adaptive compensatory processes (elongation, desaturation) to convert more toxic lipids (saturated nonesterified fatty acids) into less toxic lipids (polyunsaturated long TAGs). The progressors were also characterized by lower abundance of shorter TAGs with fewer double bonds. This probably reflects a lower abundance of fatty acids with 14 or more carbons, which become the dominant mitochondrial substrates in progressors whose mitochondrial shuttling of fatty acids with 16 or more carbons is impaired because of carnitine palmitoyltransferase I (CPT1) inhibition (50). This response leads to a relatively lower abundance and thus lower likelihood of their incorporation into corresponding (short) TAGs. Therefore, while the high abundance of longer polyunsaturated TAGs signifies upregulation of elongation and desaturation as compensatory mechanisms, the lower abundance of shorter and saturated TAGs reflects low abundance of the corresponding nonesterified fatty acids (≤ 14 C), which indicates impaired β -oxidation of fatty acids with 16 or more carbons. When stratified by levels of albuminuria, progressors with normal albumin excretion (ACR < 30 mg/g) had significantly higher AC levels compared with progressors with overt proteinuria (ACR > 300 mg/g), suggesting that β -oxidation may be upregulated as a compensatory mechanism in the early stages of DKD but that β -oxidation becomes impaired over time with progression to overt proteinuria. These observations suggest that the lipid phenotype associated with advanced DKD may be a continuum that begins mechanistically at an early stage, allowing early discrimination of progressors from nonprogressors and risk stratification during early kidney disease when no other prognostic tools are available.

The Ingenuity Pathway Analysis (QIAGEN Inc.; www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/) highlights the enrichment of the G protein signaling pathways in both glomerular and tubulointerstitial compartments. G proteins promote transcriptional activation of metabolic genes by carbohydrate-response element-binding protein and regulate downstream lipogenic genes, including *ACACA*, fatty acid synthetase, acyl CoA synthase, and glycerol phosphate acyl transferase (51). In particular, the significant association of *ACACA* gene expression with DAGs in the glomerular and TAGs in the

Table 3. Proportion of progressors by quartiles of the probabilistic risk scores calculated by various models

	Quartile 1 <i>n</i> = 23	Quartile 2 <i>n</i> = 23	Quartile 3 <i>n</i> = 23	Quartile 4 <i>n</i> = 23	<i>P</i> value (trend)
Model 1 (ACR + GFR)	13.04 (7.18)	21.74 (8.794)	52.17 (10.65)	52.17 (10.65)	0.001
Model 2 (lipids)	17.39 (8.081)	17.39 (8.081)	43.48 (10.569)	60.87 (10.405)	0.0005
Model 3 (ACR + GFR + lipids)	4.35 (4.348)	8.7 (6.007)	47.83 (10.65)	78.26 (8.794)	<0.0001

The third and fourth quartiles of the risk score model 1 (ACR + GFR) and the first and the second quartiles of model 2 (lipids) were not discriminatory. However, the most discrimination was obtained in model 3, when lipids were combined with baseline ACR and GFR. Values are percentage of progressors and SEM. *n* each quartile = 23.

