Clusters of fatty acids in the serum triacylglyceride fraction associate with the disorders of type 2 diabetes[S]

Luke W. Johnston,* Zhen Liu,* Ravi Retnakaran,[†,§] Bernard Zinman,[†,§] Adria Giacca,** Stewart B. Harris,[††] Richard P. Bazinet,* and Anthony J. Hanley[1,*,†,§§]

Department of Nutritional Sciences,* Leadership Sinai Centre for Diabetes, Division of Endocrinology,[†] Department of Physiology,** and Dalla Lana School of Public Health,[§§] University of Toronto, Toronto, Ontario, Canada; Lunenfeld Tanenbaum Research Institute,[§] Mount Sinai Hospital, Toronto, Ontario, Canada; and Centre for Studies in Family Medicine,[††] University of Western Ontario, London, Ontario, Canada

ORCID ID: 0000-0003-4169-2616 (L.W.J.)

Abstract   Our aim was to examine longitudinal associations of triacylglyceride fatty acid (TGFA) composition with insulin sensitivity (IS) and β-cell function. Adults at risk for T2D (n = 477) had glucose and insulin measured from a glucose challenge at three time points over 6 years. The outcome variables Matsuda insulin sensitivity index, homeostatic model of assessment 2–percent sensitivity (HOMA2-%S), Insulinogenic Index over HOMA-IR (IGI/IR), and Insulin Secretion-Sensitivity Index-2 were computed from the glucose challenge. Gas chromatography quantified TGFA composition from the baseline. We used adjusted generalized estimating equation (GEE) models and partial least squares (PLS) regression for the analysis. In adjusted GEE models, four TGFAs (14:0, 16:0, 14:1n-7, and 16:1n-7 as mol%) had strong negative associations with IS, whereas others (e.g., 18:1n-7, 18:1n-9, 20:2n-6, and 20:5n-3) had strong positive associations. Few associations were seen for β-cell function, except for 16:0, 18:1n-7, and 20:2n-6. PLS analysis indicated four TGFAs (14:0, 16:0, 14:1n-7, and 16:1n-7) that clustered together and strongly related with lower IS. These four TGFAs also correlated highly (*r* > 0.4) with clinically measured triacylglyceride.   We found that higher proportions of a cluster of four TGFAs strongly related with lower IS as well as hypertriglyceridemia, suggesting that only a few FAs within the TGFA composition may primarily explain lipids' role in glucose dysregulation.—Johnston, L. W., Z. Liu, R. Retnakaran, B. Zinman, A. Giacca, S. B. Harris, R. P. Bazinet, and A. J. Hanley. **Clusters of fatty acids in the serum triacylglyceride fraction associate with the disorders of type 2 diabetes.** *J. Lipid Res.* 2018. 59: 1751–1762.

Hypertriglyceridemia is a well-described metabolic disorder resulting in negative health outcomes (1). It is a risk factor for CVD (2, 3) and is associated with other metabolic disorders such as nonalcoholic fatty liver disease (4), metabolic syndrome (5), and abdominal obesity (6). Circulating triacylglyceride (TG) concentration is commonly measured during routine clinical assessment using enzymatic methods. However, clinically measured TG is limited, as it represents the full FA spectrum within the TG fraction as a summary measure. There is increasing appreciation for the importance of specific FA composition profiles in different plasma fractions on various health outcomes (7–10); however, there are relatively few studies that have explored the impact of the FA composition in the TG fraction (11, 12).

The interaction between TG and insulin sensitivity (IS) is complex and involves components of a feedback system (3). Greater resistance to insulin in both the liver and muscle may result in greater production of TG and secretion of lipoproteins that transport TG (13). Likewise, greater TG

may contribute to metabolic dysfunction and lipotoxicity in various tissues, affecting IS, and thus continue the cycle (3). Given the complexity and temporal nature of the relationship, long-term studies with multiple data collection time points are paramount to better understanding the underlying biology and subsequent risk.

Although several studies have documented prospective associations of hypertriglyceridemia with incident T2D (2, 14, 15), only a limited number of longitudinal studies (11, 12) have examined the relationship between TG and its composition with the pathophysiological factors underlying T2D, particularly β-cell function. Our objective was to examine the longitudinal role of the specific composition of the serum TG fraction on the oral glucose tolerance test (OGTT)-derived measures of IS and β-cell function compared with clinically measured TG in a Canadian population at risk for T2D.

MATERIALS AND METHODS

Recruitment for the baseline visit of the Prospective Metabolism and Islet Cell Evaluation (PROMISE) cohort took place between 2004 and 2006 in London and Toronto, Canada. Individuals were selected to participate if they met the eligibility criteria of having one or more risk factors for T2D, including obesity, hypertension, family history of diabetes, and/or a history of gestational diabetes or birth of a macrosomic infant. A total of 736 individuals attended the baseline visit. Subsequent examinations occurred every 3 years, with data from three examination visits available for the present analysis (2004–2006, 2007–2009, and 2010–2013). The present study used data on participants who did not have T2D at baseline, who returned for one or more of the follow-up examinations, and who had samples available for FA measurements (n = 477; see the CONSORT diagram in supplemental Figure S1). Metabolic characterization, anthropometric measurements, and questionnaires on lifestyle and sociodemographics were administered at each examination visit. Research ethics approval was obtained from Mount Sinai Hospital and the University of Western Ontario, and all participants provided written informed consent. Data collection methods were standardized across the two centers, and research nurses were centrally trained.

Metabolic characterization

After 8–12 h of overnight fasting, participants completed a 75 g OGTT at each examination visit, with blood samples taken at fasting, 30 min, and 2 h postglucose load. Samples were subsequently processed and frozen at −70°C. Alanine aminotransferase (ALT) was measured using standard laboratory procedures. Cholesterol, HDL, and clinically measured TG were measured using Roche Modular's enzymatic colorimetric tests (Mississauga, ON). Both insulin and glucose were measured from OGTT blood samples at fasting, 30 min, and 2 h time points. Specific insulin was measured with the Elecsys 1010 (Roche Diagnostics, Basel, Switzerland) immunoassay analyzer and electrochemiluminescence immunoassay, which shows 0.05% cross-reactivity to intact human proinsulin and the Des 31,32 circulating split form (Linco Res., Inc) and has a coefficient of variation (CV) of 9.3%. Glucose was determined using an enzymatic hexokinase method (Roche Modular, Roche Diagnostics) with a detection range of 0.11–41.6 mmol/l, an interassay CV of <1.1%, and an intraassay CV of < 1.9%. All assays were performed at the Banting and Best Diabetes Centre Core Lab at Mt. Sinai Hospital.

Triacylglyceride FA (TGFA) composition was quantified using stored fasting serum samples from the baseline visit, which had been frozen at −70°C for 4–6 years and had not been exposed to any freeze–thaw cycles. Serum FAs have been documented to be stable at these temperatures for up to 10 years (16). A known amount of triheptadecanoin (17:0; Nu-Chek Prep, Inc., Elysian, MN) was added as an internal standard prior to extracting total lipids according to the method of Folch et al. (17). Each serum lipid fraction (NEFAs, cholesteryl ester, phospholipid, and TG) was isolated using TLC. FA methyl esters were separated and quantified using a Varian-430 gas chromatograph (Varian, Lake Forest, CA) equipped with a Varian Factor Four capillary column and a flame ionization detector. FA concentrations (nmol/ml) were calculated by proportional comparison of gas chromatography peak areas to that of the internal standards (18). There were 22 FAs measured in the TGFA fraction. Findings for other lipid fractions in this cohort are reported separately (see ref. 9 for the phospholipid and cholesteryl ester fraction and ref. 10 for the NEFA fraction analysis).

Anthropometrics and sociodemographics

Height, weight, and waist circumference (WC) were measured at all clinic examinations using standard procedures. WC was measured at the natural waist, defined as the narrowest part of the torso between the umbilicus and the xiphoid process. BMI was calculated by dividing weight (kilograms) by height (meters) squared. Questionnaires administered at each examination determined sociodemographics. A version of the Modifiable Activity Questionnaire (MAQ) (19) determined estimated physical activity. The MAQ collects information on leisure and occupational activity, including intensity, frequency, and duration, over the past year. Each reported activity from the MAQ was weighted by its metabolic intensity, allowing for the estimation of metabolic equivalents of tasks (METs) hours per week (19).

Variable calculation and statistical analysis

IS and β-cell function indices were computed using the OGTT glucose and insulin data. IS was assessed using the IS Index (ISI) (20) and homeostatic model of assessment 2–percent sensitivity (HOMA2-%S) (21) using the HOMA2 Calculator. HOMA largely reflects hepatic IS, whereas ISI reflects whole-body IS (22). Beta-cell function was assessed using the Insulinogenic Index (23) over HOMA-IR (24) (IGI/IR) and the Insulin Secretion-Sensitivity Index-2 (ISSI-2) (25). IGI/IR is a measure of the early phase of insulin secretion, whereas ISSI-2 is analogous to the disposition index (but is calculated using OGTT values). Each index has been validated against gold-standard measures (20, 24–26).

