Weight-independent effects of dietary carbohydrate-to-fat ratio on metabolomic profiles: secondary outcomes of a 5-month randomized

nature communications

Article

https://doi.org/10.1038/s41467-026-68353-z

Weight-independent effects of dietary carbohydrate-to-fat ratio on metabolomic profiles: secondary outcomes of a 5-month randomized controlled feeding trial

Received: 23 September 2024

Accepted: 5 January 2026

Published online: 17 January 2026

Check for updates

A list of authors and their affiliations appears at the end of the paper

Diet plays a crucial role in health, with low-carbohydrate diets often proposed to exert metabolic benefits. We aim to investigate metabolomic adaptations in 164 adults with overweight or obesity who were randomly assigned to high- ($n = 54$), moderate- ($n = 53$), or low-carbohydrate ($n = 57$) diets during a 20-week weight-loss maintenance phase of the Framingham State Food Study [(FS)2], a controlled, parallel feeding trial (ClinicalTrials.gov: NCT02068885). We measure fasting plasma metabolites by liquid chromatography-tandem mass spectrometry using samples from 147 participants who completed the study ($n = 45$, 48, and 54 in the high-, moderate-, and low-carbohydrate diet groups, respectively). Significant associations (False Discovery Rate<0.05) are identified between carbohydrate-to-fat ratio (CFR) and diet-induced changes in 148 of 479 metabolites at 20 weeks, with nearly all showing consistent trends at 10 and 20 weeks. Phosphatidylcholines plasmanyls/plasmalogens, phosphatidylethanolamines plasmanyls/plasmalogens, and sphingomyelins generally decrease with higher CFR, whereas lysophosphatidylcholines, lysophosphatidylethanolamines, and triglycerides generally increase. Our findings are largely reproducible in an independent feeding trial involving diets with similar CFR (Popular Diets Study, ClinicalTrials.gov: NCT00315354). Eleven triglyceride species (≤3 double bonds), linked to type 2 diabetes risk, increase with higher CFR. Our findings demonstrate metabolomic changes caused by varying CFR dietary patterns, offering potential insights into mechanisms that could guide targeted dietary intervention strategies.

Diet plays a crucial role in human health and well-being and is a major contributor to noncommunicable diseases such as obesity, cardiovascular disease (CVD), and type 2 diabetes (T2D)[1,2]. Evidence from dietary intervention studies suggests that reduced-calorie diets, irrespective of specific macronutrient composition, may contribute to clinically meaningful weight loss and potentially improve cardiometabolic risk factors such as lipid profiles and insulin sensitivity[3]. However, certain dietary components are among the leading modifiable risk factors associated with disease burden and may impact individuals of all sociodemographic backgrounds[4]. Typical dietary patterns, characterized by differing carbohydrate-to-fat ratio (CFR), may exert significant influence on health outcomes, including cardiometabolic diseases, their associated risk factors, and energy expenditure[5–9]. In addition, it has been shown that although both free-sugar and ketogenic carbohydrate restriction can reduce body fat, they may elicit divergent metabolic and physiological effects, including alterations in

e-mail: cara.ebbeling@childrens.harvard.edu; joel.hirschhorn@childrens.harvard.edu

glucose tolerance, lipoprotein profiles, and gut microbial diversity, which further underscores that macronutrient composition may differentially influence cardiometabolic health[10]. Moreover, emerging metabolomic evidence highlights that individuals with obesity may exhibit distinct metabolic signatures associated with tissue-specific insulin resistance and show variable responses to dietary interventions, suggesting that individual metabolic phenotypes could play a critical role in shaping the efficacy of diet-based strategies for improving cardiometabolic outcomes[11,12].

However, much of the existing evidence in the field is derived from observational studies, which are inherently subject to confounding and other methodological limitations. Combining metabolomic technologies with dietary randomized controlled trials (RCTs) can be a potent tool in nutrition research, providing insights into the causal effects of dietary patterns on metabolism and potentially into diet-related disease risks[13,14]. In the setting of a dietary RCT, nutritional metabolomic profiles may serve as direct markers of dietary consumption and can offer deeper insights into biological mechanisms and molecular pathways, thereby paving the way for the development of targeted dietary interventions to reduce disease burden[15].

In our study, we conducted untargeted metabolite profiling using liquid chromatography tandem mass spectrometry (LC-MS) in individuals with overweight or obesity participating in an RCT with a feeding protocol, the primary outcomes of which have been previously published[8]. All analyses described in the present manuscript represent secondary outcomes of this trial. First, we aimed to explore differences in metabolite profiles for diets with different CFR (i.e., low-, moderate-, and high-carbohydrate diets) and identify changes in metabolic biomarkers caused by exposure to these diets. Second, we aimed to determine whether diet-responsive biomarkers could be used to assess adherence to a specific diet. Third, we aimed to test whether diet-induced metabolite signatures were sustained throughout the dietary intervention, thereby exploring the effects of potential metabolic adaptation. Fourth, to validate our findings, we utilized data from the Popular Diets Study, a previously published randomized crossover feeding trial[16], for which the primary outcomes have already been reported[17]. All analyses conducted using data from the Popular Diets Study also represent secondary outcomes. Finally, we leveraged data from a previously published meta-analysis focused on T2D[18] to explore whether metabolites associated with higher or lower CFR predicted future T2D incidence.

Results

Study participants

The Framingham State Food Study [(FS)2] has been described previously[19] and is summarized in Supplementary Fig. 1. All analyses described in the present manuscript are secondary outcomes of the (FS)2 trial. In the (FS)2, 164 participants achieved at least 10% weight loss during a run-in phase and then were randomly assigned to a test phase for weight-loss maintenance on one of three diets varying in CFR (carbohydrate as 20, 40 and 60% of calories, and fat as 60, 40 and 20% of calories: $n = 57$, 53, and 54, for the low-, moderate- and high-CFR groups, respectively). Of these, 148 individuals completed the study. Dropout rates across the three groups were not statistically significant ($P = 0.237$): 3 participants (5.26%) in the low-CFR group, 5 participants (9.43%) in the moderate-CFR group, and 8 participants (14.8%) in the high-CFR group (Supplementary Fig. 2). After applying a previously described quality control (QC) pipeline implemented in PAIRUP-MS[20], data from one sample did not meet the specified QC criteria, thereby excluding one participant in the high-CFR group from the analyses. The number of participants achieving weight-loss maintenance during the test phase did not differ significantly among groups, $P = 0.387$: 42, 37, and 30 participants in the low-, moderate-, and high-CFR groups respectively maintained body weight within ± 2 kg; 4, 3, and 3 respectively gained >2 kg; and 8, 8, and 12 respectively lost >2 kg (Supplementary Fig. 2). Table 1 presents the main descriptive and clinical characteristics of all study participants at randomization.

Metabolite differences across diets with higher CFR

To identify metabolites influenced by dietary CFR, we calculated changes in fasting plasma samples from randomization [START] to the end of the 20-week test phase [END]. We studied 479 plasma

Table 1 | Descriptive and clinical characteristics of the study participants at randomization

Characteristics	Total population	LC group	MC group	HC group	P-value
N	147	54	48	45	
Age (years)	38.62 (14.47)	37.18 (13.6)	38.77 (14.86)	40.21 (15.21)	0.695
Female, yes (%)	102 (69.39)	34 (62.96)	32 (66.67)	36 (80.00)	0.165
Race:*					
White (%)	115 (78.23)	41 (75.93)	38 (79.17)	36 (80.00)	0.868
Black (%)	16 (10.88)	6 (11.11)	6 (12.50)	4 (8.89)	
Asian (%)	5 (3.40)	1 (1.85)	2 (4.17)	2 (4.44)	
Other/Not Sure (%) **	11 (7.48)	6 (11.11)	2 (4.17)	3 (6.67)	
Hispanic ethnicity (%)	20 (13.61)	9 (16.67)	6 (12.50)	5 (11.11)	0.698
BMI (kg/m^2)	28.8 (4.29)	28.65 (4.3)	29.66 (4.6)	28.06 (3.85)	0.259
Body composition:					
Lean body mass (% of total mass)§	60.2 (6.39)	61.05 (5.74)	60.11 (7.57)	59.27 (5.73)	0.327
Body fat mass (% of total mass)	36.84 (6.68)	35.96 (6)	37 (7.92)	37.75 (6)	0.410
Weight loss (% of pre-weight loss)	−10.47 (1.58)	−10.35 (1.57)	−10.48 (1.56)	−10.61 (1.64)	0.714
delta Weight change during test phase (kg)	−0.58 (2.01)	−0.67 (2.07)	−0.62 (2.09)	−0.44 (1.88)	0.824

Data presented as mean (SD) for continuous variables and *n* (%) as categorical variables. Comparisons between groups were conducted using ANOVA or Kruskal-Wallis tests, depending on the distribution of the continuous variables. For categorical variables, the Chi-squared test with Yates' continuity correction was used, or Fisher's Exact Test when expected cell counts were less than 5. All statistical tests were two-sided.

*Race was determined by self-report using predefined categories that were standard when the trial was implemented: White, Black/African American, Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, Other, and Not Sure.

**Among these 11 participants, 10 reported "Other/Not Sure" for race (and also reported Hispanic ethnicity), and one specified multiple races (White, Black/African American, American Indian/Alaska Native)

§ Lean body mass does not include bone mineral content

HC High Carbohydrate group, *LC* Low Carbohydrate group, *MC* Moderate Carbohydrate group, *N* number.

metabolites identified by the Broad metabolomics platform that passed QC; of these, the majority were lipids and lipid-like molecules ($n = 325$), followed by organic acids and derivatives ($n = 70$) and organoheterocyclic compounds ($n = 37$). Of the 479 metabolites, 148 showed a significant association with CFR (trend from low- to moderate- to high-CFR), after false discovery rate (FDR) correction (FDR < 0.05). The metabolite superclasses of the 148 associated metabolites included lipids and lipid-like molecules ($n = 122$), organic acids and derivatives ($n = 19$), organic nitrogen compounds ($n = 3$), organoheterocyclic compounds ($n = 3$), and organic oxygen compounds ($n = 1$) (Supplementary Data 1). We observed a significant reduction in 87 metabolites with higher CFR, mainly in phosphatidylcholines, plasmanyls/plasmalogens, phosphatidylethanolamines plasmanyls/plasmalogens, sphingomyelins, amino acids, fatty acids, and acylcarnitines. We observed a significant increase in 61 metabolite levels with higher CFR, mainly in lysophosphatidylcholines, lysophosphatidylethanolamines, triglycerides, and diglycerides (Fig. 1). Many associated metabolites were correlated ($r > 0.6$), especially within classes of metabolites (network plot, Supplementary Fig. 3).