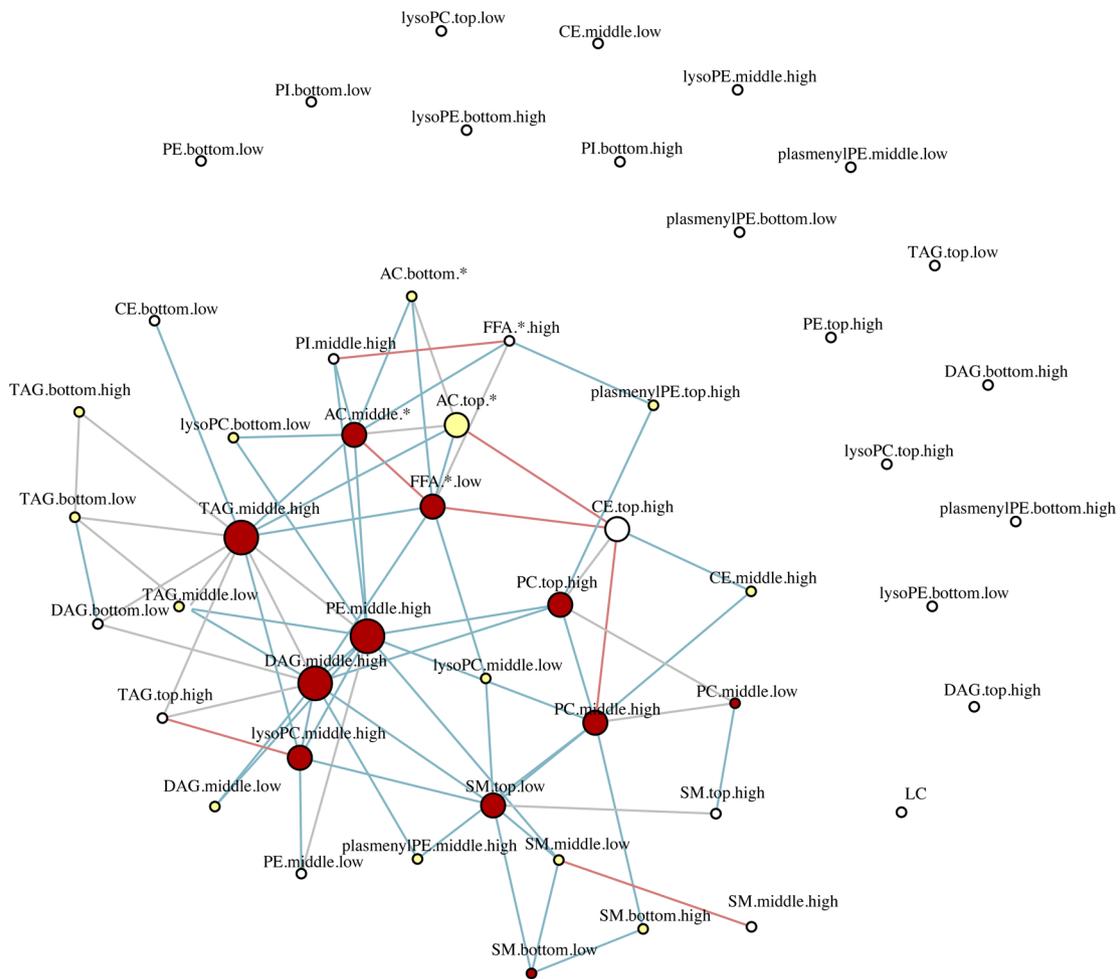


Figure 7. Differential network analysis. To identify inter- and intraclass lipid correlates, we obtained the sparse partial correlation networks that captured the interdependencies between lipids. We used lipid grouping structure, obtained the superset of the network skeleton, and finally obtained the final stable network structures, the latter based on a bootstrapping method. Differential network analysis revealed differential loss of edges between various lipid classes in progressors characterized by 547 significant edges versus 1028 in nonprogressors ($P < 0.0001$) out of 55,460 possible permutations of bivariate correlations. The lines represent significant edges that were exclusively observed in nonprogressors (shown in blue) or progressors (shown in red). Common edges are shown in gray. The node size is proportional to the number of connectivity levels within and across lipid subclasses, and node colors represent number of cross-class connections (white, low; yellow, middle; red, high). Nodes are categorized by chain length (bottom, middle, top) and double bonds (low, high), with details shown in Supplemental Table 3.

tubulointerstitial compartments supports the hypothesis that the observed differential serum lipids may be regulated, in part, by renal ACC. The direct correlation of shorter low-double-bond DAGs with elongase and desaturase imply that a higher abundance of such lipids may upregulate the renal cortical elongation and desaturation of fatty acids as a compensatory mechanism that converts shorter and relatively more saturated, toxic lipids into longer, relatively more unsaturated, less toxic products. On the other hand, the direct correlation of saturated FFAs with *ACAD10*, a β -oxidation regulatory gene, suggests they may upregulate β -oxidation, at least in early-stage DKD, a process that may be reversed with DKD progression or by unsaturated FFAs because of their inverse correlation with *ACAT2*, or β -oxidation regulatory gene. Altogether, the differential lipid alterations associated with DKD progression suggest accelerated renal de novo lipogenesis and impaired β -oxidation (18).

Acceleration of de novo lipogenesis is a major contributor to dyslipidemia in insulin resistance states, such as in type 2 diabetes (52). De novo lipogenesis of long chain fatty acids is dependent on the activity of 3 cytosolic enzymes: ATP citrate lyase, ACC, and fatty acid synthetase (53). ACC is a key regulatory enzyme in lipogenesis, and its activity is under hormonal regulation by insulin, glucagon, and epinephrine. During insulin resistance, increased insulin promotes ACC activation via its protein phosphatase-mediated dephosphorylation (54). A recent expression quantitative trait loci (eQTL)

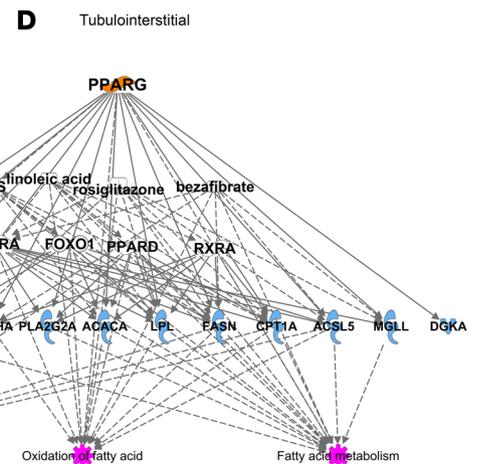
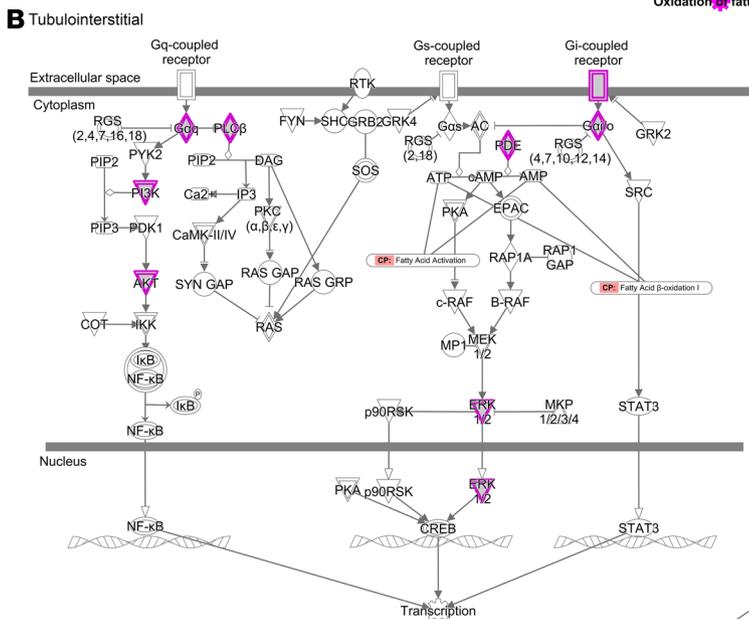
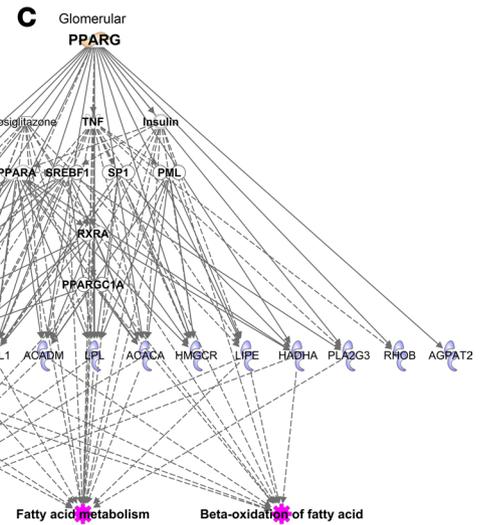
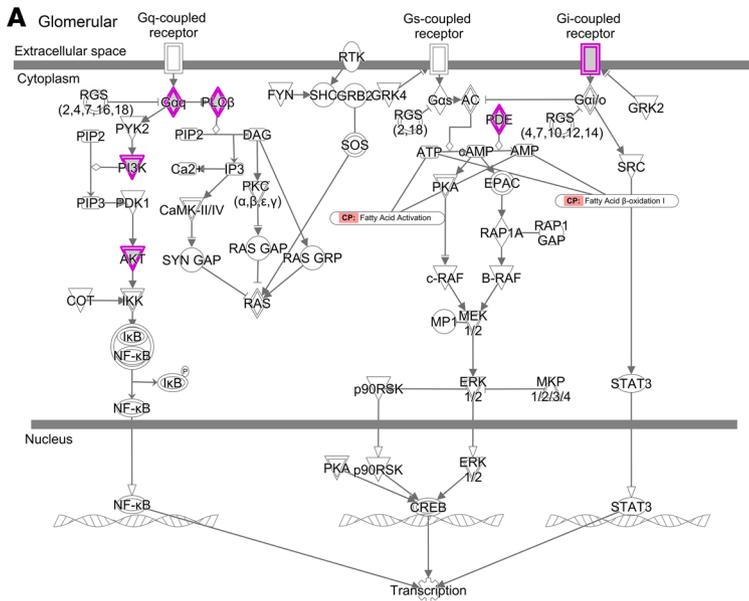