The primary outcome variables for this analysis were HOMA2-%S, ISI, IGI/IR, and ISSI-2, which were log-transformed for the statistical modeling. The primary predictor variables for this analysis were 22 individual TGFAs included as either mol% of the total fraction or as a concentration (nmol/ml). Clinically measured TG was also included as a primary predictor to allow us to test the hypothesis that specific TGFAs better predicted outcomes compared with clinical TG. Pearson correlation coefficients were computed to assess the relationships of individual TGFAs with other continuous variables. Within-TGFA composition correlations were also computed and subsequently analyzed using hierarchical clustering.

Generalized estimating equation (GEE) models (27) were used in the primary analysis to determine the longitudinal associations between the outcome variables and the predictor variables. The predictor variables were scaled (mean-centered and standardized). Given the longitudinal design, an autoregressive of order 1 working correlation matrix was specified in the GEE model.

Covariates to adjust for were selected based on the previous literature, from directed acyclic graph (DAG) (28) recommendations and from quasilikelihood information criteria. DAGs are used to identify the minimum adjustment necessary for a model by using the causal pathways to algorithmically identify potential confounding and colliding variables (see ref. 28 for more details about using DAGs). The DAG structures to understand potential confounding, shown in supplemental Figs. S2 and S3, were processed by the DAGitty software (29, 30) to generate the recommended adjustments. These DAG structures were developed based on hypothesized causal pathways between each variable, which were then input into the DAGitty software. The output from DAGitty was used, in conjunction with the other methods, to help inform the final model.

The final GEE model (M6; seen in supplemental Table S1) was adjusted for years since baseline, WC, baseline age, ethnicity, sex, ALT, MET, and total NEFA. The variables TGFA, total NEFA, sex, ethnicity, and baseline age were classified as time-independent (held constant), as they were measured only at the baseline visit or do not change throughout the study, whereas the outcome variables and remaining covariates were set as time-dependent. After transformations, the GEE estimates were interpreted as an expected percent difference in the outcome variable for every SD increase in the predictor variable, given the covariates are held constant (including time). We also tested for an interaction with sex, ethnicity, or time by the predictor term for each outcome variable.

Although GEE accounts for the longitudinal design of the data, this approach is limited in that it cannot analyze the inherent multivariate nature of the composition of the TGFA fraction. Therefore, to confirm the GEE results in a multivariate environment (i.e., all TGFAs analyzed collectively), partial least squares (PLS) regression was used to identify the patterns of TGFA composition against IS and β-cell function as outcome variables. Briefly, PLS is a technique that extracts latent structures (clusters) underlying a set of predictor variables conditional on a response variable(s) (i.e., the outcome variables). How accurately the clusters within the TGFA composition predict metabolic function is determined by using cross-validation on the PLS models.

A more detailed explanation of these statistical techniques and on the analysis process can be found in the supplemental methods for our paper in the NEFA fraction (10). All analyses were performed using R (Version 3.4.4) (31), along with the R packages geepack (Version 1.2.1) for GEE (32) and pls (Version 2.6.0) for PLS. The R code and extra analyses for this manuscript are available at https://doi.org/10.6084/m9.figshare.5143438. Results were considered statistically significant at $P < 0.05$, after adjusting for multiple testing using the Benjamini–Hochberg (BH) false discovery rate (33). STROBE was used as a guideline for reporting (34).

RESULTS

Basic characteristics of the PROMISE cohort

Table 1 shows basic characteristics of the PROMISE cohort. The mean follow-up time was 5.6 (1.0) years, where 88.5% of participants attended all three visits. There were 349 (73.2%) females and 336 (70.4%) who reported European ancestry, with a mean age in years of 50.0 (9.8) and a mean BMI of 31.1 (6.5) kg/m². As expected from the study's eligibility criteria, the majority of participants, n = 305 (65.3%), had a family history of diabetes. Between the baseline visit and the 6 year visit in this sample, IS and β-cell function measures had a significant median decline of between 14% and 21% ($P < 0.001$ from GEE; n = 357–470).

Figure 1 shows the composition of each FA in the TG fraction (see supplemental Table S2 for the raw values). Three TGFAs contributed 82.4% to the total TG concentration: 18:1 n-9 (37.8%); 16:0 (26.6%); and, 18:2 n-6 (18.0%). **Figure 2** shows a heatmap of the correlation of individual TGFAs as concentrations with the outcome variables and several basic characteristics. As expected, nearly all TGFAs had very strong positive correlations ($r = 0.34$–0.92) with clinically measured TG and moderate positive correlations with WC ($r = 0.31$–0.36). There were also moderate negative correlations with HDL ($r = -0.53$ to -0.31).

TABLE 1. Basic characteristics of PROMISE participants at each of the three clinic visits

Measure	Baseline	3 Year	6 Year
HOMA2-%S	88.8 (54.2–136.7)	76.8 (49.1–121.8)	73.7 (49.5–110.1)
ISI	13.6 (8.7–21.8)	11.6 (6.9–19.1)	11.7 (7.5–17.6)
IGI/IR	7.1 (4.2–10.6)	5.6 (3.6–9.8)	5.6 (3.4–9.2)
ISSI-2	727.5 (570.0–922.5)	611.2 (493.0–836.4)	624.8 (470.0–813.8)
ALT (U/l)	29.6 (16.0)	28.4 (19.6)	25.8 (16.9)
TG (mmol/l)	1.5 (0.8)	1.4 (0.8)	1.4 (0.7)
Chol (mmol/l)	5.2 (0.9)	5.1 (1.0)	5.1 (0.9)
HDL (mmol/l)	1.4 (0.4)	1.3 (0.4)	1.4 (0.4)
TGFA (nmol/ml)	3,137.5 (1,686.6)		
NEFA (nmol/ml)	383.1 (116.3)		
MET	46.1 (61.4)	48.2 (60.4)	43.7 (56.7)
Age (years)	50.0 (9.8)	53.2 (9.8)	56.2 (9.6)
BMI (kg/m²)	31.1 (6.5)	31.4 (6.5)	31.1 (6.6)
WC (cm)	98.5 (15.5)	99.2 (15.7)	100.5 (15.8)
Ethnicity			
European	336 (70%)		
Latino/a	59 (12%)		
Other	50 (10%)		
South Asian	32 (7%)		
Sex			
Female	349 (73%)		
Male	128 (27%)		

Values are median (interquartile range), mean (SD), or n (percent). Chol, cholesterol.
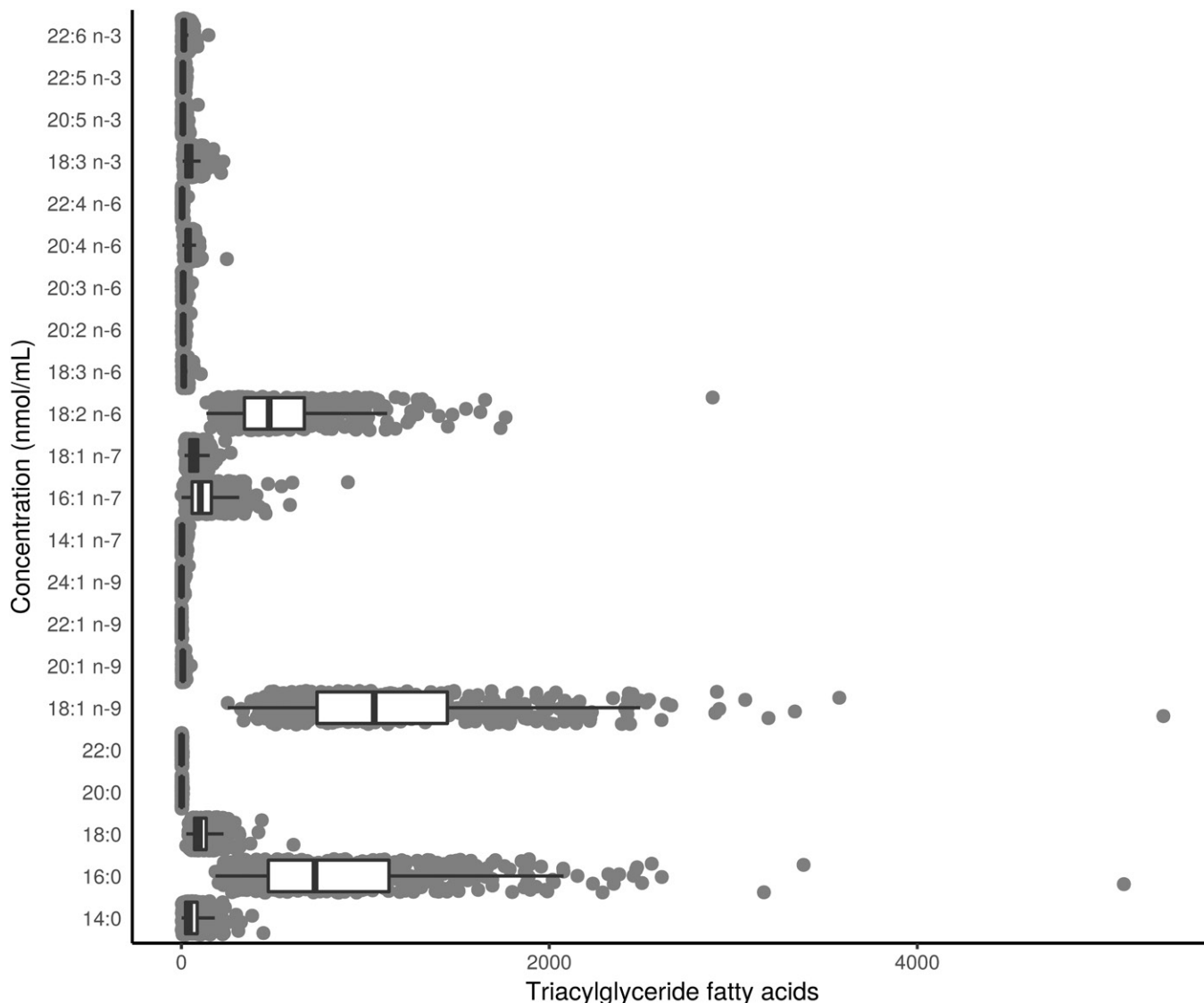
Fig. 1. Distribution of the composition of TGFAs in the baseline visit of PROMISE participants (2004–2006). Boxplots represent the median and interquartile range of the FA values.