Because many of the diet-induced changes in metabolites affected triglycerides and related lipids, we further explored the effect of CFR on triglycerides, diglycerides, and fatty acids with different numbers of carbons and double bonds. Our analyses revealed a nonlinear association between higher CFR and a higher number of carbons in triglycerides ($R^2 = 0.399$, 95% confidence interval (CI) 0.354 to 0.691 for the linear term and −0.007 to −0.004 for the quadratic term, $P = 3.82e\text{-}07$), and a linear trend of higher CFR with lower number of carbons in diglycerides ($R^2 = 0.584$, 95% CI −0.076 to −0.001, $P = 0.045$; Fig. 2). Of note, triglycerides with levels positively and significantly associated with higher CFR had carbon chain lengths ranging from 48 to 56, and mostly from 48 to 52, suggesting that the influence of CFR on circulating levels varies with fatty acid composition of the triglycerides. However, no significant association was identified between higher CFR and the number of carbon atoms in fatty acids ($R^2 = 0.051$, 95% CI −0.002 to 0.016, $P = 0.140$; Supplementary Fig. 4). The FDR-significant metabolites from fatty acids, diglycerides, and most triglycerides contained an even number of carbon atoms. No significant association was found when examining the number of double bonds with CFR ($R^2 = 0.166$, 95% CI −0.084 to 0.037, $P = 0.365$ and $R^2 = 0.054$, 95% CI −0.001 to 0.032, $P = 0.071$ for diglycerides and triglycerides, respectively). While a weak association was observed between the number of double bonds in fatty acids and the change in metabolite level with higher CFR ($R^2 = 0.176$, 95% CI 0.013 to 0.065, $P = 0.005$), all five FDR-significant fatty acids were saturated, suggesting a potential effect of CFR on this group of fatty acids.

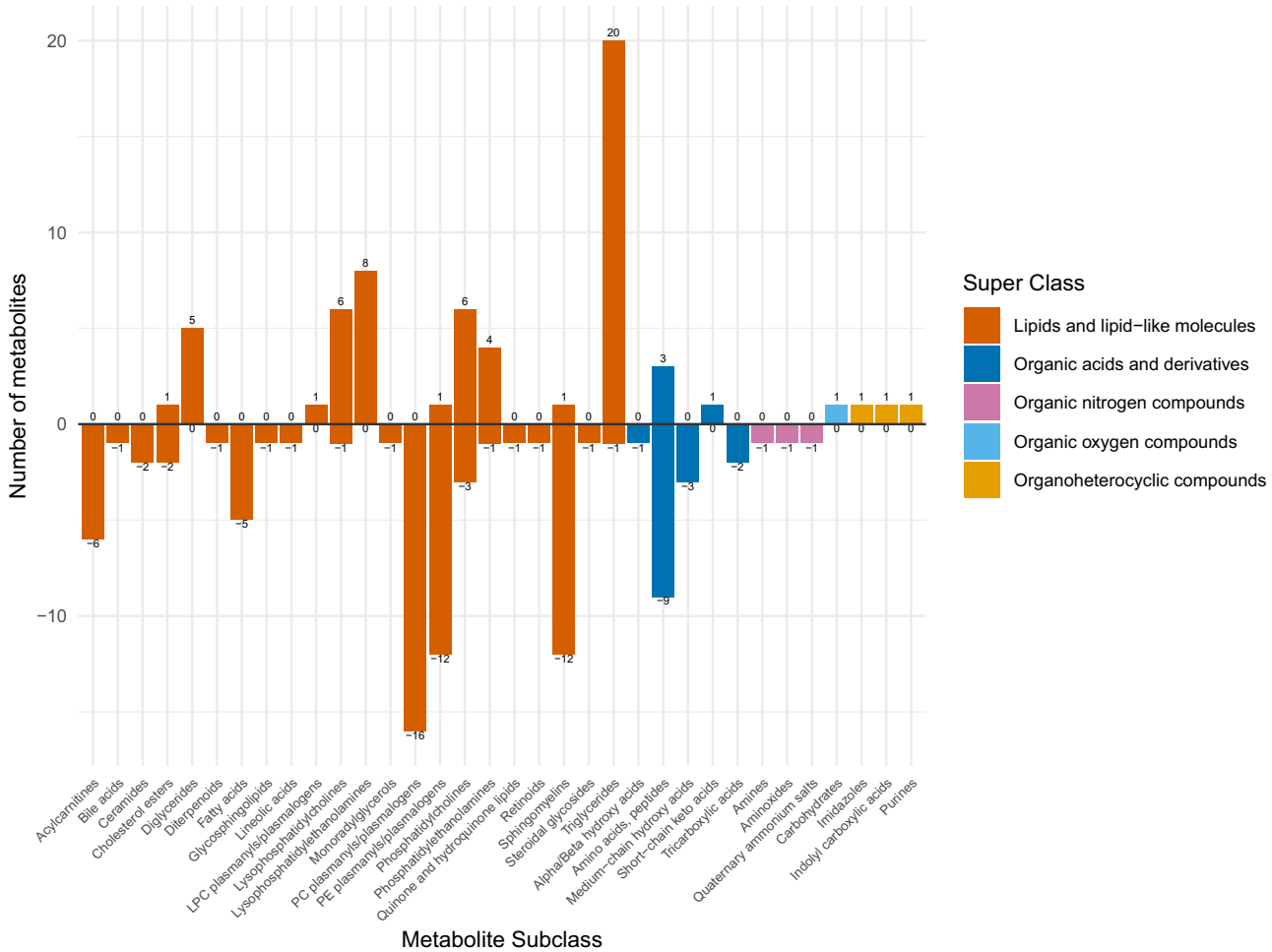


Fig. 1 | False Discovery Rate significant metabolite level changes from low- to moderate- to high-carbohydrate diets: grouped by subclass. The bar plot illustrates the metabolite level changes observed from low- to moderate- to high-carbohydrate diets (higher carbohydrate-to-fat ratio, CFR). Each column represents a group of metabolites belonging to a specific subclass. The columns indicate the number of metabolites exhibiting statistically significant changes, corrected for False Discovery Rate (FDR) using a cutoff of <0.05. Positive values on the *y*-axis correspond to the number of metabolites with increased metabolite level changes with higher CFR, while negative values correspond to the number of metabolites with decreased metabolite level changes with higher CFR. Source data are provided as a Source Data file. *CFR* carbohydrate-to-fat ratio; *LPC* lysophosphatidylcholines; *PC* phosphatidylcholines; *PE* phosphatidylethanolamines.

https://doi.org/10.1038/s41467-026-68353-z

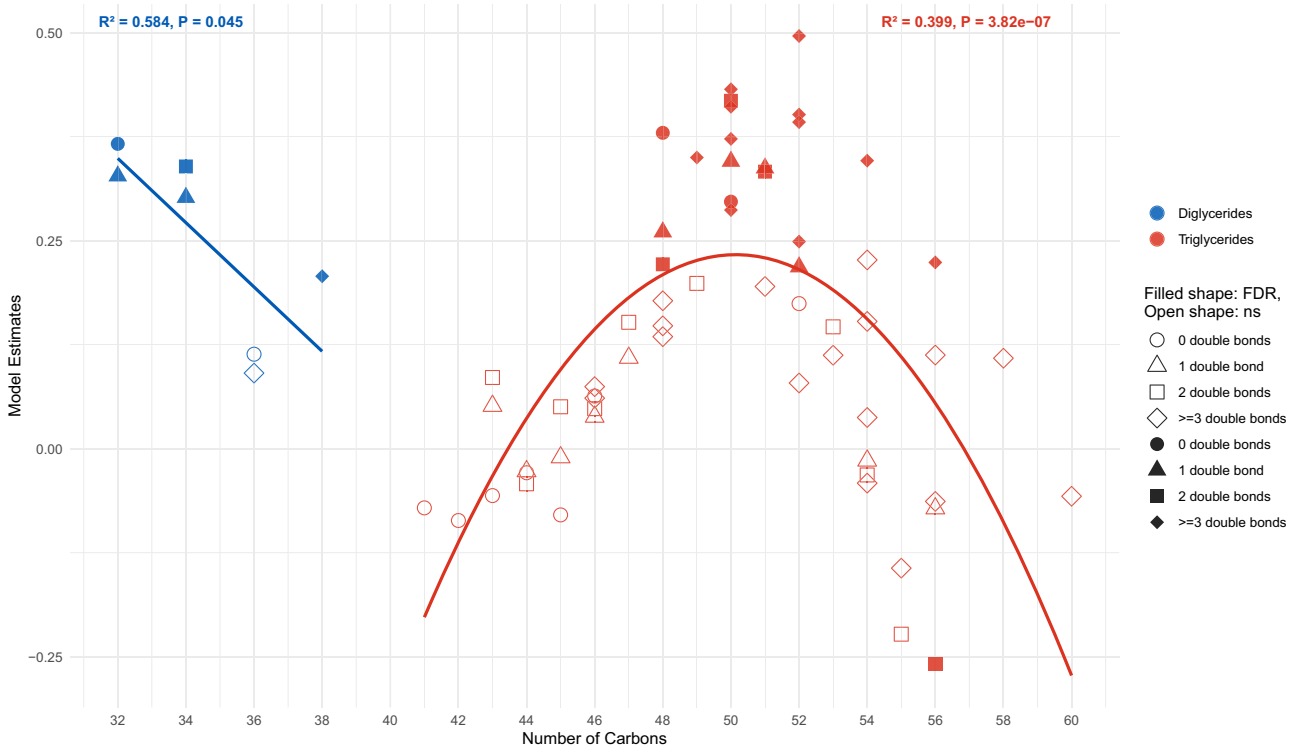


Fig. 2 | Associations between carbon chain length in triglycerides and diglycerides with metabolite level changes from low- to moderate- to high-carbohydrate diet. The scatterplot illustrates the association between carbon chain length in triglycerides and diglycerides and diet patterns, as assessed by the carb-to-fat ratio increase (i.e., from low- to moderate- to high-carbohydrate diets). Red symbols denote triglycerides, and blue symbols denote diglycerides. The symbol shapes reflect the number of double bonds in the molecules: circles for saturated, triangles for one double bond, squares for two double bonds, and diamonds for three or more double bonds. Filled shapes indicate metabolites that changed significantly with increasing carb-to-fat ratio diets, as determined by the False Discovery Rate (FDR) adjustment. The quadratic association for triglycerides is depicted with a red fitted curve (R^2 = 0.399, 95% CI 0.354 to 0.691 for the linear term and − 0.007 to −0.004 for the quadratic term, P = 3.82e-07), and a linear negative trend for diglycerides is shown in blue (R^2 = 0.584, 95% CI − 0.076 to − 0.001, P = 0.045). The horizontal axis represents the number of carbon atoms in the lipid molecules, and the vertical axis displays the beta coefficients, which are the model's estimates for each metabolite. FDR-significant triglycerides had carbon chain lengths between 48 and 56, with the majority clustered between 48 and 52. Statistical analyses were performed using linear and quadratic regression models. All reported *P*-values are two-sided. Source data are provided as a Source Data file. *FDR* False Discovery Rate; *ns* non-significant.