Figure 8. Integrative transcriptomic-lipidomic analysis identifies G protein-coupled signaling pathways and nuclear hormone-activating receptors in regulation of fatty acid synthesis and β -oxidation. Ingenuity Pathway Analysis reveals enrichment of G protein signaling pathways involved in regulation of NF- κ B, CREB, and STAT3 in glomerular (A) and tubulointerstitial compartments (B). CREB is a transcriptional regulator of de novo lipogenesis. Genes regulating the intermediaries highlighted in purple in panels A and B are significantly correlated with the corresponding serum lipids. The genes downstream of nuclear hormone-activating receptor PPAR γ known to regulate fatty acid metabolism and their β -oxidation in both glomerular (C) and tubulointerstitial compartments (D) are significantly correlated with serum lipids identified via the lipidomic analysis. NF- κ B, nuclear factor κ -light-chain-enhancer of activated B cells; CREB, cAMP response element-binding protein; PPAR γ , peroxisome proliferator-activated receptor γ ; STAT3, signal transducer and activator of transcription 3.

analysis of participants of the Nephrotic Syndrome Study Network revealed a significant differential expression of G/C *ACACA* variant in both glomerular and tubulointerstitial compartments of the participants (55), suggesting the possibility of *ACACA* polymorphism on ACC activity. ACC activation inhibits CPT1, lowering cytosolic conversion of long chain acyl CoA to long chain AC, decreases substrate for the carnitine shuttle, impairs β -oxidation of long chain fatty acids, and increases cytosolic abundance of palmitate (53). Increased palmitate promotes elongation and desaturation, as evidenced by upregulation of stearoyl-CoA desaturase (*SCD*) 1 and 2 (Figure 9 and ref. 56). It also exerts its deleterious effects on podocytes and tubulointerstitial cells by upregulating AMPK and mTOR1 signaling, intracellular serial DAG and PKC activation, induction of mitochondrial superoxide generation, endoplasmic reticulum stress, and, eventually, promotion of autophagy, apoptosis, and cell death (40, 42, 43, 46–49).

Based on these data, we propose a model to account in part for these serum lipidomic changes that predict DKD progression (Figure 9). Upregulation of renal (and possibly liver) ACC in the diabetic milieu enhances de novo lipogenesis, which generates excess saturated fatty acids (e.g., palmitate). Although kidney ACC expression is modified, the liver may also influence circulating levels of fatty acids and complex lipids. With elongation and desaturation, palmitate is converted to longer unsaturated fatty acids, which are incorporated into complex lipids (e.g., glycerolipids). In addition, ACC upregulation can inhibit CPT1, which impairs β -oxidation. Finally, elevated FFAs present in the diabetic state can exacerbate mitochondrial dysfunction, especially in the setting of impaired fatty acid oxidation. The net effect will be a pattern of complex lipid remodeling and diminished fatty acid β -oxidation, as observed in our studies.

This study has several strengths, including the longitudinal study design, rigorous quality control with good reproducibility metrics in a robust liquid chromatography-MS (LC/MS) lipidomic platform, and excellent phenotyping of a well-characterized cohort of type 2 diabetes patients. Yearly follow-ups and iothalamate-based GFR measurements provided accurate evaluation of kidney outcomes. A kidney biopsy near the time of baseline serum collection permitted a potentially novel integrative lipidomic-transcriptomic analysis. The availability of data and samples before the onset of clinical DKD provided the opportunity to identify early prognostic DKD markers.

This study also has limitations. Although existing knowledge of the de novo lipogenesis pathway, along with the results obtained from the transcriptomic-lipidomic integrative analysis, support the hypothesis that ACC activation may be responsible, in part, for the differential lipid pattern, our results cannot distinguish whether the increased renal ACC expression is a cause or a consequence of the differential plasma lipid profile. In a recent study, we assessed individual lipid features present in plasma from control and diabetic *db/db* mice (57). Plasma lipid levels and DAG saturation status correlated with the corresponding lipids in the kidney in diabetic mice compared with control, suggesting that plasma DAG metabolism may be reflected in the kidney. However, this relationship is not known in humans and technically not feasible to assess given the difficulty in obtaining an adequate kidney biopsy specimen for lipidomic analysis. The study sample size is small, and replication of the study, including the risk prediction model, in larger cohorts and in other racial/ethnic groups is warranted. Like any other omic research, this study has a large number of variables because of the high-throughput data generated by the lipidomics platform. We applied a number of strategies to minimize the need for multiple testing, including application of mixed-linear models to study alterations at the group level versus individual lipids, reduction of the data to a smaller number of principal components for downstream analysis, partial correlation-based sparsing techniques for the study of the differential networking, and FDR correction for multiple testing in lipidomic-transcriptomic integrative analysis. Although we observed a significant interaction between categories of albuminuria in the risk prediction model, with advancing DKD, the differential lipid alterations were independent of proteinuria in