For the outcome variables, the correlations for the IS measures were generally higher (HOMA2-%S: $r = -0.47$ to -0.32, ISI: $r = -0.47$ to -0.31) than for the β-cell function measures (all $r < 0.30$). For correlations of individual TGFAs using mol% with the basic participant characteristics, as shown in **Fig. 3**, differences in correlations between FAs were most evident for 14:0, 14:1n-7, 16:0, and 16:1n-7 that had a moderate positive correlation with clinical TG ($r = 0.42$–0.52), whereas all other FAs had a negative association ($r = -0.5$ to -0.34). In particular, those FAs with the negative associations with clinical TG were all the very-long-chain PUFAs (e.g., 20:4n-6 and 20:5n-3). As seen in **Fig. 4**, four FAs (14:0, 16:0, 14:1n-7, and 16:1n-7) clustered together, each highly positively correlated with each other and negatively correlated with all other FAs.

GEE models

Results from the unadjusted GEE model are shown in **Fig. 5** and for the adjusted GEE model in **Fig. 6**. The majority of

associations with β-cell function measures were attenuated after full model adjustment, whereas nearly all associations with IS remained significant for both mol% and nmol/ml results. Subsequent analysis revealed that the attenuation with β-cell function was due primarily to adjustment for WC.

In analyses using concentration values, nearly all TGFAs had a strong negative association on HOMA2-%S and ISI (estimates of percent difference ranging from -13.5 to -4.2 and -14.8 to -4.5, respectively), and a few had strong negative associations with IGI/IR and ISSI-2 (estimates ranging from -8.4 to -6.0 and -4.6 to -3.7, respectively). In analyses using TGFA mol% values, four TGFAs (14:0, 16:0, 14:1n-7, and 16:1n-7) had negative associations with HOMA2-%S and ISI (estimates between -11.9 to -6.1 and -12.8 to -6.4, respectively, lower IS for every SD increase in the TGFA), whereas several more TGFAs had positive associations with HOMA2-%S and ISI (20:0, 18:1n-9, 20:1n-9, 22:1n-9, 18:2n-6, 20:2n-6, 20:4n-6, and 22:5n-3) estimating between 4.3–14.4% and 5.9–15.9%, respectively, higher IS
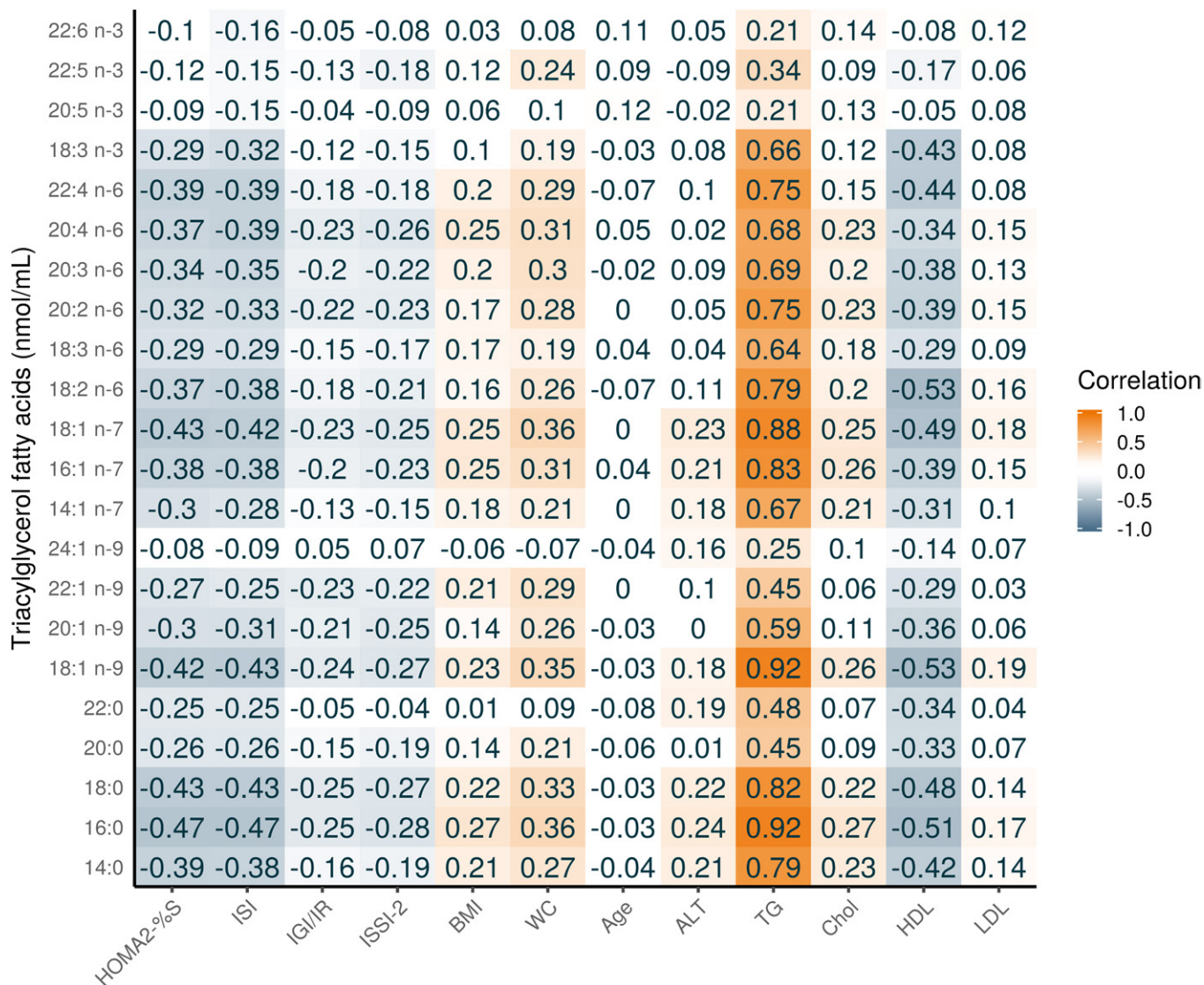
Fig. 2. Pearson correlation heatmap of TGFAs (nmol/ml) with continuous basic and metabolic characteristics of PROMISE participants from the baseline visit (2004–2006). Darker orange represents a positive correlation; darker blue represents a negative correlation. Chol., cholesterol.

for every SD increase in the TGFA. One TGFA, 20:2n-6, had a very strong positive association with the IS measures, with a 14.4–9.3% higher IS for every SD increase. Both clinically measured TG and total TGFA concentration had very strong negative associations with all outcome variables.

Although there were a few significant interactions by time in unadjusted models, after inclusion of covariates in the model, these interactions were attenuated (data not shown). There were no significant interactions by sex or ethnicity for any of the TGFAs (data not shown). Results of the sensitivity analyses identifying WC as the covariate that attenuated the β-cell function associations from the unadjusted model are shown in supplemental Fig. S4. A tabular presentation of the GEE results is shown in supplemental Table S3 for unadjusted models and supplemental Table S4 for adjusted models.

Clustering of TGFAs by metabolic measures

The PLS analysis corroborated the findings from the GEE models. The PLS results conditioned on IS as the outcome showed a clustering of the FAs 14:0, 14:1n-7, 16:0, and 16:1n-7 as mol% (**Fig. 7**). These TGFAs loaded strongly and negatively on HOMA2-%S and ISI in the first component, suggesting that this cluster of TGFAs tracks together with lower IS. The TGFAs 20:2n-6, 20:5n-3, 22:5n-3, and 22:6n-3 loaded positively on both IS measures. No other TGFAs loaded strongly. In the second component, 18:1n-9 and 18:1n-7 loaded positively, but not strongly, whereas 20:5n-3 and 22:6n-3 loaded strongly and negatively with both HOMA2-%S and ISI; however, this component only explained <10% of the variance. The PLS model for IS had good predictive ability, with a high correlation between the predicted outcome values against the observed values (HOMA2-%S: $r = 0.46$, $P < 0.001$; ISI: $r = 0.39$, $P < 0.001$).