External validation of CFR-associated metabolites

To validate associations with CFR in an external study, we utilized data from the Popular Diets Study, a crossover feeding trial, in which metabolites were measured using an earlier iteration of the Broad Metabolomics Platform[16]. All analyses using data from the Popular Diets Study are secondary outcomes. The study design was similar to the (FS)2, including participants with overweight or obesity who achieved 10–15% weight loss. In the Popular Diets Study, each participant then consumed three test diets in random order for 4-week periods: a low-fat diet with 60% carbohydrate, 20% fat, and 20% protein; a low-glycemic index diet with 40% carbohydrate, 40% fat, and 20% protein; and a very-low carbohydrate diet with 10% carbohydrate, 60% fat, and 30% protein)[17]. Metabolites were measured in fasting plasma samples at the beginning and end of each dietary intervention. After applying the same analytical approaches (including the same QC pipeline and covariates), we tested for the association of metabolite levels with increasing CFR in the three diets. Of the 289 known metabolites that passed QC in the Popular Diets Study, 186 were also present in the (FS)2 dataset and were included in subsequent analyses. A strong correlation was observed among the 186 metabolite level changes with a higher CFR in the (FS)2 and the Popular Diets Study (rho = 0.618; P <2.2e-16; Fig. 3). Furthermore, among these 186 metabolites, a significant association (FDR < 0.05) was observed in (FS)2 for 54 metabolites (Supplementary Data 2). Of these 54 metabolites, 46 (85.2%) showed directional concordance of association with CFR in the Popular Diets Study. Even more strikingly, of the 38 metabolites that had FDR-significant associations in both studies, 37 (97.4%) showed the same direction of association (Supplementary Table 1), providing strong validation of the associations observed in (FS)2.

Duration-dependent effects of CFR on metabolite levels

To explore potential adaptation of the metabolome to dietary CFR, we assessed associations of changes in metabolites across low-, moderate-, and high-CFR diets at two different timepoints in the (FS)2 (Supplementary Fig. 1). We compared changes in the early test phase (calculated as the metabolite level measured at the study midpoint [MID] minus START) with changes during the entire test phase (END minus START). Nearly all (218 out of 222) of the metabolites that showed significant associations with dietary pattern in the early phase, the entire study, or both had the same direction of association with CFR, with a strong correlation (rho = 0.872; P < 2.2e-16; Fig. 4a and Supplementary Data 3). Of these 222 FDR metabolites, 136 (61.3%) were significant for both the early and entire test phase, 74 (33.3%) were significant only for the early test phase, and 12 (5.41%) were significant only for the entire test phase (Fig. 4b). The metabolites that were associated exclusively during the early test phase, and hence potentially indicating metabolic adaptation, included acylcarnitines, fatty acids (both showing decreased levels with higher CFR), and triglycerides (showing increased levels with higher CFR). The metabolites that were associated both during the early and entire test phase primarily included amino acids, phosphatidylcholines, plasmanyls/plasmalogens, phosphatidylethanolamines plasmanyls/plasmalogens, and

https://doi.org/10.1038/s41467-026-68353-z

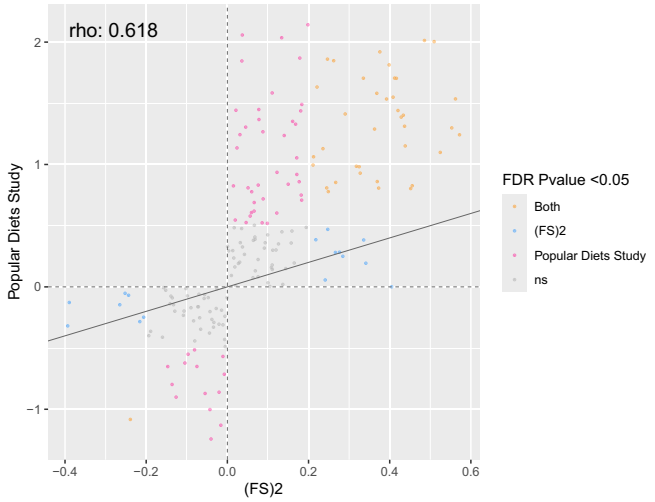


Fig. 3 | Comparative analysis of metabolite level changes with higher CFR in (FS)2 and popular diets study (validation study). The signs of the *X*- and *Y*-values have been set to both be positive if the directions of effect agree and both negative if they are discordant. As a result, points on the scatterplot located in the upper right quadrant indicate consistent effect directions, while points in the lower left quadrant represent inconsistent effect directions. The more consistent the data, the more the plot will resemble a fish along the $y = x$ line, with a smaller tail in the lower left quadrant and a larger head and body in the upper right quadrant ("fish" plot). See Fig. 4a for an example with even more highly correlated data. The *x*- and *y*-axes represent values derived from the regression t-statistics for each metabolite, adjusted by dividing by the square root of the number of participants in the respective analysis ((FS)2: $n = 147$; Popular Diets Study: $n = 21$). Metabolites on the plot are color-coded based on their FDR significance. Blue ("(FS)2") represents FDR-significant metabolite changes only in (FS)2. Pink ("Popular Diets Study") represents FDR-significant metabolite changes only in the Popular Diets Study (validation study). Yellow ("Both") represents FDR-significant metabolite changes in both studies, and gray ("ns") indicates no FDR-significant metabolite changes in both studies. The R coefficient indicates Spearman's correlation (rho = 0.618; $P < 2.2e\text{-}16$, two-sided). Source data are provided as a Source Data file. (FS)2, Framingham State Food Study; *FDR* False Discovery Rate; *ns* non-significant.

sphingomyelins (showing decreased levels with higher CFR), lysophosphatidylcholines, lysophosphatidylethanolamines, and triglycerides (showing increased levels with higher CFR).

Classification of diet groups and discriminative metabolites

Given the large number of metabolites associated with dietary CFR, we assessed how well models based on FDR-significant metabolites could separate individuals by diet type, employing a Partial Least Squares Discriminant Analysis (PLS-DA). We first focused on the two most extreme diets (i.e., low-CFR and high-CFR). Cross-validation of the PLS-DA model indicated that two components minimized the classification error rate, explained the largest proportion of variance, and were retained in the final model. The model achieved a prediction accuracy of 0.849 (95% CI 0.762 to 0.913) with sensitivity and specificity of 0.867 and 0.833, respectively (Supplementary Table 2). Not surprisingly, similar but less accurate results were observed when attempting to discriminate between all three diets, with model accuracy of 0.667 (95% CI 0.584 to 0.742).

Metabolites contributing to the discriminative power between the diet groups could serve as biomarkers of dietary CFR and potentially adherence. The strongest contributing metabolites were identified based on the Variable Importance in the Project (VIP) score value of the PLS-DA model (see methods for further details), with a threshold set at 1 (Supplementary Data 4). The metabolites with VIP > 1 ($n = 61$) primarily belonged to the subclasses of triglycerides, phosphatidylcholines, plasmanyls/plasmalogens, amino acids and peptides, phosphatidylethanolamines, and sphingomyelins. By applying stricter criteria (VIP > 1.5), we identified 2-aminooctanoic acid and PC P-36:0/PC O-36:1 as potential biomarkers for the extreme diet groups, and 2-aminooctanoic acid, PC P-36:0/PC O-36:1, and SM 18:1;O2/20:0 for biomarkers that could best distinguish all three diet groups.

Effect of weight-loss maintenance on metabolite profiles

To test whether individuals who did not maintain stable weights during the test phase could have altered metabolite profiles, we first identified participants who maintained their weight within ± 2 kg during the test phase, serving as a possible surrogate marker of adherence according to the study protocol. No statistical differences in characteristics were observed between groups (Supplementary Table 3). Notably, participants who gained >2 kg showed the lowest accuracy in the multi-metabolite PLS-DA classification model; by contrast, the model accurately predicted diet type for all participants who lost >2 kg for both diet types (Fig. 5), suggesting that excessive weight gain but not excessive weight loss during the test phase might be a surrogate for lower adherence. Similar findings were observed when incorporating all three diets into the PLS-DA models (Supplementary Fig. 5). To further explore our observation, we repeated the PLS-DA analyses using the same training datasets but excluding participants who gained > 2 kg from the testing set. This resulted in an increased accuracy of the model when examining both the two most extreme diets (0.880, 95% CI 0.796 to 0.939) and all three diets (0.686, 95% CI 0.601 to 0.763, Supplementary Table 2). When comparing only participants maintaining weight loss or losing > 2 kg vs. those gaining > 2 kg during the test phase, accurate identification of the type of diet was higher for those losing weight in the low-CFR group ($P = 0.012$ comparing the 2 extreme diets and $P = 0.004$ when all 3 diets were included in the PLS-DA models).

We repeated association analyses limited to the 137 individuals who were potentially more adherent to the diet (i.e., excluding those who gained weight >2 kg during the test phase, Supplementary Data 5, Supplementary Fig. 6, and Supplementary Data 6). Effects on metabolite levels were strongly correlated with effects seen in the entire sample (rho = 0.989, $P < 2.2e\text{-}16$, Supplementary Fig. 7). Notably, effect sizes of associations for metabolites that were FDR-significant in the total population were larger in this subsample (mean absolute values (standard deviation, SD) of t-statistics weighted by sample size: 0.346 (0.110) vs. 0.374 (0.124) in the subsample, $P < 1.90e\text{-}20$).

Furthermore, we conducted a logistic regression analysis to identify potential associations between metabolite level changes that were FDR significantly associated with higher CFR and weight gain (more than 2 kg), considering weight gain as a potential surrogate marker for lower dietary adherence in each diet. A number of nominally significant associations were observed for each diet (i.e., 26, 2, and 2 for low-, moderate-, and high-carbohydrate diets, respectively, Supplementary Data 7). However, no metabolite remained significant after applying FDR correction (Supplementary Data 7).

Sensitivity analyses of weight-loss maintenance thresholds

To evaluate the robustness of our results with different thresholds for weight change during the test phase, we repeated our analyses using alternative thresholds of ± 1.5 kg, ± 2.5 kg, and ± 3 kg. Our results were strongly associated with those using the ± 2 kg threshold (all rho >0.97), no statistically significant differences in changes in metabolite levels were observed using different thresholds, and all FDR-significant metabolites exhibited concordant directions of effect across all thresholds (Supplementary Data 8).