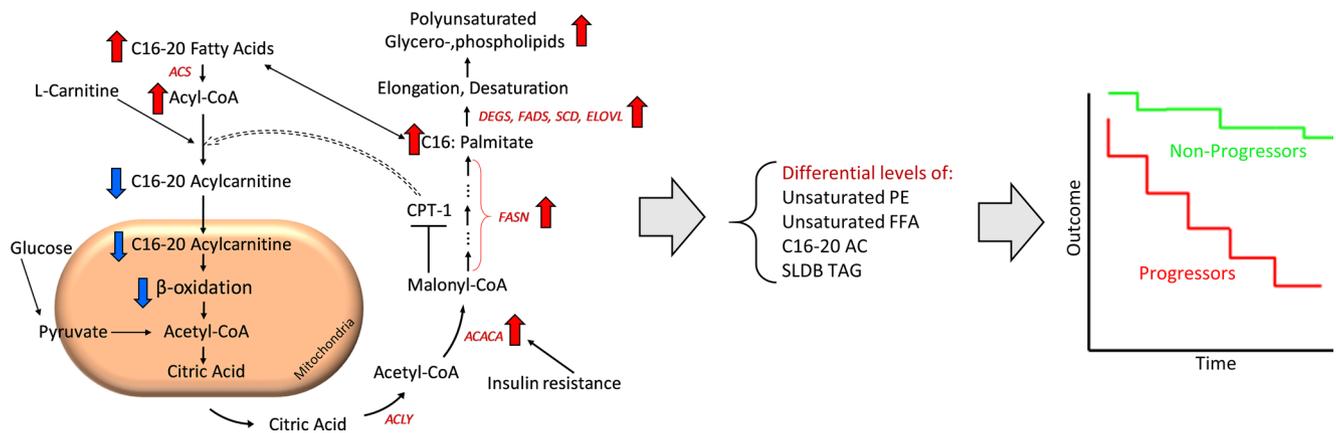


Figure 9. Proposed mechanisms underlying lipid abnormalities that predict early renal function decline in DKD. Upregulation of ACC, mediated by insulin resistance, enhances de novo lipogenesis characterized by increased abundance of palmitate, a C16 fatty acid. With elongation and desaturation, palmitate is converted into longer unsaturated fatty acids, which are incorporated into complex lipids (e.g., glycerolipids). In concert, upregulation of ACC also inhibits CPT1, which in turn decreases the conversion of L-carnitine to C16–C20 ACs. C16–C20 ACs are efficient β -oxidation substrates, and, therefore, their diminished mitochondrial transfer downregulates β -oxidation. The net effect of upregulated de novo lipogenesis is characterized by higher abundance of longer chain polyunsaturated glycerolipids and lower abundance of C16–C20 ACs and shorter low-double-bond glycerolipids. *ACACA*, acetyl-CoA carboxylase alpha; *ACLY*, ATP citrate lyase; *ACS*, acetyl-CoA synthetase; *CPT*, carnitine palmitoyltransferase; *DEGS*, delta 4-desaturase; *ELOVL*, elongation of very long chain fatty acids; *FASN*, fatty acid synthase; *FADS*, fatty acid desaturase; *SCD*, stearoyl-CoA desaturase; *SLDB TAG*, short-low-double-bond triacylglycerol.

our prior cross-sectional study (18). Routinely, frequency of follow-up declines over time, and participants with worse kidney function tend to attend research examinations less often than participants with preserved kidney function, raising the possibility of differential censoring by outcome (29, 30). However, in this particular subset of participants, the duration and number of follow-up visits were almost equal in progressors and nonprogressors, so the potential for ascertainment bias because of differences in follow-up is low. The current study is also limited by its sample size to investigate *ACACA* polymorphism despite evidence for *ACACA* polymorphism in other renal eQTL studies (55).

This study has important clinical implications and illustrated that a panel of lipid biomarkers may predict decline at early stage, when kidney function is still preserved (GFR > 90 mL/min). It provides a valuable opportunity for early-stage risk stratification when there is a paucity of reliable biomarkers at early stage. The proposed underpinning mechanism suggests targeted interventions at early stage. Inhibition of kidney fibrosis by modulating fatty acid receptors (58) provides further encouraging evidence for the success of such interventions. We conclude that lipid alterations that typify advanced DKD are present before the onset of clinical DKD and are characterized by higher abundance of unsaturated PEs, and longer polyunsaturated TAGs, but lower abundance of unsaturated FFAs and C16–C20 ACs. In patients with GFR more than 90 mL/min, a panel of lipids consisting of unsaturated FFAs and PEs, short-low-double-bond TAGs, and long chain ACs predicts the onset and progression of DKD. The underlying mechanism may depend on impairment of fatty acid β -oxidation and renal ACC activation, thus providing a potential therapeutic target (Figure 9).

Methods

Patients. Details of the study population and participant recruitment are published elsewhere (31).

Sample preparation and MS. We applied previously published methods for sample preparation (18, 19, 59). In brief, we used the modified Bligh-Dyer method for lipid extraction. We retrieved 10 μ L of serum and added water/methanol/dichloromethane at room temperature with 2:2:2 volume ratio, followed by spiking internal standards PC 17:0/17:0, LPC 17:0, PG 17:0/17:0, PE 17:0/17:0, TAG 17:0/17:0/17:0, SM 18:1/17:0, MAG 17:0, DAG 16:0/18:1, CE 17:0, ceramide phosphate 18:1/17:0, PA 17:0, PI 17:0/20:4, and PS 17:0/17:0. After collecting the organic layer, we dried the extracts under nitrogen and reconstituted them by adding 100 μ L of acetonitrile/water/isopropyl alcohol (10:5:85) followed by 10 mM ammonium acetate (NH_4OAc). Then we subjected the extracts to LC/MS, utilizing ABSciex quadrupole time of flight 5600 mass spectrometer equipped with a Turbo V ion source (AB Sciex, Concord, Ontario, Canada) and

Shimadzu CTO-20A Nexera X2 UHPLC with ACQUITY UPLC HSS T3 1.8- μ m column (Waters, Milford, Massachusetts, USA). ACs were quantified by LC/MS using an Agilent 6410 Triple Quadrupole tandem mass spectrometer (Agilent, Santa Clara, California, USA) with a targeted method described previously (60).

Quality control. A pool of study samples was injected at the beginning and after every 20 MS runs in the lipidomic study and after every 15 MS runs in the AC study to assess the stability of the measures over time and to identify any batch effects (Supplemental Figures 1 and 2).