The β-cell function PLS results showed a similar clustering of FAs; however, there was a lower correlation (although significant at $P < 0.001$) between the predicted values and the observed values ($r = 0.25$–0.24), suggesting that TGFA composition poorly predicts β-cell function. Given the low predictability, only the IS measures are
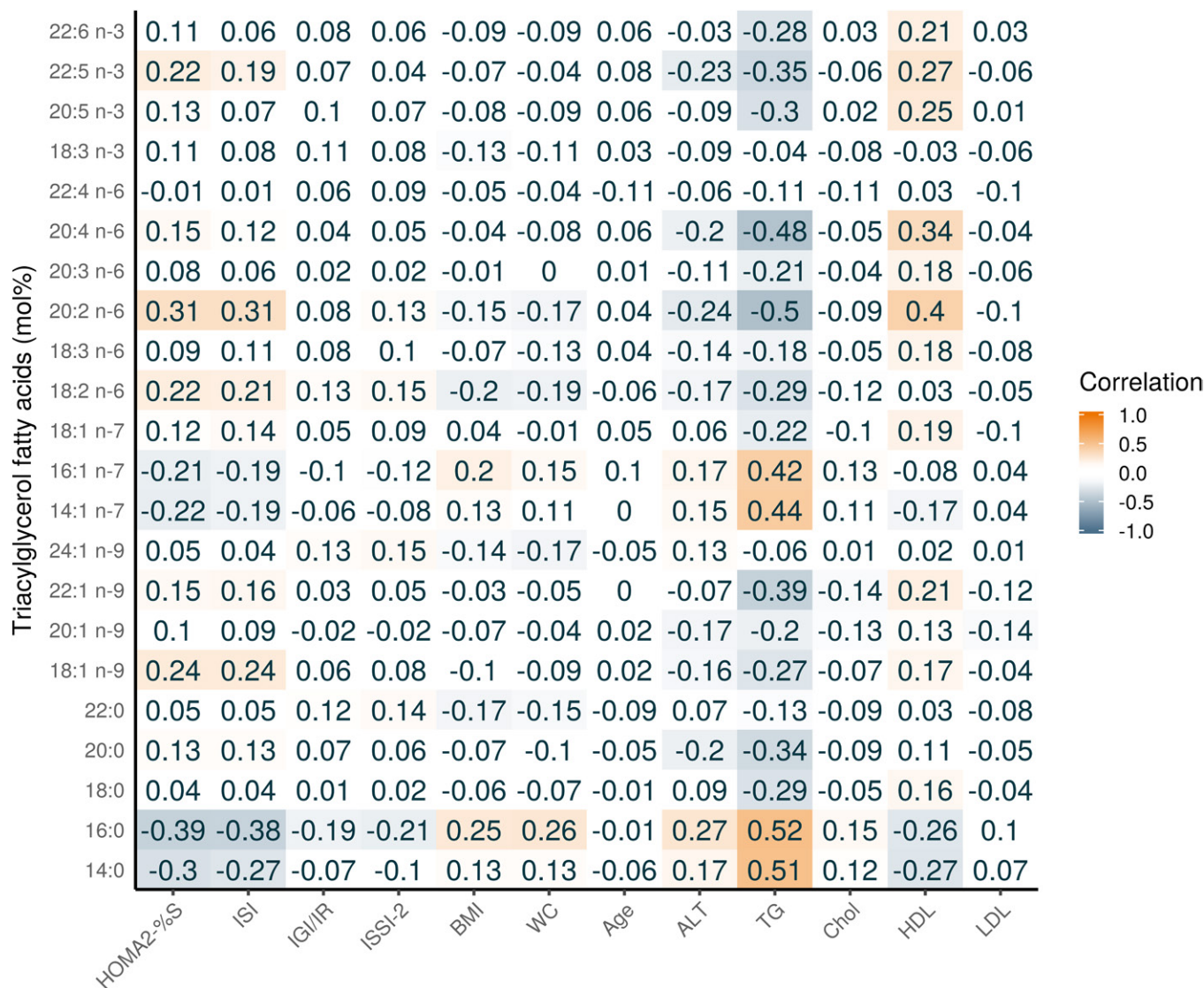
Triacylglycerol fatty acids (mol%)	HOMA2-%S	ISI	IGI/IR	ISSI-2	BMI	WC	Age	ALT	TG	Chol	HDL	LDL
22:6 n-3	0.11	0.06	0.08	0.06	-0.09	-0.09	0.06	-0.03	-0.28	0.03	0.21	0.03
22:5 n-3	0.22	0.19	0.07	0.04	-0.07	-0.04	0.08	-0.23	-0.35	-0.06	0.27	-0.06
20:5 n-3	0.13	0.07	0.1	0.07	-0.08	-0.09	0.06	-0.09	-0.3	0.02	0.25	0.01
18:3 n-3	0.11	0.08	0.11	0.08	-0.13	-0.11	0.03	-0.09	-0.04	-0.08	-0.03	-0.06
22:4 n-6	-0.01	0.01	0.06	0.09	-0.05	-0.04	-0.11	-0.06	-0.11	-0.11	0.03	-0.1
20:4 n-6	0.15	0.12	0.04	0.05	-0.04	-0.08	0.06	-0.2	-0.48	-0.05	0.34	-0.04
20:3 n-6	0.08	0.06	0.02	0.02	-0.01	0	0.01	-0.11	-0.21	-0.04	0.18	-0.06
20:2 n-6	0.31	0.31	0.08	0.13	-0.15	-0.17	0.04	-0.24	-0.5	-0.09	0.4	-0.1
18:3 n-6	0.09	0.11	0.08	0.1	-0.07	-0.13	0.04	-0.14	-0.18	-0.05	0.18	-0.08
18:2 n-6	0.22	0.21	0.13	0.15	-0.2	-0.19	-0.06	-0.17	-0.29	-0.12	0.03	-0.05
18:1 n-7	0.12	0.14	0.05	0.09	0.04	-0.01	0.05	0.06	-0.22	-0.1	0.19	-0.1
16:1 n-7	-0.21	-0.19	-0.1	-0.12	0.2	0.15	0.1	0.17	0.42	0.13	-0.08	0.04
14:1 n-7	-0.22	-0.19	-0.06	-0.08	0.13	0.11	0	0.15	0.44	0.11	-0.17	0.04
24:1 n-9	0.05	0.04	0.13	0.15	-0.14	-0.17	-0.05	0.13	-0.06	0.01	0.02	0.01
22:1 n-9	0.15	0.16	0.03	0.05	-0.03	-0.05	0	-0.07	-0.39	-0.14	0.21	-0.12
20:1 n-9	0.1	0.09	-0.02	-0.02	-0.07	-0.04	0.02	-0.17	-0.2	-0.13	0.13	-0.14
18:1 n-9	0.24	0.24	0.06	0.08	-0.1	-0.09	0.02	-0.16	-0.27	-0.07	0.17	-0.04
22:0	0.05	0.05	0.12	0.14	-0.17	-0.15	-0.09	0.07	-0.13	-0.09	0.03	-0.08
20:0	0.13	0.13	0.07	0.06	-0.07	-0.1	-0.05	-0.2	-0.34	-0.09	0.11	-0.05
18:0	0.04	0.04	0.01	0.02	-0.06	-0.07	-0.01	0.09	-0.29	-0.05	0.16	-0.04
16:0	-0.39	-0.38	-0.19	-0.21	0.25	0.26	-0.01	0.27	0.52	0.15	-0.26	0.1
14:0	-0.3	-0.27	-0.07	-0.1	0.13	0.13	-0.06	0.17	0.51	0.12	-0.27	0.07

Correlation
- 1.0
- 0.5
- 0.0
- -0.5
- -1.0

Fig. 3. Pearson correlation heatmap of TGFAs (mol%) with continuous basic and metabolic characteristics of PROMISE participants from the baseline visit (2004–2006). Darker orange represents a positive correlation; darker blue represents a negative correlation. Chol., cholesterol.

presented. We used the extracted PLS scores as the predictor variable in the GEE models and found negative associations of the first component on all outcome variables, with the strongest association being with the IS variables (estimate of 10.4, all $P < 0.001$; using PLS scores constrained by ISI). See supplemental Fig. S5 for a plot of the loadings of each TGFA on the two components.

DISCUSSION

In the present study, we found that in a Canadian cohort at risk for T2D, several specific TGFAs and groups of TGFAs were strongly associated with IS and moderately associated with β-cell function. In particular, the TGFAs myristic acid (14:0), 7-tetradecenoic acid (14:1n-7), palmitic acid (16:0), and palmitoleic acid (16:1n-7) all strongly and negatively associated with lower IS. Although most TGFAs were not associated with β-cell function, three FAs, palmitic acid

(16:0), cis-vaccenic acid (18:1n-7), and eicosadienoic acid (20:2n-6), were associated negatively and positively, respectively, with measures of β-cell function. Using PLS, we also found that four TGFAs (14:0, 14:1n-7, 16:0, and 16:1n-7) clustered together and that this cluster strongly predicted lower IS.