Metabolites and groups associated with CFR and T2D risk

We examined the directional trends of FDR-significant metabolites identified in the main analysis of the (FS)2 in relation to their

https://doi.org/10.1038/s41467-026-68353-z

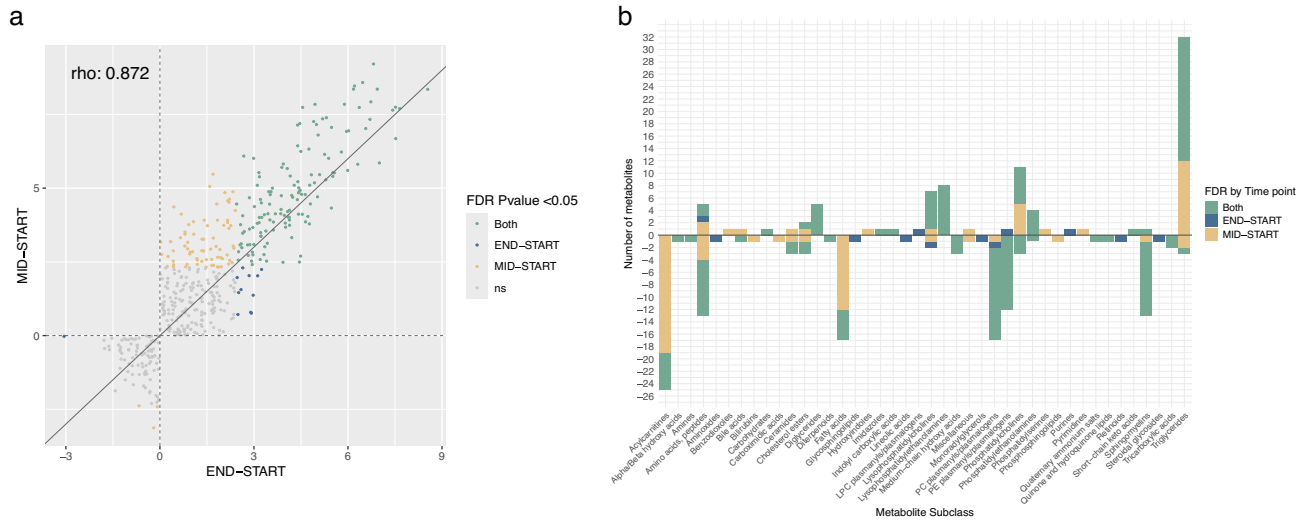


Fig. 4 | Metabolite level changes from low- to moderate- to high-carbohydrate diet: comparative results between early vs. entire test phase. a Metabolite changes: entire vs. early test phase. This "fish" plot (see Fig. 3 for explanation), illustrates the relationship between the metabolite changes for all participants ($n = 147$) across two distinct timeframes. The first analysis was conducted for the entire test phase (END-START), and the second analysis focused on metabolite changes during the early test phase (MID-START). The *x*- and *y*-axes represent values derived from *t*-statistics of the corresponding linear regression analyses. The R coefficient indicates Spearman's correlation (rho = 0.872; $P < 2.2e\text{-}16$, two-sided). Each point corresponds to a specific metabolite. Metabolites on the plot are color-coded based on their False Discovery Rate (FDR) P-value significance. Blue ("END-START") indicates FDR-significant metabolite changes only during the entire test phase, while yellow ("MID-START") represents FDR-significant metabolite changes solely during the early test phase. Green ("Both") represents FDR-significant metabolite changes in both periods (early and entire test phases), and gray ("ns") indicates the absence of FDR-significant metabolite changes in either period. Source data are provided as a Source Data file. **b** FDR significant metabolite changes: entire vs. early test phase. This bar plot illustrates the number of metabolites that exhibited FDR-significant changes from low- to moderate- to high-carbohydrate diets (higher carbohydrate-to-fat ratio, CFR) across different intervals of the test phase. Each bar represents a group of metabolites belonging to a specific subclass. The bars indicate the number of metabolites showing statistically significant changes, corrected for FDR using a cutoff of <0.05. Positive values on the y-axis correspond to the number of metabolites with increased metabolite level changes with higher CFR, while negative values correspond to the number of metabolites with decreased metabolite level changes with higher CFR. "END-START" represents FDR-significant metabolite changes only throughout the entire test phase. "MID-START" represents FDR-significant metabolite changes only during the early test phase. "Both" signifies FDR-significant metabolite changes in both periods (early and entire test phases). Source data are provided as a Source Data file. *CFR* carbohydrate-to-fat ratio; *END*,end of test phase; *FDR* False Discovery Rate; (FS)2, Framingham State Food Study; *LPC* lsophosphatidylcholines; *MID* middle of test phase; ns, non-significant; *PC* phosphatidylcholines; *PE* phosphatidylethanolamines; START, start of test phase.

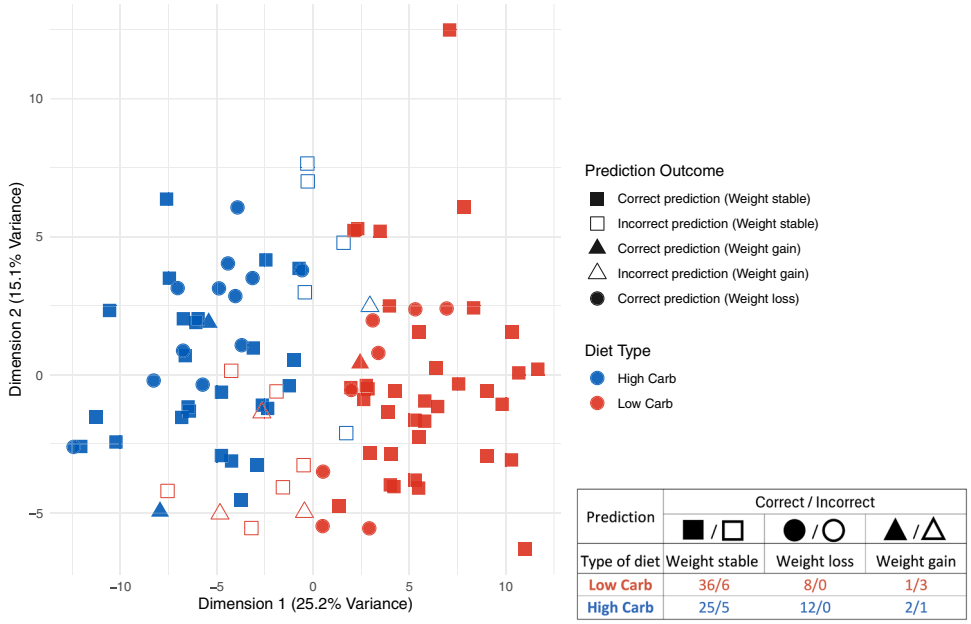
association with risk of incident T2D (Supplementary Table 4), as determined in a recent published meta-analysis, which considered metabolites associated with incident T2D in at least two independent prospective cohort studies[18]. Twenty metabolites showed increased levels associated with higher CFR and also with an elevated T2D risk (Fig. 6a). The majority of these metabolites belong to the triglycerides and diglycerides classes. One metabolite (LPC 18:2) showed decreased levels associated with higher CFR and decreased diabetes risk. In contrast, five metabolites displayed the opposite pattern (Fig. 6b), where metabolite levels either increased with higher CFR but were associated with a lower T2D risk (two metabolites), or where metabolite levels decreased with higher CFR but were associated with an increased T2D risk (three metabolites).

Finally, to identify groups of metabolites associated with higher CFR and T2D risk and could therefore potentially represent connections between dietary patterns and T2D, we conducted Metabolite Set Enrichment Analysis (MSEA) to evaluate enrichment scores in the (FS)2 main analysis and T2D meta-analysis across 18 metabolite classes (Fig. 7). A concordant pattern of increased levels with higher CFR and T2D risk was observed for triglycerides, diglycerides, phosphatidylcholines, phosphatidylethanolamines, and carboxylic acids and derivatives, with significant enrichment scores observed, particularly for triglycerides with ≤3 double bonds in both analyses. A concordant pattern of decreased levels with both higher CFR and T2D risk was observed for sphingomyelins and carnitines. Notably, not all groups of metabolites associated with both CFR and T2D risk had concordant direction of effects: an opposite trend between the two analyses was observed mainly for lysophosphatidylcholines and lysophosphatidylethanolamines (Fig. 7). Thus, if some of these associations with T2D represent causal mechanisms, the consequences of dietary CFR could include different metabolic effects, with some increasing and others decreasing future disease risk.

Metabolite and cardiometabolic biomarker mediation analyses

To assess whether diet-responsive metabolites from the (FS)2, validated in the Popular Diets Study, mediate changes in cardiometabolic biomarkers, we tested 160 metabolite–outcome pairs derived from 10 cardiometabolic biomarkers and 16 validated metabolites (after excluding highly correlated metabolites). After applying FDR correction, PC 38:3, SM 18:1;O2/14:0, and SM 18:1;O2/16:0 had significant evidence of mediating the associations with low-density lipoprotein (LDL) (ACME = 5.87, − 6.67, − 6.50; 95% CI 2.55 to 9.8, − 11.3 to − 2.86, − 10.8 to − 3.26), total cholesterol (ACME = 11.36, − 13.68, − 9.11; 95% CI 5.83 to 17.1, − 19.8 to − 8.34, − 14.6 to − 4.42), and non-HDL cholesterol (ACME = 9.64, − 9.03, − 7.56; 95% CI 5.13 to 14.7, − 14.8 to − 4.2, − 12.3 to − 3.51), respectively; all $P < 2.71e\text{-}15$. For high-density lipoprotein (HDL), SM 18:1;O2/14:0 and CE 18:2 showed significant evidence of negative mediation effects (ACME = − 4.56 and − 2.54; 95% CI − 6.72 to − 2.69 and − 4.44 to − 0.981; $P < 2.71e\text{-}15$ and 0.02, respectively). For lipoprotein insulin resistance index (LP-IR), four metabolites had significant evidence of mediating the dietary effect: PC 38:3 (ACME = 6.62, 95% CI 3.16 to 10.6, P = 2.7e-15), CE 18:2 (ACME = 10.32, 95% CI 6.27 to 15.1, P = 2.7e-15), gamma-Aminobutyric acid (GABA) (ACME = 6.18, 95% CI 2.58 to 10.5, P = 2.7e-15), and TG 54:8 (ACME = 3.22, 95% CI 0.869 to 6.55, $P = 0.02$). Finally, for systolic blood pressure, only TG 54:8 showed significant evidence of mediation (ACME = 2.12, 95% CI

https://doi.org/10.1038/s41467-026-68353-z



Prediction	Correct / Incorrect		
	■ / □	● / ○	▲ / △
Type of diet	Weight stable	Weight loss	Weight gain
Low Carb	36/6	8/0	1/3
High Carb	25/5	12/0	2/1

Fig. 5 | PLS-DA Score olot for diet-related metabolites: predictive outcomes and weight-loss maintenance annotations forlow and high carbohydrate diets. The PLS-DA score plot demonstrates the separation of participants by diet, based on metabolites exhibiting FDR-significant changes between the two diets. Scores are derived from the data of both dietary groups. Predictive outcomes result from a 5-fold cross-validation process, utilizing the Mahalanobis distance to measure model efficacy. Further participant annotations are based on weight-loss maintenance during the test phase of the study, which was used as a surrogate marker of adherence. "Weight stable" refers to participants who maintained their weight loss, while the "weight loss" and "weight gain" categories identify individuals who experienced a change of more than 2 kg. The accompanying table provides a summary of the correct and incorrect diet predictions for participants in each class. Source data are provided as a Source Data file. *FDR* False Discovery Rate; *PLS-DA* Partial Least Squares Discriminant Analysis.