Transcriptomic analysis. We manually microdissected kidney biopsy specimens to isolate glomerular and tubulointerstitial tissue. Glomerular and tubular gene expression profiling was performed as described previously on Affymetrix Gene Chip Array Human Genome series U133A and Plus 2.0 (Affymetrix, Inc., Santa Clara, California, USA) (61, 62). Cell files were processed, normalized (Robust Multi Array method), and batch corrected (COMBAT) on the R statistical platform. We used Human Entrez Gene custom CDF from Brainarray for annotations (<http://brainarray.mbni.med.umich.edu>). Normalized and \log_2 -transformed expression profiles were used in all the downstream analyses.

Statistics. We applied mean \pm SD or frequency (percentage) for description of normally distributed continuous and categorical variables, respectively. Median and IQR were used to describe non-normally distributed variables. To compare the baseline characteristics of progressors and nonprogressors, we used the 2-tailed *t* test for normally distributed continuous variables, the Kolmogorov-Smirnov test for skewed continuous variables, and the χ^2 test for categorical variables. We used the relative abundance of the peak intensities for the downstream analysis. To prepare the lipidomics data for downstream analysis, the batch normalized raw peak intensities were sum normalized by lipid members within each lipid subclass, logit transformed, and Z-score standardized (18). We used principal component analysis to reduce the number of lipids of the TAG, DAG, PE, pPE, CE, FFA, and AC classes to subclasses to generate secondary variables representative of various chain lengths within these classes (Supplemental Table 2). Long chain AC (C16–C20) to intermediate chain AC (C6–C14) ratio was applied as a marker of efficiency of β -oxidation (63). Multiple linear regression analysis was applied to explore the relationships between the long chain/intermediate chain AC ratio (predictor variable) with the resulting principal components (secondary variables) of complex lipids. We used generalized linear mixed models to test the intraclass alteration in lipid abundance (standardized, logit transformed, normalized lipid intensities used as the dependent variable) by change in carbon number and number of double bonds (saturation status) as the predicting variables.

We applied a Cox proportional hazard model to identify the independent lipid predictors of DKD progression and to estimate the risk associated with their change. Violation of proportional hazard assumption was ruled out by confirming the random distribution of the scaled Schoenfeld residuals, also known as partial residuals, of the predicting variables over time. We adjusted the models by BMI, FPG, HbA1c, and ACR followed by elimination of nonsignificant covariates from the model. The probability of progression by various panels was estimated using the logit score of the corresponding predictors in each panel and their conversion to probabilistic risk scores (64). Risk score calculation was performed separately for the baseline ACR and GFR, the lipids proposed by the Cox model, and the proposed lipids plus baseline ACR and GFR. The risk analysis was performed separately for the baseline model (model 1: ACR + GFR), the model incorporating lipid components with a significant interaction term by ACR categories (model 2: lipids), and the model adjusted further by baseline GFR, ACR, FPG, and HbA1c (model 3: ACR + GFR + lipids).

Because baseline TAG levels were different between progressors and nonprogressors, the glycerolipid factors and their upstream regulators (FFAs and ACs) were tested for their correlation with corresponding renal tissue transcripts. We used MetScape to map lipid identifiers to gene symbols responsible for lipid metabolism (65). The \log_2 Z-score standardized values of the corresponding genes were calculated, aimed at finding their lipid correlates. Then, the Z-score standardized values of the principal components derived from the glycerolipids (TAGs and DAGs), FFAs, and ACs were correlated with the transcriptomic profiles for the mapped genes using Pearson's correlations. We used matching samples from both compartments to run the correlation analysis. Ingenuity Pathway Analysis based on the top lipid transcript significant correlations was applied to identify the corresponding enriched pathways. Benjamini-Hochberg procedure was applied to minimize the FDR (66).

For differential network analysis aimed at identifying differential lipid correlates (inter- and intra-lipid class) by progression, we obtained the sparse partial correlation networks that capture interdependencies between lipids for the nonprogressor and progressor participant groups. We utilized the following approach

that leveraged (a) similarity information among lipids that is concordant with data obtained from a diabetic mouse model (Supplemental Table 3) and (b) the assumption that many interconnections, especially across lipid classes, are not present in both groups and hence we can use all 92 samples to estimate them more robustly. Throughout the technical developments, P denotes the total number of lipids under consideration, $n_{\text{nonprogressor}}$ and $n_{\text{progressor}}$ denote the sample sizes for nonprogressors and progressors, respectively, while x_{ij}^k denotes the relative abundance of lipid j for sample i in condition k (nonprogressor, progressor). Further, we arrange the data x_{ij}^k s in matrix form \mathbf{X}^k as:

$$\mathbf{X}^k := \begin{bmatrix} x_{11}^k & \cdots & x_{1p}^k \\ \vdots & \ddots & \vdots \\ x_{n_k 1}^k & \cdots & x_{n_k p}^k \end{bmatrix}, \quad k \in \{\text{non-progressor}, \text{progressor}\} \text{ (Equation 1)}$$

and let \mathbf{X}_j^k denote its j th column.

The main steps of the proposed estimation procedure include obtaining the lipid grouping structure, the superset of the network skeleton, and the final stabilized network structure. To obtain lipid grouping structure (step 1), we constructed groups among lipids based on their correlation structure using spectral clustering in progressors and nonprogressors (67) and denoted the collections of groups as $G_{\text{progressor}}$ and $G_{\text{nonprogressor}}$, respectively. In particular, the number of clusters is prespecified at 20 for both conditions; hence $|G_{\text{nonprogressor}}| = |G_{\text{progressor}}| = 20$. Further, we retained only the groups that were common in both conditions; that is, let $G = G_{\text{nonprogressor}} \cap G_{\text{progressor}}$, and the resulting common group G contains 7 subsets (groups), encompassing primarily FFAs and selected PCs, TAGs, lysoPCs and lysoPEs, and pPEs. As a confirmatory step, these groups were also detected in the mice serum data (57), where they exhibited strong correlations (coexpression signal). We allowed G^c to denote the indices of lipids that did not belong in any of these 7 groups; that is:

$$G^c = \{1, \dots, p\} \setminus \{\cup_g g \mid g \in G\} \text{ (Equation 2)}$$