To our knowledge, no longitudinal study to date has examined the role of the composition of the TGFA fraction on detailed OGTT-derived metabolic measures. Two large prospective studies have been published that similarly examined TGFA composition and T2D outcomes. Rhee et al. presented a nested case-control analysis (n = 189 cases and n = 189 controls) within the Framingham offspring cohort (11), which found that subjects with a TGFA composition characterized by a lower carbon chain and fewer double bonds (e.g., 14:0 or 16:0) had a higher risk for T2D after 12 years, whereas those with a profile characterized by higher carbon chain and more double-bond TGFAs had a lower risk for T2D. A similar pattern of TGFAs was also associated
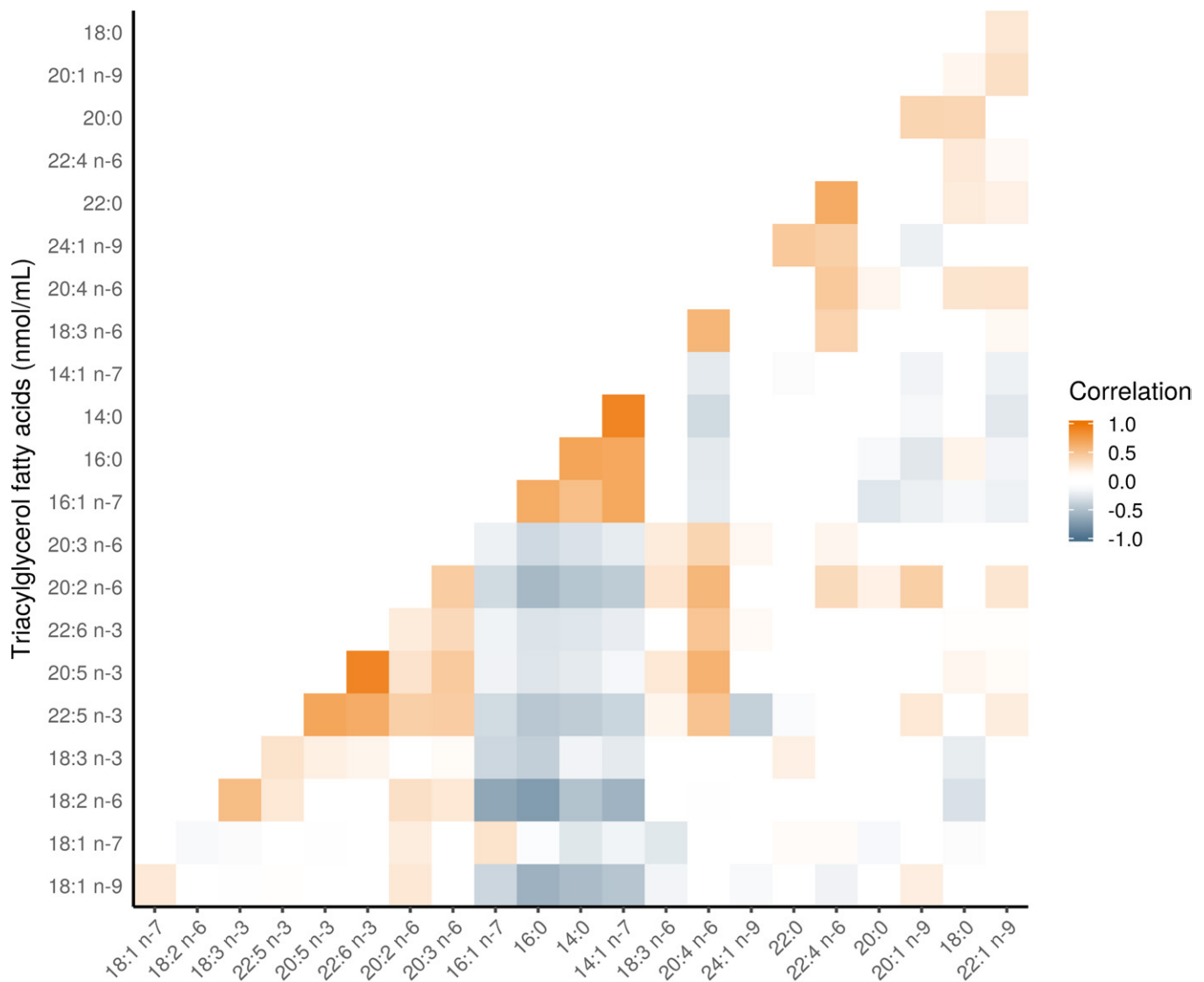
Fig. 4. Pearson correlation heatmap of TGFAs in the PROMISE participants from the baseline visit (2004–2006). The correlations of FAs grouped using hierarchical cluster analysis; FA along the *x* and *y* axes are ordered according to this analysis. Darker orange represents a positive correlation; darker blue represents a negative correlation.

with HOMA-IR cross-sectionally at the baseline visit. In addition, Lankinen et al. reported on a prospective cohort of males in Finland (12), for which TGFA data were available for 831 participants after 6 years of follow-up. In their cohort, OGTT data were only available at the 6 year visit. In the cross-sectional analysis, they found that most saturated FAs had negative associations with IS and β-cell function, whereas linoleic acid (18:2n-6), docosapentaenoic acid (22:5n-3), eicosapentaenoic acid (20:5n-3), and arachidonic acid (20:4n-6) had positive associations with IS. The magnitude of the associations were larger for the IS results compared with the β-cell function results, similar to what we observed. Our study extends these findings by using multiple measurements of metabolic function and as well as multivariate statistical approaches that allowed us to identify clusters of TGFAs. In another, much smaller study (n = 16) of mostly females (35), the authors reported a positive correlation between total esterified (of which TG make up the majority) 16:0, 16:1n-7, and 18:1n-9 with

HOMA-IR, findings which were largely similar to the present analysis.

There are a few possible explanations for these findings. Circulating TGFAs derive from three sources: adipose lipolysis, dietary fat, and de novo lipogenesis (DNL). Dietary carbohydrates and fat can influence DNL activity (36–38). Determining the specific source of TGFA is extremely difficult to ascertain outside of highly controlled experimental settings. Many previous studies that have examined DNL as the source have used markers of estimated DNL, such as the ratio between 18:2n-6 to 16:0 or 16:1n-7 to 16:0 (12, 39, 40). However, there are limitations to using these ratios, as the FAs used in their calculation can also be obtained from the diet, in addition to being created through DNL (39). An experimental feeding trial (n = 24) was conducted to identify the FAs that most accurately reflected DNL as potential biomarkers (41). The study found that palmitoleic acid (16:1n-7), directly measured DNL using isotopes, and liver fat were all highly correlated with each
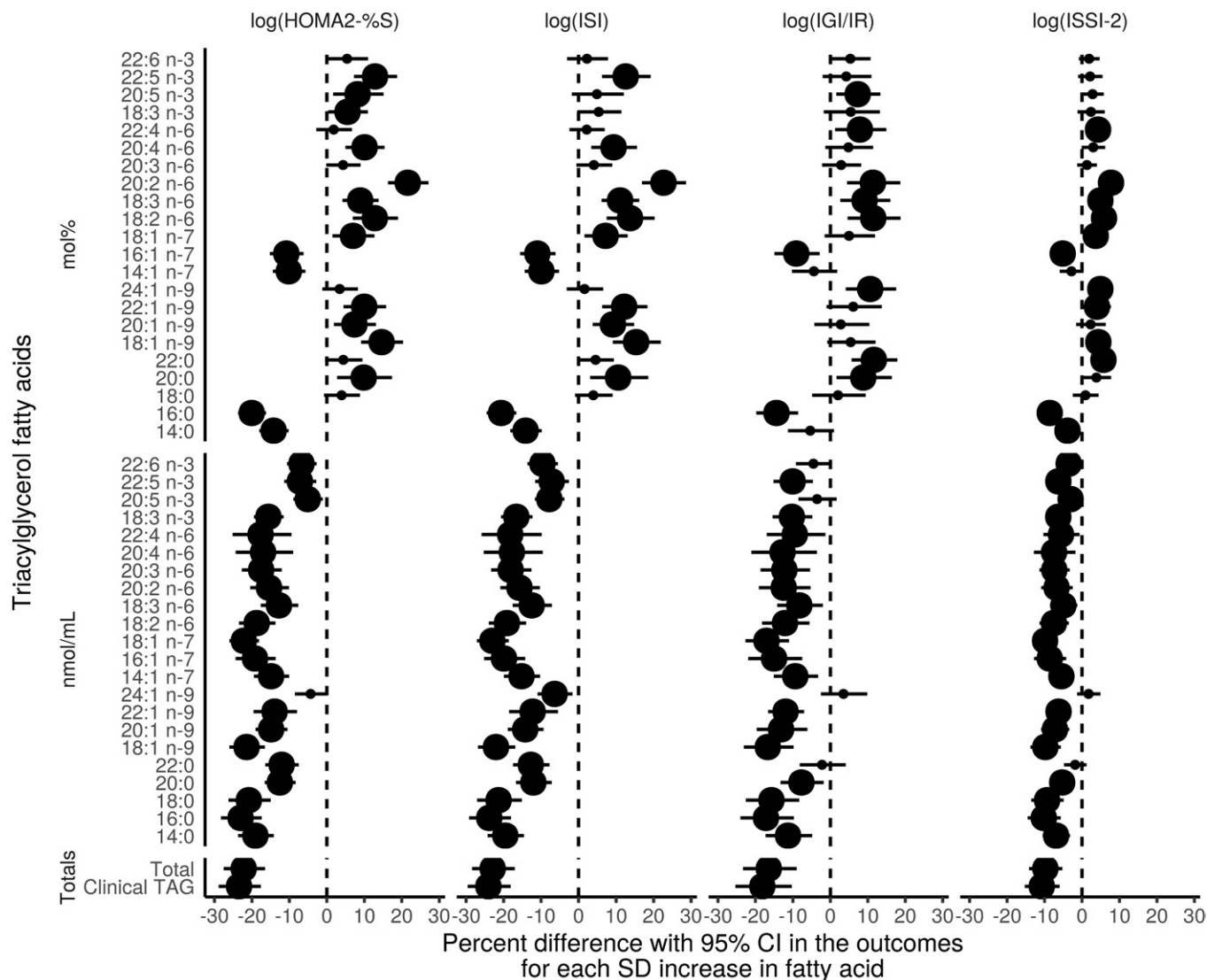
Fig. 5. Time-adjusted GEE models of the association of the TGFAs (mol% and nmol/ml) and total clinically measured TG with IS and β-cell function outcomes using the 6 year longitudinal data from the PROMISE cohort. The x-axis values represent a percent difference in the outcome per SD increase in the FA. *P* values were adjusted for the BH false discovery rate, with the largest dot representing a significant (*P* < 0.05) association.