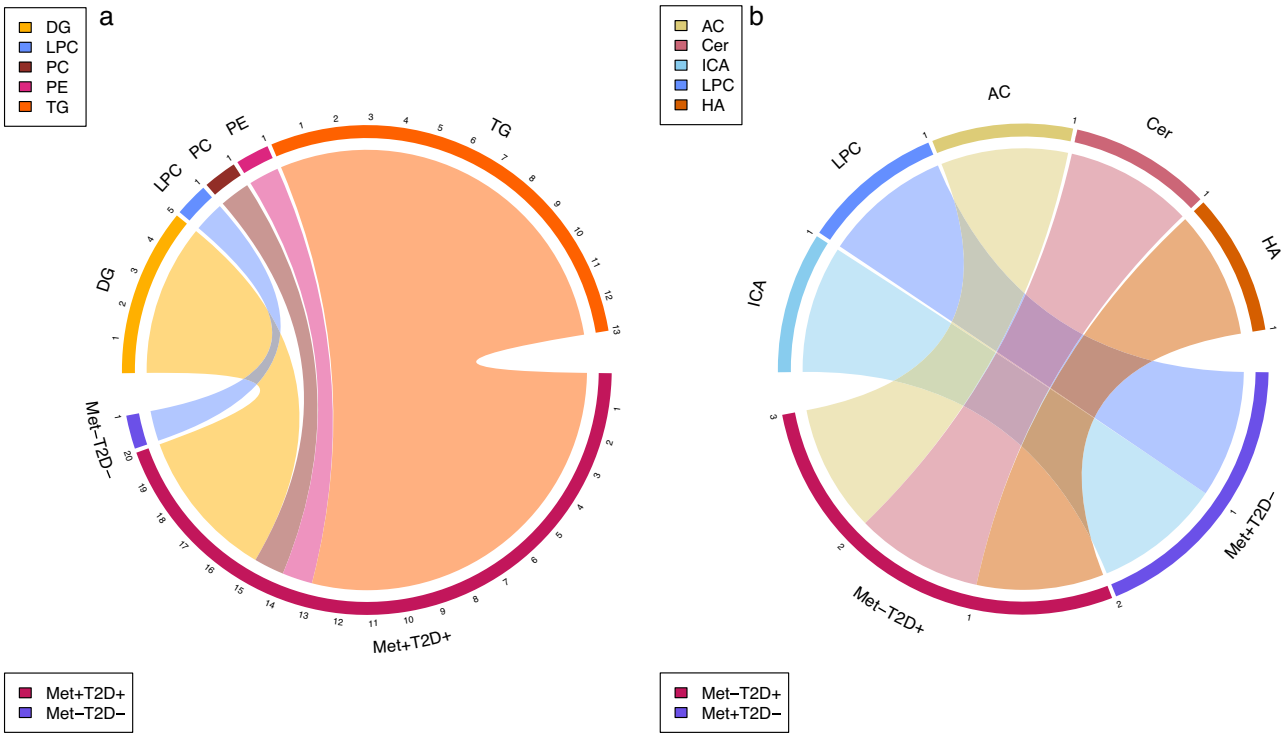


Fig. 6 | Concordant and discordant trends in metabolite distribution across (FS)2 and T2D risk. Chord diagram illustrating the number of metabolites grouped by classes, depicting (**a**) concordant or (**b**) discordant trends observed in both the (FS)2 (assessed from low- to moderate- to high-carbohydrate diet) and their association with T2D risk. The thickness of the sectors and ribbons is proportional to the number of metabolites. Only metabolites exhibiting FDR significance in both analyses are included. Source data are provided as a Source Data file.Met + : Metabolite levels increase across low to moderate to high carb diets in the (FS)2; Met-: Metabolite levels decrease across low to moderate to high carb diets in the (FS)2. T2D + : Associated with increased T2D risk; T2D-: Associated with decreased risk. *(FS)2* Framingham State Food Study; *AC* Acylcarnitines; *Cer* Ceramides; *DG* Diglycerides; *FDR* False Discovery Rate; *ICA* Indolyl carboxylic acids and derivatives, *LPC* Lysophosphatidylcholines, *PC* Phosphatidylcholines, *PE* Phosphatidylethanolamines, *T2D* Type 2 Diabetes; *TG* Triglycerides, and *HA* Alpha/Beta hydroxy acids.

https://doi.org/10.1038/s41467-026-68353-z

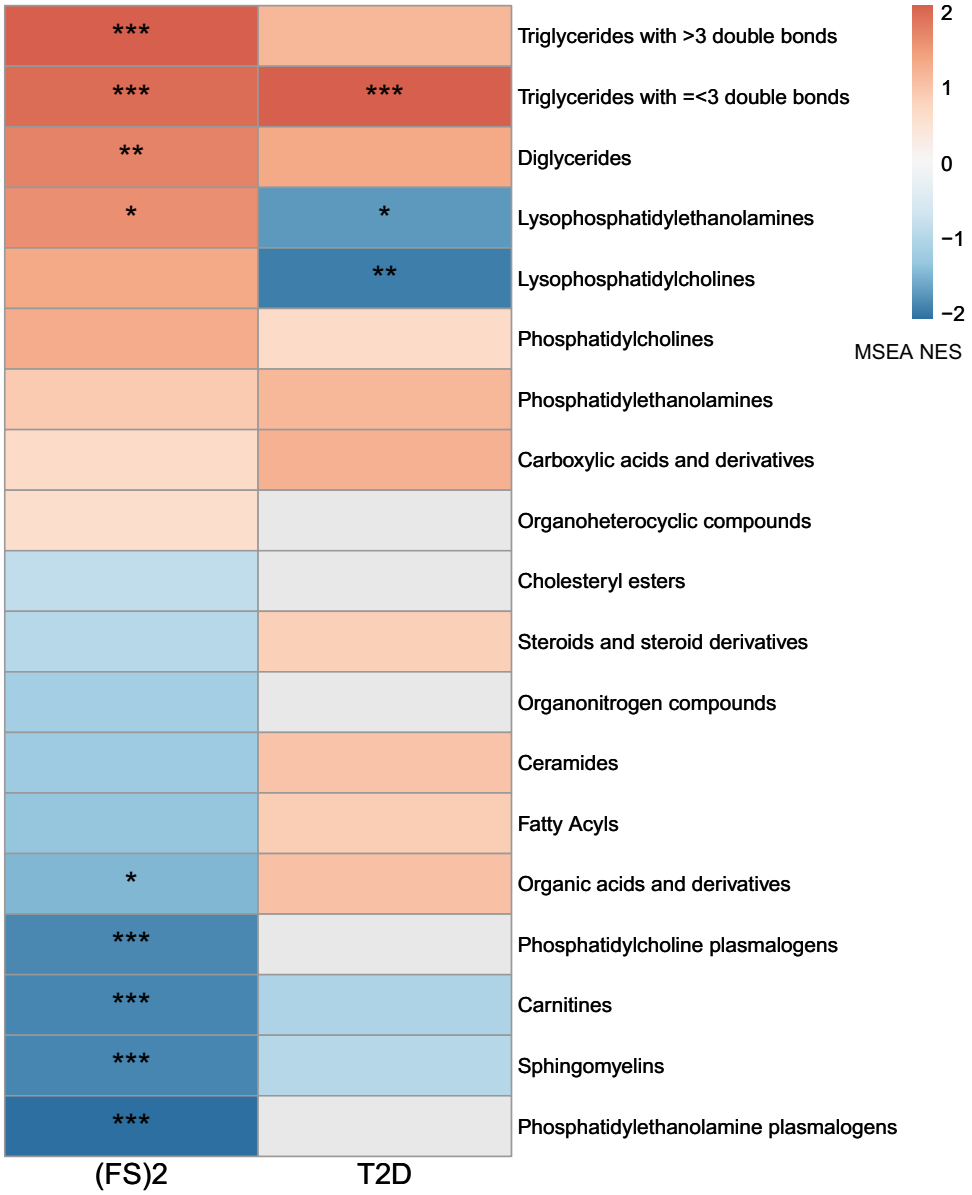


Fig. 7 | Metabolite set enrichment analysis heatmaps of (FS)2 and T2D risk. The heatmaps display the results of the Metabolite Set Enrichment Analysis (MSEA) for metabolites identified in the (FS)2 and in the Type 2 Diabetes (T2D) meta-analysis. Annotations on the heatmaps indicate the significance of the Normalized Enrichment Scores obtained from the MSEA. All statistical tests were two-sided, with the False Discovery Rate (FDR) correction applied to adjust for multiple comparisons. Source data are provided as a Source Data file. * FDR < 0.2; ** FDR < 0.05; *** FDR < 0.01. *(FS)2* Framingham State Food Study; *FDR* False Discovery Rate; *MSEA* Metabolite Set Enrichment Analysis, *NES* Normalized Enrichment Score; *T2D* Type 2 Diabetes.

0.409 to 4.32, $P = 0.02$). No significant mediation effects were observed for the remaining biomarkers [1,5-anhydroglucitol, diastolic blood pressure, lipoprotein(a), and high-sensitivity C-reactive protein] (Supplementary Data 9).

Discussion

In this study, our primary aim was to identify metabolite changes associated with dietary interventions differing in CFR. Leveraging metabolomic data from 147 participants in the FS(2), we identified significant changes in fasting levels of 148 of 479 measured metabolites caused by changes in dietary CFR. Specifically, higher CFR was significantly associated with a notable increase in triglycerides, diglycerides, lysophosphatidylcholines, and lysophosphatidylethanolamines. Conversely, a significant decrease in metabolite levels was observed in phosphatidylcholines plasmanyls/plasmalogens, phosphatidylethanolamines plasmanyls/plasmalogens, sphingomyelins, amino acids, acylcarnitines, and fatty acids. External validation with metabolomics data from the Popular Diets Study, a previous crossover feeding trial, confirmed that dietary CFR changes consistently altered fasting levels of many metabolite levels, and that these changes were not idiosyncratic to the FS(2). Among the validated metabolites, several lipid classes were consistently influenced by dietary CFR changes, including ceramides, sphingomyelins, diglycerides, and triglycerides, highlighting a robust impact on lipid metabolism. Acylcarnitines, particularly short- and medium-chain species, exhibited consistent patterns, underscoring the role of dietary macronutrient composition in energy metabolism[8,21]. In addition, dietary-associated changes were evident for amino acids, including serine, aminoisobutyric acid[22], and gamma-aminobutyric acid (GABA), supporting potential links between dietary patterns and amino acid metabolism. Notably, GABA, a primary inhibitory neurotransmitter in the central nervous system, has been implicated in the regulation of eating behavior, inflammatory

pathways, and pancreatic islet function, with potential effects on glucose metabolism[23–25]. Moreover, indole-3-propionic acid[26,27], a microbial tryptophan metabolite with potential anti-inflammatory and protective effects in metabolic diseases, also exhibited consistent dietary responsiveness. Results from our study are also consistent with changes observed in participants following a very-low-carbohydrate high-fat (VLCHF) diet, wherein triglycerides and glycerophospholipids (including phosphatidylcholines, lysophosphatidylcholines, and phosphatidylethanolamines) levels decreased, while acylcarnitines, plasmalogens, sphingomyelin, ceramides, and fatty acids increased[28].