To obtain a superset of the network skeleton (step 2), we estimated the skeleton (edges present) of the nonprogressor and progressor partial correlation networks based on a variation of the node-wise regression approach (68), while incorporating lipid group information extracted from step 1 and encouraging common sparsity structure (absence of edges in both networks) between the 2 groups (69). Toward this end, by considering the distribution of lipid j conditional on all the other lipids, their relative abundance levels satisfied the following relationship:

$$\mathbf{X}_j^k = \mathbf{X}_{-j}^k \beta_j^k + \mathbf{E}_j, \quad \beta_j^k \in \mathbb{R}^{p-1} \text{ (Equation 3)}$$

where each coordinate of the vector β_j^k encompassed the *scaled partial covariance* of lipid j with all other lipids j' ($j' \neq j$). Built upon the original node-wise regression formulation (68), to encourage similar sparsity structure incurred by lipids that were within the same group $g \in G$, we imposed a group penalty on their coefficients in the form of a vector ℓ_2 norm. On the other hand, for lipids that did not belong to any group, i.e., the elements in G^c , we modeled the coefficient as $\beta_{j,j'}^k$ as $\beta_{j,j'}^k = \gamma_{j,j'}^k \theta_{j,j'}^k, \theta_{j,j'}^k \geq 0$ (69) and penalized their individual absolute values to encourage the presence of common nodes (thus, absence of edges in the networks) across the nonprogressor and progressor groups. Formally, the optimization problem was formulated as:

$$\min_{\beta_j^k, \gamma_{j,j'}^k, \theta_{j,j'}^k, j \in \{1, \dots, p\}, k \in \{\text{non-progressor}, \text{progressor}\}} \sum_k \left\{ \sum_{j=1}^p \frac{1}{2n_k} \|\mathbf{X}_j^k - \mathbf{X}_{-j}^k \beta_j^k\|_2^2 + \sum_{g \in G} \lambda_g \|\beta_{j|g}^k\|_2 + \eta \sum_{j' \in G^c} |\gamma_{j,j'}^k| \right\} + \sum_{j \in G^c} \theta_{j,j'}^k \text{ . (*)}$$

(Equation 4)

The optimization problem in (*) estimated the nonprogressor and progressor partial correlation networks *jointly*, thus efficiently utilizing all 106 samples. Further, this objective function is separable in $j = 1, \dots, p$, and thus can be solved by splitting it into p parallel subproblems. For each subproblem indexed by j , it is equivalent to solving the following optimization problem obtained after some algebraic manipulations (69) (Lemma 2):

$$\min_{\beta_j^k, k \in \{\text{non-progressor}, \text{progressor}\}} \sum_k \left\{ \frac{1}{2n_k} \|\mathbf{X}_j^k - \mathbf{X}_{-j}^k \beta_j^k\|_2^2 + \sum_{g \in G} \lambda_g \|\beta_{j|g}^k\|_2 \right\} + \rho \sum_{j' \in G^c} \left(\sum_K |\beta_{j,j'}^k| \right)^{\frac{1}{2}}, \text{ where } \rho = 2\eta^{\frac{1}{2}} \text{ . (**)}$$

(Equation 5)

In the above formulation, λ_g, η, ρ were all prespecified tuning parameters, and they determined the ultimate sparsity level of the estimated networks, through the tuning of the corresponding norms. The solution to (*) contained the skeleton (presence/absence of edges) information of the partial correlation networks of interest; in particular, the nonzero elements in $\widehat{\beta}_j^k$ point to the presence of edges in the network. At this step,

we set the tuning parameters at a small relative value, to eliminate edges with a weak signal and to obtain a superset of the network skeleton. We denoted the obtained skeleton superset as \mathcal{S}^{super} , and at this step, the estimated skeletons for both conditions corresponded to networks with density level at around 10% (i.e., out of the possible $p^2 / 2$ edges, only 10% are present) for a technical justification of why only sparse partial correlation networks can be estimated from data when the sample size is smaller than the number of lipids (70).

To obtain the final stabilized network structures, we employed the technique of stability selection (71) coupled with the graphical lasso (72) network estimation procedure and used the skeleton information in \mathcal{S}^{super} as follows. We imposed a small lasso penalty for edges that are in \mathcal{S}^{super} and a larger one if in its complement. The stability selection step yielded the final network skeleton structure that was stable and robust to the choice of the tuning parameters, and the networks had a respective density level of 3% (non-progressor) and 2.7% (progressor). Finally, based on the skeleton, we estimated the magnitude of the edges (correlations), which after proper normalization, corresponded to the strength of the partial correlation among lipids in the 2 groups.

Study approval. The study was approved by the Institutional Review Board (IRB 0000006) at the National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland, USA. All participants gave written informed consent before their participation in the study.

Author contributions

FA designed the lipidomic study, prepared serum samples for mass spectrometry, analyzed and interpreted the data, and wrote the first draft. VN performed the transcriptomic-lipidomic analysis and contributed to manuscript drafting. JL and GM performed the differential network analysis and contributed to manuscript drafting. TMR and JB helped with serum sample preparation for lipidomic analysis and MS runs. TS retrieved MS data. HCL and RGN contributed to clinical study design, data collection, manuscript drafting, and critical evaluation of the paper. SP contributed to the lipidomic study design, MS, data interpretation, and manuscript drafting. FA, VN, JL, TMR, TS, JB, KS, PEF, TWG, HCL, RGN, FCB, ELF, GM, MK, and SP critically evaluated the paper and have approved the final version.

Acknowledgments

Authors would like to thank Lois Jones, RN; Enrique Diaz, RN; Bernadine Waseta; and Camille Waseta for seeing the patients, performing the research studies, and collecting and processing the samples.

This work was supported by the NIH grants R24DK082841, K08DK106523, R03DK121941, P30DK089503, P30DK081943, and P30DK020572; the Program for Neurology Research and Discovery at Michigan Medicine (<https://www.pnrdfeldman.org>); and the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases.

Address correspondence to: Subramaniam Pennathur, Division of Nephrology, 5309 Brehm Center, 1000 Wall Street, University of Michigan, Ann Arbor, Michigan 48105, USA. Phone: 734.936.5645; Email: spennath@umich.edu.