other (*r* > 0.50), suggesting that 16:1n-7 may be a good biomarker for hepatic DNL. In another small (n = 14) feeding trial, meal type (high-fat vs low-fat) was tested to determine its effect on DNL and TGFA composition (42). The authors reported that 14:0, 16:0, 16:1, and 18:2 were higher in the low-fat (high-carbohydrate) group. In another recent overfeeding trial, 1,000 kcal of saturated fat, unsaturated fat, or carbohydrates over 3 weeks was given to 38 overweight individuals (38) to test changes in liver fat and hepatic DNL. At the end of the study, the carbohydrate group had higher DNL activity as well as an increase in liver fat, although the saturated fat group had the highest increase in liver fat. This link between carbohydrate intake and DNL activity has been well documented (1, 3, 4, 37, 39, 43, 44).

In our findings, the four FAs were highly positively correlated among each other and negatively or neutrally with all other TGFAs, in addition to clustering together on their

negative association with IS. This may suggest that greater DNL activity is the source of these TGFAs. Several studies have shown a link between higher estimated DNL activity and an increased risk for metabolic dysfunction (8, 12, 40, 45). How DNL may influence metabolic dysfunction is not well understood. Possible reasons may be that higher DNL produces more of certain FAs or that higher DNL increases circulating TG, which itself is well documented to contribute to metabolic dysfunction and which we found in our study as a high positive correlation between the four TGFA and clinical TG.

Regardless of the exact source of these FAs, our results, in addition to the available scientific evidence, emphasize the importance of the FA composition on metabolic health, as individual FAs can have specific physiological functions. For instance, a higher concentration of circulating 14 and 16 carbon FAs may expose tissues to greater lipotoxicity, for instance, from palmitic acid (16:0), which is well known
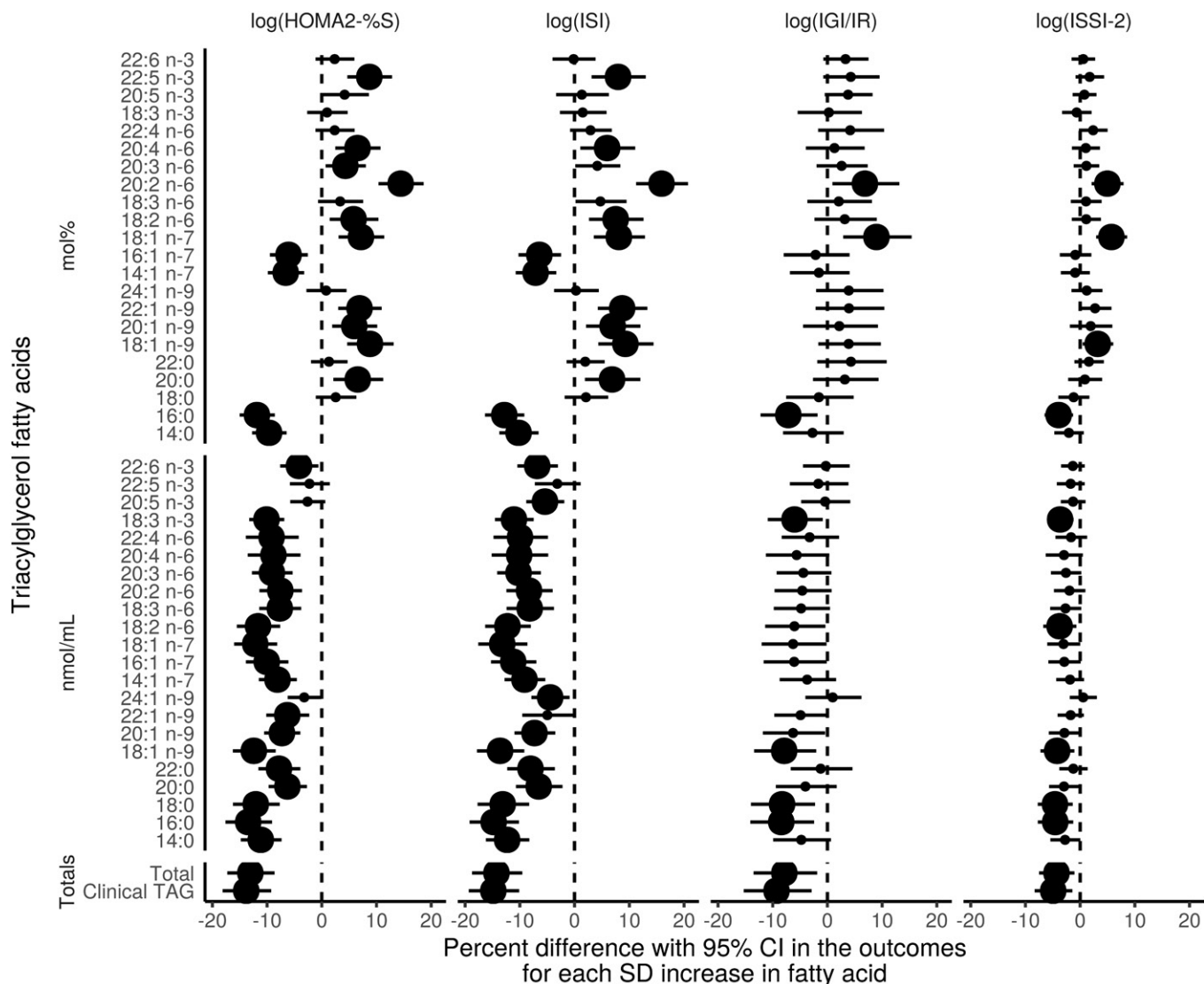
Fig. 6. Fully-adjusted GEE models of the association of the TGFAs (mol% and nmol/ml) and total clinically measured TG with IS and β-cell function outcomes using the 6 year longitudinal data from the PROMISE cohort. Variables controlled for were follow-up time, WC, baseline age, ethnicity, sex, ALT, physical activity, and total NEFA. The x-axis values represent a percent difference in the outcome per SD increase in the FA. P values were adjusted for the BH false discovery rate, with the largest dot representing a significant ($P < 0.05$) association.

to have harmful effects on tissues (46, 47). Our study extends these findings by showing that TGFAs with 14 to 16 carbons clustered together and this pattern strongly associated with lower IS. Although some of these FAs also had a significant association with β-cell function, the magnitude of associations were more modest compared with those for IS.

The direction of association between TGFAs and IS is unclear from previous cross-sectional studies due to the physiological feedback mechanisms involved. For example, although higher TGFAs may promote muscle insulin resistance, the reverse may also be true (48). As we found no interaction by time of TGFAs on IS, this study cannot determine the exact role of the feedback mechanism. However, by combining the lack of a time interaction and the consistent negative association in models without the time interaction, these results at least suggest that the feedback mechanism may not be strongly influential and that TGFAs

may predict IS at least over a 6 year period. Given the complex biological mechanisms and feedback loops involved, disentangling whether IS influences TGFAs more strongly than TGFAs influences IS will require more complex research designs and analyses.