In our study, we noticed that participants who gained more than 2 kg during the intervention had the lowest accuracy in prediction of diet type using models based on multiple metabolites. Consistent with this observation, diet-related changes were greater in those who maintained weight or lost more than 2 kg. In addition, discrimination of diets using multiple metabolites was sharper in participants who did not gain more than 2 kg, suggesting that metabolite measurements may have the potential to reflect adherence to a prescribed diet.

By exploring the changes in metabolite levels across timepoints of the test phase, we observed that the majority of the metabolite alterations exhibited a sustained response, evident from the early phase of the dietary intervention. However, some metabolites, primarily within the acylcarnitines, triglycerides, and fatty acids groups, showed a significant change during the early phase of intervention, but these changes were not maintained after 10 weeks. In the acylcarnitine group, most of the changes in metabolite levels occurred during the early phase of the dietary intervention, with only 6 out of 25 acylcarnitines remaining significant throughout the entire study, suggesting a potential adaptation. A similar pattern was observed for fatty acids, with 5 out of 17 remaining significant throughout the entire study. These early responses highlight the complexity of metabolic adaptation to dietary modifications and underscore the necessity for long-term monitoring to fully understand the metabolic shifts associated with changes in dietary carbohydrate and fat. Of note, the acylcarnitines that remained significant throughout the study were mostly short- and medium-chain acylcarnitines. Short-chain acylcarnitines are considered the most abundant group, and both short- and medium-chain acylcarnitines have been investigated for their roles in various diseases and pathologies, including cardiometabolic disorders[29–31].

Lipids and lipid-like molecules constituted the main superclass among the potential biomarkers identified, revealing specific and distinct patterns for particular lipid subclasses. We observed a more pronounced effect of higher CFR on triglyceride levels. A significant positive association for triglycerides was identified, evident from the early test phase. Notably, the majority of triglycerides (18 out of 21 that were FDR-significant) had an even number of carbon atoms. Furthermore, our analysis unveiled a parabolic relationship between triglyceride composition and dietary patterns. The carbon chain lengths of the FDR-significant triglycerides ranged from 48 to 56, with the peak of this curve noted for triglycerides with carbon chain lengths ranging from 48 to 52 (18 out of 21 FDR-significant triglycerides). It is noteworthy that triglycerides within this same range of total carbon atoms were linked to an increased risk of T2D[32]. In addition, we explored the role of the number of double bonds in triglycerides with CFR and T2D risk by performing MSEA using data from (FS)2 and a meta-analysis of published associations of metabolites with incident T2D[18]. We observed a significant association of triglyceride class with higher CFR and T2D risk, more prominently for triglycerides with lower saturation (i.e., ≤ 3 double bonds, Fig. 7). In alignment with our findings, previous studies have highlighted the pattern of association of triglycerides with a similarly low number of carbon atoms and double bond content with increased T2D risk[32,33]. Recent studies have also underscored the distinct physiological behaviors between small (≤48 carbons) and large (≥49 carbons) triglycerides, impacting energy metabolism, lipidemic and inflammatory patterns[34], as well as the association of triglycerides with ≤ 56 carbons and ≤3 double bonds with increased risk of all-cause mortality[35]. Our observations suggest an influence of CFR on triglycerides with fatty acid compositions potentially associated with T2D and other related health outcomes, highlighting the necessity for further research to elucidate the potential underlying functional roles within the triglyceride spectrum. As triglycerides demonstrated the highest technical reproducibility in a previous study, they could be further explored as potential biomarker candidates at the species and subclass levels[34].

In our study, we observed distinct behaviors between ester-linked and ether-linked glycerophospholipids. Specifically, we found that phosphatidylethanolamines plasmanyls/plasmalogens (PE-P/PE-O) and phosphatidylcholines plasmanyls/plasmalogens (PC-P/PC-O) levels decreased significantly with higher CFR. These changes were significant even from the early test phase (i.e., 10 weeks) and persisted through 20 weeks. In contrast, we mainly observed an increase in phosphatidylcholine and phosphatidylethanolamine levels with higher CFR. Previous lipidomic studies have also reported distinct behaviors between ester-linked and ether-linked phosphatidylethanolamines. A positive association of phosphatidylcholine and phosphatidylethanolamine levels with insulin resistance and metabolic disorders like obesity, prediabetes, and T2D has been observed, while the inverse association has been noted for ether-phosphatidylcholine and ether-phosphatidylethanolamines[36–38]. In addition, phosphatidylethanolamines plasmanyls/plasmalogens may exhibit antioxidant properties[39], alleviate inflammation, and are overall linked to healthy phenotypes[34]. A study involving monozygotic twins showed that obesity, independent of genetic influences, was associated with decreases in ether phospholipids and increases in lysophosphatidylcholines, indicating alterations in the lipid profile involved in pathways linking obesity with insulin resistance, inflammation, and atherogenesis[40]. However, further research is needed to fully understand the distinct pathophysiological roles of different lipid groups in relation to insulin sensitivity and metabolic disorders[41], as well as the impact of diet on lipid composition and the resulting biological responses.

In our study, besides triglycerides, acylcarnitines, ester/ether-linked phosphatidylethanolamines, and ester/ether-linked phosphatidylcholines, we observed significant changes in additional metabolite subclasses such as lysophosphatidylcholines, lysophosphatidylethanolamines, and sphingomyelins. All the above metabolite groups have been studied for their involvement in insulin resistance and related metabolic pathways and disorders[29,40,42,43], with diet potentially playing a crucial role in determining circulating lipid profiles and levels[29,42,44]. A low-carbohydrate diet may offer potential metabolic benefits – such as enhancing insulin sensitivity, attenuating postprandial blood glucose spikes, improving lipid profiles, and reducing body fat – and represents an important dietary strategy for managing T2D[45], while also linked to higher rates of possible T2D remission[46]. In the present study, we identified changes between diets with distinct CFR and specific metabolite subgroups associated with metabolic disorders such as insulin resistance and T2D. These findings may partially explain the underlying mechanisms behind the beneficial effects of low-carbohydrate diets. Our mediation analysis further suggests that specific diet-responsive metabolites, particularly lipids, may partially mediate the relationship between dietary CFR and cardiometabolic biomarkers (such as LDL cholesterol), offering insight into potential biological pathways influenced by dietary composition. However, further studies are required to assess longer-term health outcomes and fully elucidate the effects of low-carbohydrate diets on the metabolome. Additional research should investigate how changes in dietary patterns, including the type and amount of lipid intake, as well as interventions targeting specific lipid pathways, can influence plasma lipid profiles and clinical outcomes like insulin resistance and T2D. In addition, it is crucial to explore which specific

diet-associated metabolite changes might be linked to altered cardiometabolic risk, as some metabolites are associated with increased risk and others with decreased risk.

Our study has several strengths. To our knowledge, (FS)2 is the longest RCT of diets varying in CFR, with comprehensive metabolomic profiling. Furthermore, the dietary interventions were designed to differ in CFR without imposing extreme restrictions on either carbohydrate or fat, potentially facilitating translation. The duration of the test phase was sufficient to eliminate confounding by transient metabolic adaptations to dietary changes after randomization[47]. Moreover, we conducted additional analyses to identify metabolite changes in the early test phase that could reflect transient metabolic adaptations. Also, dietary adherence was enhanced by implementing a feeding protocol rather than solely relying on nutrition education and dietary counseling, ensuring differentiation among intervention groups consistent with specified macronutrient targets. In addition, after implementing advanced discriminant analysis techniques, we observed that participants who gained more than 2 kg during the intervention had the lowest correct diet discrimination and smaller average changes in metabolite levels, suggesting that this magnitude of weight change is a potential marker of lower adherence. Finally, comparison of the results from the present study with those from the Popular Diets Study, a previous crossover feeding trial with similar test diets[16] used for external validation, indicated consistency and thereby providing some assurance regarding the reproducibility of the metabolomic profiling, despite differences in intervention duration between the two studies.

Our study has some limitations. Participants in the (FS)2 were middle-aged adults who were generally healthy, albeit with overweight or obesity. Although there was some ethnic-racial diversity, the majority of participants were non-Hispanic white, which limits the generalizability of our findings and requires cautious interpretation for other demographic groups. Moreover, the need to generalize these findings to free-living cohorts is crucial, as the context of a controlled feeding trial does not reflect real-world conditions. The results also may not be generalizable to individuals consuming more restrictive diets (e.g., very-low-carbohydrate or very-low-fat diets with < 10% of total energy derived from these macronutrients). Even within the specified targets for carbohydrate and fat, diets may vary in factors such as glycemic index of carbohydrate sources, amounts and types of fiber, fatty acid profiles, micronutrient content, and availability of bioactive compounds. Therefore, diets with similar CFR may not necessarily lead to the same changes in metabolomic profiles, and our conclusions may not apply to all diets with similar macronutrient content, depending on the food source. Nevertheless, we aimed to employ healthful, palatable, and pragmatic representations of each diet type, with relevance to clinical translation. Inconsistent dietary adherence could be a potential limitation of the study, but weight gain served as a surrogate marker, with sensitivity analyses confirming strongly consistent results. While our study cannot provide evidence of causality for the effect of diet-associated metabolites on cardiometabolic outcomes, it generates hypotheses about potential links between diet-induced metabolic changes and health outcomes, which can guide future research to explore these associations. The present study highlights the need for further research, including comprehensive analyses of these diet-responsive metabolites in large prospective longitudinal cohorts with clinically relevant outcomes, as well as the application of genetic causal inference methods (such as Mendelian randomization) to investigate potential causal relationships and elucidate the clinical relevance and biological roles of these metabolites. Finally, metabolomic platforms used in this study, despite employing a comprehensive approach to metabolomics, identified only a subset of circulating metabolites, did not always determine the exact molecule identity and structure, and did not include analyses of completely unknown signals in our data.