- Morgenstern H, Robinson B, Saran R. Chapter 1: Incidence, prevalence, patient characteristics, and treatment modalities. In: United States Renal Data System. *2017 Annual Data Report*. Bethesda, Maryland, USA: NIH and National Institute of Diabetes and Digestive and Kidney Diseases; 2017: 247–276. End Stage Renal Disease in the United States; vol 2.
- United States Renal Data System. *2018 Annual Data Report*. Bethesda, Maryland, USA: NIH and National Institute of Diabetes and Digestive and Kidney Diseases; 2018.
- Institute for Alternative Futures. Diabetes 2030: United States diabetes data & forecasts. <http://www.altfutures.org/pubs/diabetes2030/UNITEDSTATESDataSheet.pdf>. Updated 2015. Accessed October 10, 2019.
- Rahman M, et al. Relation of serum lipids and lipoproteins with progression of CKD: the CRIC study. *Clin J Am Soc Nephrol*. 2014;9(7):1190–1198.
- Reis A, et al. Top-down lipidomics of low density lipoprotein reveal altered lipid profiles in advanced chronic kidney disease. *J Lipid Res*. 2015;56(2):413–422.
- Reiss AB, Voloshyna I, De Leon J, Miyawaki N, Mattana J. Cholesterol metabolism in CKD. *Am J Kidney Dis*. 2015;66(6):1071–1082.
- Tsuruya K, et al. Impact of the triglycerides to high-density lipoprotein cholesterol ratio on the incidence and progression of CKD: a longitudinal study in a large Japanese population. *Am J Kidney Dis*. 2015;66(6):972–983.
- Wahl P, Ducasa GM, Fornoni A. Systemic and renal lipids in kidney disease development and progression. *Am J Physiol Renal Physiol*. 2016;310(6):F433–F445.

9. Quehenberger O, et al. Lipidomics reveals a remarkable diversity of lipids in human plasma. *J Lipid Res.* 2010;51(11):3299–3305.
10. Quehenberger O, Dennis EA. The human plasma lipidome. *N Engl J Med.* 2011;365(19):1812–1823.
11. Graessler J, Schwudke D, Schwarz PE, Herzog R, Shevchenko A, Bornstein SR. Top-down lipidomics reveals ether lipid deficiency in blood plasma of hypertensive patients. *PLoS ONE.* 2009;4(7):e6261.
12. Kontush A, Lhomme M, Chapman MJ. Unraveling the complexities of the HDL lipidome. *J Lipid Res.* 2013;54(11):2950–2963.
13. Pietiläinen KH, 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(2):e218.
14. Stegemann C, et al. Comparative lipidomics profiling of human atherosclerotic plaques. *Circ Cardiovasc Genet.* 2011;4(3):232–242.
15. Stegemann C, et al. Lipidomics profiling and risk of cardiovascular disease in the prospective population-based Bruneck study. *Circulation.* 2014;129(18):1821–1831.
16. Rhee EP, et al. Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *J Clin Invest.* 2011;121(4):1402–1411.
17. Rhee EP, et al. A combined epidemiologic and metabolomic approach improves CKD prediction. *J Am Soc Nephrol.* 2013;24(8):1330–1338.
18. Afshinnia F, et al. Impaired β -oxidation and altered complex lipid fatty acid partitioning with advancing CKD. *J Am Soc Nephrol.* 2018;29(1):295–306.
19. Afshinnia F, et al. Lipidomic Signature of progression of chronic kidney disease in the chronic renal insufficiency cohort. *Kidney Int Rep.* 2016;1(4):256–268.
20. Niewczas MA, et al. Circulating TNF receptors 1 and 2 predict ESRD in type 2 diabetes. *J Am Soc Nephrol.* 2012;23(3):507–515.
21. Niewczas MA, et al. Uremic solutes and risk of end-stage renal disease in type 2 diabetes: metabolomic study. *Kidney Int.* 2014;85(5):1214–1224.
22. Looker HC, et al. Biomarkers of rapid chronic kidney disease progression in type 2 diabetes. *Kidney Int.* 2015;88(4):888–896.
23. Saulnier PJ, et al. Association of circulating biomarkers (adrenomedullin, TNFR1, and NT-proBNP) with renal function decline in patients with type 2 diabetes: a French prospective cohort. *Diabetes Care.* 2017;40(3):367–374.
24. Choudhury D, Tuncel M, Levi M. Diabetic nephropathy -- a multifaceted target of new therapies. *Discov Med.* 2010;10(54):406–415.
25. Nowak N, et al. Markers of early progressive renal decline in type 2 diabetes suggest different implications for etiological studies and prognostic tests development. *Kidney Int.* 2018;93(5):1198–1206.
26. Fufaa GD, et al. Structural predictors of loss of renal function in American Indians with type 2 diabetes. *Clin J Am Soc Nephrol.* 2016;11(2):254–261.
27. Nelson RG, et al. Development and progression of renal disease in Pima Indians with non-insulin-dependent diabetes mellitus. Diabetic Renal Disease Study Group. *N Engl J Med.* 1996;335(22):1636–1642.
28. Pavkov ME, Knowler WC, Hanson RL, Nelson RG. Diabetic nephropathy in American Indians, with a special emphasis on the Pima Indians. *Curr Diab Rep.* 2008;8(6):486–493.
29. Saulnier PJ, et al. Advanced glycation end products predict loss of renal function and correlate with lesions of diabetic kidney disease in American Indians with type 2 diabetes. *Diabetes.* 2016;65(12):3744–3753.
30. Tanamas SK, et al. Long-term effect of losartan on kidney disease in American Indians with type 2 diabetes: a follow-up analysis of a randomized clinical trial. *Diabetes Care.* 2016;39(11):2004–2010.
31. Weil EJ, et al. Effect of losartan on prevention and progression of early diabetic nephropathy in American Indians with type 2 diabetes. *Diabetes.* 2013;62(9):3224–3231.
32. Lemley KV, et al. Glomerular permselectivity at the onset of nephropathy in type 2 diabetes mellitus. *J Am Soc Nephrol.* 2000;11(11):2095–2105.
33. Lemley KV, et al. modeling GFR trajectories in diabetic nephropathy. *Am J Physiol Renal Physiol.* 2005;289(4):F863–F870.
34. Levey AS, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. *Am J Kidney Dis.* 2014;64(6):821–835.
35. Luo S, et al. Serum metabolomic alterations associated with proteinuria in CKD. *Clin J Am Soc Nephrol.* 2019;14(3):342–353.
36. Grove KJ, et al. Diabetic nephropathy induces alterations in the glomerular and tubule lipid profiles. *J Lipid Res.* 2014;55(7):1375–1385.
37. Calzada E, Onguka O, Claypool SM. Phosphatidylethanolamine metabolism in health and disease. *Int Rev Cell Mol Biol.* 2016;321:29–88.
38. Emoto K, Toyama-Sorimachi N, Karasuyama H, Inoue K, Umeda M. Exposure of phosphatidylethanolamine on the surface of apoptotic cells. *Exp Cell Res.* 1997;232(2):430–434.
39. Rockenfeller P, et al. Phosphatidylethanolamine positively regulates autophagy and longevity. *Cell Death Differ.* 2015;22(3):499–508.
40. Karaskov E, Scott C, Zhang L, Teodoro T, Ravazzola M, Volchuk A. Chronic palmitate but not oleate exposure induces endoplasmic reticulum stress, which may contribute to INS-1 pancreatic beta-cell apoptosis. *Endocrinology.* 2006;147(7):3398–3407.
41. Sieber J, Jehle AW. Free fatty acids and their metabolism affect function and survival of podocytes. *Front Endocrinol (Lausanne).* 2014;5:186.
42. Sieber J, et al. Regulation of podocyte survival and endoplasmic reticulum stress by fatty acids. *Am J Physiol Renal Physiol.* 2010;299(4):F821–F829.
43. Soumura M, et al. Oleate and eicosapentaenoic acid attenuate palmitate-induced inflammation and apoptosis in renal proximal tubular cell. *Biochem Biophys Res Commun.* 2010;402(2):265–271.
44. Martínez-García C, et al. Renal lipotoxicity-associated inflammation and insulin resistance affects actin cytoskeleton organization in podocytes. *PLoS ONE.* 2015;10(11):e0142291.
45. Lennon R, et al. Saturated fatty acids induce insulin resistance in human podocytes: implications for diabetic nephropathy. *Nephrol Dial Transplant.* 2009;24(11):3288–3296.
46. Jiang XS, Chen XM, Wan JM, Gui HB, Ruan XZ, Du XG. Autophagy protects against palmitic acid-induced apoptosis in podocytes in vitro. *Sci Rep.* 2017;7:42764.
47. Lee E, Choi J, Lee HS. Palmitate induces mitochondrial superoxide generation and activates AMPK in podocytes. *J Cell Physiol.* 2017;232(12):3209–3217.