Given the close biological relationship between circulating NEFAs and TG, NEFA may act as a confounding factor and was thus adjusted for. In our published analysis of the NEFA fraction (10), we found that higher total NEFAs, but not the specific composition, associated with lower β-cell function. This is in contrast to the TGFA findings that the specific composition does differentially associate with IS and β-cell function, adjusting for total NEFAs. There was no difference in results in models that did not include NEFA as a confounder (data not shown). This difference in results between NEFA and TGFA suggests that TGFA may independently and strongly influence the pathophysiology of T2D, when compared with other lipid fraction
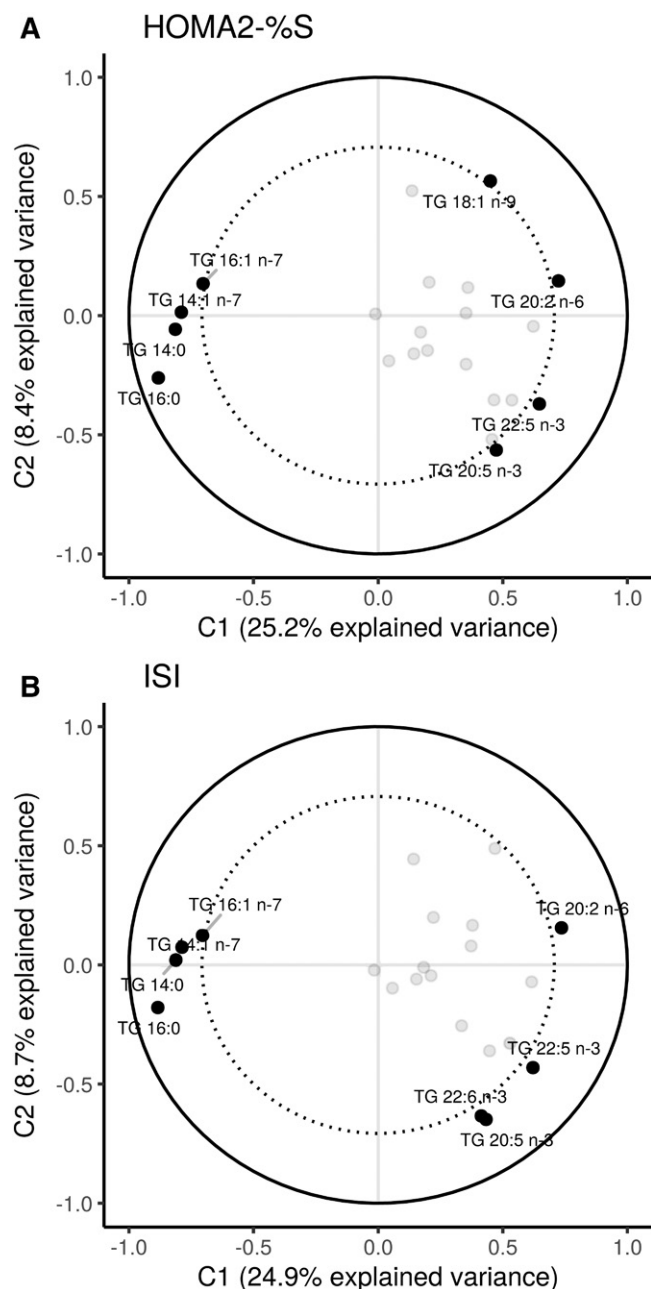
Fig. 7. PLS models showing the clustering of TGFAs on IS measures. The percent explained variance of each component is shown in brackets on each axis. The solid line represents an explained variance of 100%, while the dashed line represents an explained variance of 50%. FAs between these lines represent variables that strongly explain the underlying structure of the data.

compositions, including the phospholipid and cholesteryl ester fractions (9). This may be due to TG being biologically destined for uptake by nonhepatic tissue, as they are found mainly in VLDLs, at least during fasting. This is in contrast to NEFAs that are mostly taken up by the liver and used in TG production (49).

Our study has potential limitations that need to be considered when interpreting the results. First, this is an observational cohort, and as such, there may be some residual confounding we were not able to control for or were unaware of. However, we have taken extensive, empirically based precautions in identifying potential confounders and mediators through the use of the DAG modeling, relying on previous literature, and through information criteria model fit comparison methods. Only fasting TGFAs were quantified and only at the baseline visit. TGFA composition can fluctuate substantially throughout the day, so in order to control for this, PROMISE participants came for the clinic visit in the morning and fasted. There is some evidence to suggest that fasting TG is better able to discriminate diabetes cases compared with a postprandial state (11). Because TGFAs were only measured at the baseline visit, we cannot investigate whether there are concomitant changes in TGFAs and the metabolic measures over time. However, to optimally use GEE to analyze the data and for interpretation, we used the model to infer that a given value of TGFA could predict values of IS or β-cell function over a 6 year period. This, in our view, is a strength of our analysis, as it reduces the chance of reverse causality given the tight integration of the glucose and FA metabolism pathways, as well as maximizes the specific usage of the GEE modeling.

PLS is a well-established technique for constructing predictive models of high-dimensionality data structures (i.e., FA composition); however, a limitation is that the initial models analyzed through PLS and the final computed scores are not able to control for potential confounders and other effect modifiers. PLS is also not able to handle longitudinal data, so only the baseline visit was used in the PLS analysis, although we analyzed the extracted scores using the GEE modeling to overcome this limitation and observed concordant results between the PLS and GEE analyses.

Our study has several notable strengths, including the longitudinal design and the use of advanced statistical techniques for data analysis. These statistical techniques take advantage of the longitudinal data to allow appropriate investigation of temporal relationships and are able to handle the multidimensional nature of the data. Finally, our cohort contains highly detailed and comprehensive variable measurements for the FAs and outcomes, which were collected at each visit.

CONCLUSION

In conclusion, we found that a TGFA composition containing higher proportions of 14:0, 14:1n-7, 16:0, and 16:1n-7 associated strongly with lower IS and (more moderately) with lower β-cell function. We also found that most other TGFAs (e.g., 20:0, most omega 6 and 9 TGFAs, and 20:5n-3) associated positively with IS. Only a few TGFAs associated positively and consistently with β-cell function (e.g., 18:1n-7 and 20:2n-6). These results provide more insight into how individual TGFAs contribute to the pathogenesis of T2D while reinforcing the importance and value of clinically measured total TG as an indicator of metabolic health.

and Sheila Porter and Mauricio Marin of the Centre for Studies in Family Medicine, University of Western Ontario, London, Canada for their expert technical assistance and dedication in their work for PROMISE.

REFERENCES

1. Chehade, J. M., M. Gladysz, and A. D. Mooradian. 2013. Dyslipidemia in type 2 diabetes: prevalence, pathophysiology, and management. *Drugs.* **73:** 327–339.

2. D'Agostino, R. B., R. F. Hamman, A. J. Karter, L. Mykkanen, L. E. Wagenknecht, and S. M. Haffner. 2004. Cardiovascular disease risk factors predict the development of type 2 diabetes: The Insulin Resistance Atherosclerosis Study. *Diabetes Care.* **27:** 2234–2240.

3. Vergès, B. 2015. Pathophysiology of diabetic dyslipidaemia: where are we? *Diabetologia.* **58:** 886–899.

4. Kawano, Y., and D. E. Cohen. 2013. Mechanisms of hepatic triglyceride accumulation in non-alcoholic fatty liver disease. *J. Gastroenterol.* **48:** 434–441.

5. Alberti, K. G., R. H. Eckel, S. M. Grundy, P. Z. Zimmet, J. I. Cleeman, K. A. Donato, J-C. Fruchart, W. P. T. James, C. M. Loria, and S. C. Smith. 2009. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation.* **120:** 1640–1645.

6. Lemieux, I., A. Pascot, C. Couillard, B. Lamarche, A. Tchernof, N. Alméras, J. Bergeron, D. Gaudet, G. Tremblay, D. Prud'homme, et al. 2000. Hypertriglyceridemic waist: a marker of the atherogenic metabolic triad (hyperinsulinemia; hyperapolipoprotein B; small, dense LDL) in men? *Circulation.* **102:** 179–184.

7. Forouhi, N. G., A. Koulman, S. J. Sharp, F. Imamura, J. Kröger, M. B. Schulze, F. L. Crowe, J. M. Huerta, M. Guevara, J. W. J. Beulens, et al. 2014. Differences in the prospective association between individual plasma phospholipid saturated fatty acids and incident type 2 diabetes: the EPIC-InterAct Case-Cohort Study. *Lancet Diabetes Endocrinol.* **2:** 810–818.

8. Ma, W., J. H. Y. Wu, Q. Wang, R. N. Lemaitre, K. J. Mukamal, L. Djoussé, I. B. King, X. Song, M. L. Biggs, J. A. Delaney, et al. 2015. Prospective association of fatty acids in the de novo lipogenesis pathway with risk of type 2 diabetes: The Cardiovascular Health Study. *Am. J. Clin. Nutr.* **101:** 153–163.

9. Johnston, L. W., S. B. Harris, R. Retnakaran, B. Zinman, A. Giacca, Z. Liu, R. P. Bazinet, and A. J. Hanley. 2016. Longitudinal associations of phospholipid and cholesteryl ester fatty acids with disorders underlying diabetes. *J. Clin. Endocrinol. Metab.* **101:** 2536–2544.

10. Johnston, L. W., S. B. Harris, R. Retnakaran, A. Giacca, Z. Liu, R. P. Bazinet, and A. J. Hanley. 2017. Association of non-esterified fatty acid composition with insulin sensitivity and beta cell function in the Prospective Metabolism and Islet Cell Evaluation (PROMISE) cohort. *Diabetologia.* **61:** 821–830.

11. Rhee, E. P., S. Cheng, M. G. Larson, G. A. Walford, G. D. Lewis, E. McCabe, E. Yang, L. Farrell, C. S. Fox, C. J. O'Donnell, et al. 2011. Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *J. Clin. Invest.* **121:** 1402–1411.