In summary, we measured a diverse set of lipids and various other plasma metabolites to identify the effects of higher CFR diets (i.e., from low to moderate to high carbohydrate) on plasma metabolites and potential diet-responsive biomarkers. We identified profound, diet-induced changes in the metabolome that were largely reproducible in an independent study. Distinct patterns were observed in specific groups of metabolites, and several of these diet-responsive metabolites have associations with T2D risk, suggesting that diet-induced lipid and metabolite changes may provide insights into the potential effects of different dietary macronutrient compositions on long-term metabolic disease risk. Therefore, the present work provides an important basis for future research assessing potential causal relationships between these diet-responsive metabolites and clinical variables, including both cardiometabolic risk factors and outcomes. Understanding how diet influences cardiometabolic risk factors and related health outcomes is crucial for assessing disease development and progression. Utilizing metabolomic methodologies, we identified diet-related metabolites, providing markers of dietary patterns and potentially valuable insights into pathophysiological mechanisms linked to these patterns. These findings could enhance our understanding of the role of diet in cardiometabolic conditions such as T2D, especially among individuals with overweight and obesity, thereby facilitating targeted dietary treatment approaches, improving preventive strategies and outcomes, and mitigating disease burdens.

Methods

Study design and participants

The study protocol, including metabolomics analyses of all cohorts, was approved by the Institutional Review Board of Boston Children's Hospital and conducted in accordance with the Declaration of Helsinki and all relevant ethical regulations. Participants in both trials, the (FS)2 trial and the Popular Diets Study, provided written informed consent. Each consent form indicated that de-identified biospecimens would be saved for future analyses. Patients or members of the public were not involved in conceptualizing the research question, specifying outcome measures, developing plans for design and implementation of study protocols, or interpreting and reporting results.

The Framingham State Food Study, known as (FS)2, was a randomized controlled parallel trial of dietary interventions (ClinicalTrials.gov: NCT02068885)[19]. The inclusion and exclusion criteria, along with the study protocol, and the primary outcomes have been previously published[8,19]. Briefly, the study enrolled participants aged 18 to 65 years with a body mass index (BMI) of ≥ 25 kg/m^2. Exclusion criteria included physician-diagnosed major medical illness or abnormal laboratory screening results (hemoglobin A1c, thyroid-stimulating hormone, hematocrit [<30%], blood urea nitrogen, creatinine, and alanine transaminase [>200% of the normal upper limit]).

The trial comprised run-in and test phases, as illustrated in Supplementary Fig. 1 and previously described in detail[19]. According to feeding protocols during both phases, all foods were provided to participants to promote adherence with specified dietary targets for carbohydrate, fat, and protein. The run-in phase involved pre-weight-loss assessments (PRE), reducing energy intake to 60% of estimated weight-maintenance needs (to achieve weight loss of 12 ± 2% relative to PRE body weight), stabilizing body weight after weight loss, and conducting assessments again at random assignment (START). Participants who completed the run-in phase were randomized into three diet arms (with an allocation ratio of 1:1:1) using a blocked randomization design to ensure balance throughout the study. Randomization was stratified by sex, ethnicity/race (non-Hispanic White, other), age (18–39.9, 40.0–65.9 years), BMI category (overweight, obesity), and feeding site. Within each stratum, diet assignments were randomly generated and permuted within blocks of 3, 6, or 9, and the blocks also were randomly permuted to ensure unpredictability and prevent any bias. The diet assignment lists were stored in a secure electronic folder

accessible only to the data and quality manager responsible for randomization. Blinding was maintained throughout the study for all staff involved in outcome assessment and laboratory analyses. Staff responsible for randomization were separate from those conducting data processing and analyses.

The test phase focused on comparing the effects of low-, moderate-, and high-carbohydrate diets—differing in CFR and controlled for protein—during weight-loss maintenance. According to the study protocol, energy intake was adjusted to maintain body weight within ± 2 kg of the stabilized body weight achieved after weight loss, with weight-loss maintenance serving as a surrogate marker of adherence. Assessments were repeated at 10 (MID) and 20 (END) weeks of the test phase.

The percentages of total energy (carbohydrate-fat-protein) were 60-20-20 for the high-carbohydrate, 40-40-20 for the moderate-carbohydrate, and 20-60-20 for the low-carbohydrate diets. We used the same foods across diets, albeit in differing amounts, and replaced foods when necessary to reach designated macronutrient targets. Targets for the high-carbohydrate diet were reached using a variety of carbohydrate-containing foods such as whole grains, vegetables, fruits, legumes, and low-fat dairy products. Targets for the moderate-carbohydrate diet, compared to the high-carbohydrate diet, were reached by decreasing the quantity of grains and fruits, adding foods containing fat (e.g., nuts, seeds, sauces, spreads, toppings), decreasing amounts of legumes when necessary, and including some higher-fat dairy products. Targets for the low-carbohydrate diet, compared to the moderate-carbohydrate diet, were reached by eliminating all grains, removing some fruits, adding more foods containing fat, further decreasing amounts of legumes when necessary, and increasing some higher-fat dairy products. The quantities of non-starchy vegetables were similar across diets.

Data collection took place at Framingham State University and Assabet Valley Regional Technical High School in Massachusetts. Recruitment of participants for three cohorts occurred each spring from 2014 to 2016, during the semester preceding the academic year of participation. Study implementation, including data collection and controlled feeding, began each fall and continued through the spring. Body composition was assessed using dual-energy X-ray absorptiometry (Horizon A, Hologic Inc., Bedford, MA) equipped with the Hologic APEX 5™ operating system. Data collected for the original trial were used in the exploratory analyses for the present study. The sex of participants was determined based on self-report. Additional metabolomic analyses were subsequently performed using the Metabolomics Platform at the Broad Institute to achieve the present aims.

Outcomes

All analyses described in this manuscript are secondary outcomes from metabolomic analyses of samples collected during the main trial. The main secondary outcome was to evaluate metabolomic responses to diets differing in CFR, assessed by untargeted LC–MS profiling, including identification of diet-responsive biomarkers. Additional secondary outcomes included: assessing the utility of diet-responsive biomarkers for monitoring dietary adherence; evaluating the consistency of diet-induced metabolite signatures throughout the intervention to assess potential metabolic adaptation; and examining the association of CFR-linked metabolites to T2D risk using data from a published meta-analysis.

Metabolite profiling, data preprocessing and missing values

Metabolomic measurements were made using four complementary LC-MS methods that provided a hybrid analysis of metabolites of known identity and signals from yet to be identified compounds that were traced by measured mass-to-charge ratios (m/z) and retention times (RT). Plasma metabolites were profiled for samples collected at PRE, START, MID, and END following an overnight fast (>8 hours). The metabolomic measurements were conducted by the Metabolomics Platform at the Broad Institute, blinded to the intervention type. For signals that match more than one chemical isomer indistinguishable by the LC-MS method, potential compound names are presented, separated by either "or" or "/".

Hydrophilic liquid chromatography analyses of water-soluble metabolites in the positive ionization mode (HILIC-pos) were conducted using an LC-MS system composed of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp) coupled to a Q Exactive hybrid quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific). Plasma samples (10 µL) were prepared via protein precipitation with the addition of 9 volumes of 74.9:24.9:0.2 v/v/v acetonitrile/methanol/formic acid containing stable isotope-labeled internal standards (valine-d_8, Sigma-Aldrich; and phenylalanine-d_8, Cambridge Isotope Laboratories). The samples were centrifuged (10 min, 9000 × *g*, 4 °C), and the supernatants were injected directly onto a 150 × 2 mm, 3 µm Atlantis HILIC column (Waters). The column was eluted isocratically at a flow rate of 250 µL/min with 5% mobile phase A (10 mmol/L ammonium formate and 0.1% formic acid in water) for 0.5 min followed by a linear gradient to 40% mobile phase B (acetonitrile with 0.1% formic acid) over 10 min. Mass spectroscopic (MS) analyses were performed by using electrospray ionization in the positive ion mode using full scan analysis over 70 to 800 *m/z* at 70 000 resolution and 3 Hz data acquisition rate. Other MS settings were as follows: sheath gas 40, sweep gas 2, spray voltage 3.5 kV, capillary temperature 350 °C, S-lens RF 40, heater temperature 300 °C, microscans 1, automatic gain control target 1e6, and maximum ion time 250 ms.

HILIC analyses of water soluble metabolites in the negative ionization mode (HILIC-neg) were conducted using an Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.; Marlborough, MA) coupled to a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific; Waltham, MA). Metabolites were extracted from plasma (30 µL) using 120 µL of 80% methanol containing inosine-15N4, thymine-d4 and glycocholate-d4 internal standards (Cambridge Isotope Laboratories; Andover, MA). The samples were centrifuged (10 min, 9000 x *g*, 4 °C), and the supernatants were injected directly onto a 150 × 2.0 mm Luna NH2 column (Phenomenex; Torrance, CA). The column was eluted at a flow rate of 400 µL/min with initial conditions of 10% mobile phase A (20 mM ammonium acetate and 20 mM ammonium hydroxide in water) and 90% mobile phase B (10 mM ammonium hydroxide in 75:25 v/v acetonitrile/methanol) followed by a 10 min linear gradient to 100% mobile phase A. MS analyses were carried out using electrospray ionization in the negative ion mode using full scan analysis over m/z 70–750 at 70,000 resolution and 3 Hz data acquisition rate. Additional MS settings were: ion spray voltage, − 3.0 kV; capillary temperature, 350 °C; probe heater temperature, 325 °C; sheath gas, 55; auxiliary gas, 10; and S-lens RF level 50.

Reversed-phase C18 chromatography, negative ion mode MS profiling of free fatty acids, bile acids, and metabolites of intermediate polarity (C18-neg) was performed using a Nexera X2 U-HPLC (Shimadzu Corp.; Marlborough, MA) coupled to a Q Exactive hybrid quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific; Waltham, MA). Plasma samples (30 µL) were extracted using methanol containing 15R-15-methyl-PGA_2, 15R-15-methyl-PGF_{2alpha}, 15S-15-methyl-PGD_2, 15S-15-methyl-PGE_1, and 15S-15-methyl-PGE_2 (Cayman Chemical Co.) internal standards. Supernatants were injected onto a 150 × 2.1 mm ACQUITY T3 column (Waters; Milford, MA). The column was eluted isocratically at a flow rate of 450 µL/min with 20% mobile phase A (0.01% formic acid in water) for 3 min followed by a linear gradient to 100% mobile phase B (0.01% acetic acid in acetonitrile) over 12 min. MS analyses were carried out using electrospray ionization in the negative ion mode using full scan analysis over m/z 70–850 at 70,000 resolution and 3 Hz data acquisition rate. Additional MS settings were: ion spray voltage, − 3.5 kV; capillary temperature, 320 °C;

probe heater temperature, 300 °C; sheath gas, 45; auxiliary gas, 10; and S-lens RF level 60.