48. Xu S, et al. Palmitate induces ER calcium depletion and apoptosis in mouse podocytes subsequent to mitochondrial oxidative stress. *Cell Death Dis.* 2015;6:e1976.
49. Yasuda M, et al. Fatty acids are novel nutrient factors to regulate mTORC1 lysosomal localization and apoptosis in podocytes. *Biochim Biophys Acta.* 2014;1842(7):1097–1108.
50. Rhee EP, et al. Metabolite profiling identifies markers of uremia. *J Am Soc Nephrol.* 2010;21(6):1041–1051.
51. Gopinathrao G, D'Eustachio P. Integration of energy metabolism. Reactome. <http://www.reactome.org/content/detail/R-HSA-163685>. Published May 11, 2005. Accessed October 11, 2019.
52. Otero YF, Stafford JM, McGuinness OP. Pathway-selective insulin resistance and metabolic disease: the importance of nutrient flux. *J Biol Chem.* 2014;289(30):20462–20469.
53. Song Z, Xiaoli AM, Yang F. Regulation and metabolic significance of *de novo* lipogenesis in adipose tissues. *Nutrients.* 2018;10(10):E1383.
54. Carlson CA, Kim KH. Regulation of hepatic acetyl coenzyme A carboxylase by phosphorylation and dephosphorylation. *Arch Biochem Biophys.* 1974;164(2):478–489.
55. Gillies CE, et al. An eQTL landscape of kidney tissue in human nephrotic syndrome. *Am J Hum Genet.* 2018;103(2):232–244.
56. Sieber J, et al. Susceptibility of podocytes to palmitic acid is regulated by stearoyl-CoA desaturases 1 and 2. *Am J Pathol.* 2013;183(3):735–744.
57. Sas KM, et al. Shared and distinct lipid-lipid interactions in plasma and affected tissues in a diabetic mouse model. *J Lipid Res.* 2018;59(2):173–183.
58. Li Y, et al. Fatty acid receptor modulator PBI-4050 inhibits kidney fibrosis and improves glycemic control. *JCI Insight.* 2018;3(10):120365.
59. Afshinnia F, et al. Lipidomics and biomarker discovery in kidney disease. *Semin Nephrol.* 2018;38(2):127–141.
60. Sas KM, et al. Tissue-specific metabolic reprogramming drives nutrient flux in diabetic complications. *JCI Insight.* 2016;1(15):e86976.
61. Cohen CD, Frach K, Schlöndorff D, Kretzler M. Quantitative gene expression analysis in renal biopsies: a novel protocol for a high-throughput multicenter application. *Kidney Int.* 2002;61(1):133–140.
62. Schmid H, et al. Modular activation of nuclear factor-kappaB transcriptional programs in human diabetic nephropathy. *Diabetes.* 2006;55(11):2993–3003.
63. Overmyer KA, et al. Maximal oxidative capacity during exercise is associated with skeletal muscle fuel selection and dynamic changes in mitochondrial protein acetylation. *Cell Metab.* 2015;21(3):468–478.
64. Muller CJ, MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. *Int J Epidemiol.* 2014;43(3):962–970.
65. Karnovsky A, et al. Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics.* 2012;28(3):373–380.
66. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* 1995;57(1):289–300.
67. Ng AY, Jordan MI, Weiss Y. On spectral clustering: analysis and an algorithm. *Adv Neural Inf Process Syst.* 2002;14:849–56.
68. Shojaie A, Michailidis G. Penalized likelihood methods for estimation of sparse high-dimensional directed acyclic graphs. *Biometrika.* 2010;97(3):519–538.
69. Guo J, Levina E, Michailidis G, Zhu J. Joint estimation of multiple graphical models. *Biometrika.* 2011;98(1):1–15.
70. Basu S, Duren W, Evans CR, Burant CF, Michailidis G, Karnovsky A. Sparse network modeling and Metscape-based visualization methods for the analysis of large-scale metabolomics data. *Bioinformatics.* 2017;33(10):1545–1553.
71. Meinshausen N, Bühlmann P. Stability selection. *J R Stat Soc Series B Stat Methodol.* 2010;72(4):417–473.
72. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics.* 2008;9(3):432–441.