12. Lankinen, M. A., A. Stančáková, M. Uusitupa, J. Ågren, J. Pihlajamäki, J. Kuusisto, U. Schwab, and M. Laakso. 2015. Plasma fatty acids as predictors of glycaemia and type 2 diabetes. *Diabetologia.* **58:** 2533–2544.

13. Yu, S. S., D. C. Castillo, A. B. Courville, and A. E. Sumner. 2012. The triglyceride paradox in people of African descent. *Metab. Syndr. Relat. Disord.* **10:** 77–82.

14. Chien, K., T. Cai, H. Hsu, T. Su, W. Chang, M. Chen, Y. Lee, and F. B. Hu. 2009. A prediction model for type 2 diabetes risk among Chinese people. *Diabetologia.* **52:** 443–450.

15. Schulze, M. B., C. Weikert, T. Pischon, M. M. Bergmann, H. Al-Hasani, E. Schleicher, A. Fritsche, H-U. Haring, H. Boeing, and H-G. Joost. 2009. Use of multiple metabolic and genetic markers to improve the prediction of type 2 diabetes: The EPIC-Potsdam Study. *Diabetes Care.* **32:** 2116–2119.

16. Matthan, N. R., B. Ip, N. Resteghini, L. M. Ausman, and A. H. Lichtenstein. 2010. Long-term fatty acid stability in human serum cholesteryl ester, triglyceride, and phospholipid fractions. *J. Lipid Res.* **51:** 2826–2832.

17. Folch, J., M. Lees, and G. H. Sloane Stanley. 1957. A simple method for the isolation and purification of total lipides from animal tissues. *J. Biol. Chem.* **226:** 497–509.

18. Nishi, S., C. W. C. Kendall, A-M. Gascoyne, R. P. Bazinet, B. Bashyam, K. G. Lapsley, L. S. A. Augustin, J. L. Sievenpiper, and D. J. A. Jenkins. 2014. Effect of almond consumption on the serum fatty acid profile: a dose-response study. *Br. J. Nutr.* **112:** 1137–1146.

19. Kriska, A. M., W. C. Knowler, R. E. LaPorte, A. L. Drash, R. R. Wing, S. N. Blair, P. H. Bennett, and L. H. Kuller. 1990. Development of questionnaire to examine relationship of physical activity and diabetes in Pima Indians. *Diabetes Care.* **13:** 401–411.

20. Matsuda, M., and R. A. DeFronzo. 1999. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. *Diabetes Care.* **22:** 1462–1470.

21. Levy, J. C., D. R. Matthews, and M. P. Hermans. 1998. Correct homeostasis model assessment (HOMA) evaluation uses the computer program. *Diabetes Care.* **21:** 2191–2192.

22. Abdul-Ghani, M. A., M. Matsuda, B. Balas, and R. A. DeFronzo. 2007. Muscle and liver insulin resistance indexes derived from the oral glucose tolerance test. *Diabetes Care.* **30:** 89–94.

23. Wareham, N. J., D. I. Phillips, C. D. Byrne, and C. N. Hales. 1995. The 30 minute insulin incremental response in an oral glucose tolerance test as a measure of insulin secretion. *Diabet. Med.* **12:** 931.

24. Matthews, D. R., J. P. Hosker, A. S. Rudenski, B. A. Naylor, D. F. Treacher, and R. C. Turner. 1985. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia.* **28:** 412–419.

25. Retnakaran, R., Y. Qi, M. I. Goran, and J. K. Hamilton. 2009. Evaluation of proposed oral disposition index measures in relation to the actual disposition index. *Diabet. Med.* **26:** 1198–1203.

26. Hermans, M. P., J. C. Levy, R. J. Morris, and R. C. Turner. 1999. Comparison of insulin sensitivity tests across a range of glucose tolerance from normal to diabetes. *Diabetologia.* **42:** 678–687.

27. Zeger, S. L., and K. Y. Liang. 1986. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* **42:** 121–130.

28. Greenland, S., J. Pearl, and J. M. Robins. 1999. Causal diagrams for epidemiologic research. *Epidemiology.* **10:** 37–48.

29. Textor, J., J. Hardt, and S. Knüppel. 2011. DAGitty: a graphical tool for analyzing causal diagrams. *Epidemiology.* **22:** 745.

30. Shrier, I., and R. W. Platt. 2008. Reducing bias through directed acyclic graphs. *BMC Med. Res. Methodol.* **8:** 70.

31. R Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/

32. Højsgaard, S., U. Halekoh, and J. Yan. 2006. The R package geepack for generalized estimating equations. *J. Stat. Softw.* **15:** 1–11.

33. Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **57:** 289–300.

34. Vandenbroucke, J. P., E. von Elm, D. G. Altman, P. C. Gøtzsche, C. D. Mulrow, S. J. Pocock, C. Poole, J. J. Schlesselman, M. Egger, and STROBE Initiative. 2007. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med.* **4:** e297.

35. Kotronen, A., V. R. Velagapudi, L. Yetukuri, J. Westerbacka, R. Bergholm, K. Ekroos, J. Makkonen, M-R. Taskinen, M. Oresic, and H. Yki-Järvinen. 2009. Serum saturated fatty acids containing triacylglycerols are better markers of insulin resistance than total serum triacylglycerol concentrations. *Diabetologia.* **52:** 684–690.

36. Aarsland, A., and R. R. Wolfe. 1998. Hepatic secretion of VLDL fatty acids during stimulated lipogenesis in men. *J. Lipid Res.* **39:** 1280–1286.

37. Harding, S. V., K. P. Bateman, B. P. Kennedy, T. C. Rideout, and P. J. H. Jones. 2015. Desaturation index versus isotopically measured de novo lipogenesis as an indicator of acute systemic lipogenesis. *BMC Res. Notes.* **8:** 49.

38. Luukkonen, P. K., S. Sädevirta, Y. Zhou, B. Kayser, A. Ali, L. Ahonen, S. Lallukka, V. Pelloux, M. Gaggini, C. Jian, et al. Saturated fat is more metabolically harmful for the human liver than unsaturated fat or simple sugars. *Diabetes Care.* Epub ahead of print. May 29, 2018; doi:10.2337/dc18-0071.

39. Hodson, L., C. M. Skeaff, and B. A. Fielding. 2008. Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. *Prog. Lipid Res.* **47:** 348–380.

40. Kröger, J., V. Zietemann, C. Enzenbach, C. Weikert, E. H. Jansen, F. Döring, H-G. Joost, H. Boeing, and M. B. Schulze. 2011. Erythrocyte membrane phospholipid fatty acids, desaturase activity, and dietary fatty acids in relation to risk of type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. *Am. J. Clin. Nutr.* **93:** 127–142.

41. Lee, J. J., J. E. Lambert, Y. Hovhannisyan, M. A. Ramos-Roman, J. R. Trombold, D. A. Wagner, and E. J. Parks. 2015. Palmitoleic acid is elevated in fatty liver disease and reflects hepatic lipogenesis. *Am. J. Clin. Nutr.* **101:** 34–43.

42. Wilke, M. S., M. A. French, Y. K. Goh, E. A. Ryan, P. J. Jones, and M. T. Clandinin. 2009. Synthesis of specific fatty acids contributes to VLDL-triacylglycerol composition in humans with and without type 2 diabetes. *Diabetologia.* **52:** 1628–1637.

43. Hudgins, L. C. 2000. Effect of high-carbohydrate feeding on triglyceride and saturated fatty acid synthesis. *Proc. Soc. Exp. Biol. Med.* **225:** 178–183.

44. Parks, E. J., R. M. Krauss, M. P. Christiansen, R. A. Neese, and M. K. Hellerstein. 1999. Effects of a low-fat, high-carbohydrate diet on VLDL-triglyceride assembly, production, and clearance. *J. Clin. Invest.* **104:** 1087–1096.

45. Zong, G., J. Zhu, L. Sun, X. Ye, L. Lu, Q. Jin, H. Zheng, Z. Yu, Z. Zhu, H. Li, et al. 2013. Associations of erythrocyte fatty acids in the de novo lipogenesis pathway with risk of metabolic syndrome in a cohort study of middle-aged and older Chinese. *Am. J. Clin. Nutr.* **98:** 319–326.

46. Risérus, U. 2008. Fatty acids and insulin sensitivity. *Curr. Opin. Clin. Nutr. Metab. Care.* **11:** 100–105.

47. Iggman, D., J. Arnlöv, B. Vessby, T. Cederholm, P. Sjögren, and U. Risérus. 2010. Adipose tissue fatty acids and insulin sensitivity in elderly men. *Diabetologia.* **53:** 850–857.

48. Flannery, C., S. Dufour, R. Rabøl, G. I. Shulman, and K. F. Petersen. 2012. Skeletal muscle insulin resistance promotes increased hepatic de novo lipogenesis, hyperlipidemia, and hepatic steatosis in the elderly. *Diabetes.* **61:** 2711–2717.

49. Nielsen, S., and F. Karpe. 2012. Determinants of VLDL-triglycerides production. *Curr. Opin. Lipidol.* **23:** 321–326.