Positive ion mode analyses of polar and nonpolar plasma lipids using C8-reveresed phase chromatography (C8-pos) were conducted using an LC-MS system composed of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp) coupled to an Exactive Plus orbitrap mass spectrometer (Thermo Fisher Scientific). Plasma samples (10 µL) were extracted for lipid analyses using 190 µL of isopropanol containing 1,2-didodecanoyl-*sn*-glycero-3-phosphocholine (Avanti Polar Lipids). After centrifugation, supernatants were injected directly onto a 100 × 2.1 mm, 1.7-µm ACQUITY BEH C8 column (Waters). The column was eluted isocratically with 80% mobile phase A (95:5:0.1 v/v/v 10 mmol/L ammonium acetate/methanol/formic acid) for 1 min followed by a linear gradient to 80% mobile phase B (99.9:0.1 v/v methanol/formic acid) over 2 min, a linear gradient to 100% mobile phase B over 7 min, then 3 min at 100% mobile phase B. MS analyses were performed using electrospray ionization in the positive ion mode using full scan analysis over 200 to 1000 *m/z* at 70 000 resolution and 3 Hz data acquisition rate. Other MS settings were as follows: sheath gas 50, in source collision-induced dissociation 5 eV, sweep gas 5, spray voltage 3 kV, capillary temperature 300 °C, S-lens RF 60, heater temperature 300 °C, microscans 1, automatic gain control target 1e6, and maximum ion time 100 ms. Lipid identities were determined based on comparison with standards representative of each lipid class and reference plasma extracts. Identities were denoted by the total number of carbons in the lipid acyl chain(s) and the total number of double bonds in the lipid acyl chain(s).

Data quality was assured by monitoring internal standard performance in every sample during data acquisition and by regular analysis of pooled quality control (QC) samples. Pairs of pooled plasma QC samples were analyzed after intervals of approximately study samples. One pooled QC sample from each pair was used to correct for MS sensitivity drift across the run using "nearest-neighbor" scaling. The other pooled QC was used to calculate coefficients of variation (CVs) for every metabolite and unknown as a measure of analytical precision for the feature. Raw data were processed using TraceFinder 3.3 software (Thermo Fisher Scientific; Waltham, MA) and Progenesis QI (Nonlinear Dynamics; Newcastle upon Tyne, UK). Metabolite identities in the HILIC-pos, HILIC-neg, and C18-ng methods were confirmed using authentic reference standards.

A total of 758 known metabolites were initially identified. To further minimize data noise and enhance analytical power, a suite of computational methods called PAIRUP-MS[20], including a QC pipeline and filtering procedures, was employed alongside the initial QC conducted by the Broad Institute. Metabolites exhibiting more than 30% missing values across participant samples and samples with more than 25% missing values were excluded. Following QC, the remaining missing values for metabolites were imputed by leveraging the correlation structure among signals, utilizing the MICE R package (v3.16.0). For a minor subset of missing values that could not be imputed by MICE (<0.01% of all data), these were replaced with the median value of each signal. Consequently, 479 unique known metabolites were retained for analysis. However, samples from one participant failed to meet the QC criteria. We manually annotated all metabolites for the superclass, classes, and subclasses based on the Human Metabolome Database using their HMDB annotators (HMDB, http://www.hmdb.ca).

Covariates

Analysis performed on each metabolite among post-weight loss test diets included adjustments for the following covariates: study site, cohort, enrollment wave, age, sex, race, ethnicity, BMI, percentage of fat mass, and percentage of lean mass. In addition, the values were adjusted for the percentage of weight loss during the run-in phase, delta weight change over the test phase, and metabolite levels at randomization. Furthermore, a rank-based inverse normal transformation was applied to each metabolite signal, with the resulting abundance *z*-scores being utilized in downstream analyses.

Analytic approach

A total of 147 participants (102 women and 45 men; mean age 38.62 ± 14.47 years) were included in the analyses, conducted according to originally assigned dietary groups: low-carbohydrate ($n = 54$; 34 women; 37.18 ± 13.6 years), moderate-carbohydrate ($n = 48$; 32 women; 38.77 ± 14.86 years), and high-carbohydrate ($n = 45$; 36 women; 40.21 ± 15.21 years; Table 1). Descriptive data are presented as mean (standard deviation) for continuous variables and *n* (%) for categorical variables. Comparisons between groups were conducted using ANOVA or Kruskal-Wallis tests and T-tests or Wilcoxon tests, as appropriate, depending on the distribution of the continuous variables. The Chi-squared test with Yates' continuity correction and Fisher's Exact Test (when the expected cell counts were less than 5) were applied for the categorical variables, as appropriate.

For the analysis of individual metabolites, a rank-based inverse normal transformation was applied to the values of each metabolite to mitigate skewness, resulting in abundance *z*-scores utilized in downstream analyses adjusted for the aforementioned covariates. Linear regression analyses were conducted using the delta change of the metabolites throughout the test phase (END minus START) to assess trends across post-weight loss test diets. CFR was modeled as an ordinal variable in the linear models, with the low-carbohydrate diet coded as 1, the moderate-carbohydrate diet as 2, and the high-carbohydrate diet as 3, reflecting increasing CFR across diets. In addition, the metabolite changes during the early test phase (MID minus START timepoints) were evaluated among the diets in the same way. We evaluated the correlation between metabolite level changes across low-, moderate-, and high-CFR diets, comparing both early and entire test phases, as well as between the main analysis and sensitivity analyses. This assessment was performed using Pearson (r)/Spearman correlation (rho) coefficients, as appropriate, by comparing the t-statistics of the linear regression from both analyses for all included metabolites. For analyses involving differing participant populations, we adjusted the t-statistic value by dividing it by the square root of the total number of participants in each analysis.

Furthermore, to elucidate the complex interrelationships among FDR statistically significant diet-related metabolites, we employed the Fruchterman-Reingold force-directed layout algorithm to construct a network plot. This visualization delineates the metabolite interconnections (and their respective classes) characterized by significant correlations ($r > 0.6$).

In addition, we conducted PLS-DA to evaluate intergroup differences and classify participants according to their diet types. Metabolite data were used as predictors, with diet type serving as the outcome variable. We included the FDR-significant metabolites in the PLS-DA models to enhance the signal-to-noise ratio, reduce the risk of overfitting, and improve model interpretability. Performance of the models was evaluated using 5-fold cross-validation (CV), In each fold, the data were randomly split into 80:20 training/testing ratios, and predictions were based on Mahalanobis distance. The optimal number of components was determined by selecting the model that minimized the classification error rate across cross-validation repeats, while also considering the proportion of variance explained by each component. The first two components yielded the lowest balanced error rate and together accounted for the majority of the variance in the predictor space.

Our analysis focused on the most extreme diets (low-carbohydrate and high-carbohydrate) for a targeted approach, while also encompassing all three diet types. By convention, variables with a VIP score greater than 1 are often considered potential contributors to the

PLS-DA model. In our study, we also adopted an even more stringent criterion of VIP greater than 1.5 based on the PLS-DA model.

To evaluate the robustness of metabolite selection and classification performance, a LASSO (Least Absolute Shrinkage and Selection Operator) logistic regression model was implemented in parallel with the PLS-DA analysis. The model was trained on 80% of the data (80:20 training/testing ratios) using stratified sampling and 5-fold cross-validation. The optimal penalty parameter (λ) was selected based on the minimization of the classification error. Model performance was assessed on the testing set and evaluated based on classification accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and balanced accuracy. These metrics and metabolite selection were comparable to those of the PLS-DA models, and therefore, the PLS-DA model was retained as the modeling approach.

Furthermore, we conducted an exploratory logistic regression analysis to test for associations of metabolites with weight gain (more than 2 kg) throughout the test phase (END minus START). Weight gain >2 kg was considered a potential surrogate marker for lower dietary adherence in each diet group, so this analysis aimed to assess whether any CFR-associated metabolites might also reflect behavioral lower adherence.

External validation – data from the Popular Diets Study

To assess the results of our primary analysis and identify diet-responsive metabolites, we utilized metabolomic data from the Popular Diets Study, an earlier randomized crossover feeding trial[16], as an external validation study to ensure the robustness of our findings. The trial was registered at ClinicalTrials.gov (identifier: NCT00315354). These data were obtained using the Broad Metabolomics Platform, employing the same four complementary LC-MS methods as in the (FS)2. The primary outcomes of the trial have been previously published[17]. The study included both a run-in phase and a test phase, incorporating a dietary intervention and a weight-loss maintenance period. We assigned 21 participants to three test diets in random order, each lasting four weeks: low-fat (60% carbohydrate, 20% fat, 20% protein), low glycemic index (40% carbohydrate, 40% fat, 20% protein), or very-low carbohydrate (10% carbohydrate, 60% fat, 30% protein). We employed the same data processing, QC, handling of missing values, and analytic approaches as those applied in the (FS)2 to identify changes in metabolite levels among the diets with higher CFR order. Consistent with the main analysis, CFR was modeled as an ordinal variable in the linear models, with the low-, moderate-, and high-carbohydrate diets coded as 1, 2, and 3, respectively, to reflect increasing CFR. Analysis performed included adjustments for the following covariates: sequence of diets, age, sex, race, ethnicity, BMI, percentage of fat mass, percentage of lean mass, metabolite levels at randomization, percentage of weight loss during the run-in phase, and delta weight change over the test phase. The known metabolites identified in both the Popular Diets Study and the (FS)2 were included in the validation analyses to facilitate direct comparisons between the two RCTs.

Metabolite trends and classes in (FS)2 and T2D risk

We investigated the association between metabolite trends in the (FS)2 and the potential risk of developing type 2 diabetes (T2D) by leveraging data from a recent meta-analysis[18] encompassing plasma and serum metabolites associated with T2D risk, in order to extrapolate additional insights and potential metabolic implications from our findings. We identified metabolites that exhibited FDR significance in both analyses and compared their trends to ascertain concordance or discordance.

Furthermore, we used the metabolite groups provided by the Broad Institute, based on biological or molecular similarities, and conducted Metabolite Set Enrichment Analysis (MSEA)[48] to capture broader metabolic patterns associated with higher CFR and T2D risk, incorporating FDR adjustment for multiple comparisons. To perform MSEA and enable comparisons across metabolite subclasses, we converted the summary relative risk values from the published meta-analysis into beta coefficients using the natural logarithm. MSEA prioritizes metabolites based on the beta coefficients associated with CFR and T2D risk to identify enriched metabolite groups at both ends of the ordered beta coefficient spectrum, indicating either positive or negative associations. MSEA combines the effect estimates of individual metabolites within each defined group to calculate a summary enrichment score and a normalized enrichment score (NES), adjusted for group size. The NES reflects the degree to which a metabolite group is overrepresented relative to other groups, with a positive NES indicating positive enrichment and a negative NES indicating negative enrichment with respect to CFR and T2D risk. A significance level of FDR <0.05 was set for MSEA, with FDR <0.2 considered nominally significant for subsequent analyses, given the exploratory nature of our study.

Sensitivity analyses

To enhance the robustness of our findings, sensitivity analyses were performed using the same analytical procedures as those applied in the main study. With weight-loss maintenance within ± 2 kg specified as a surrogate marker of adherence to the feeding protocol (weight trajectories illustrated in Supplementary Fig. 8), one sensitivity analysis excluded participants who gained > 2 kg during the test phase. Additional sensitivity analyses were conducted by excluding one participant who developed hypothyroidism during the intervention and by excluding this participant as well as those who gained weight >2 kg during the test phase, with findings consistent when compared to the primary analysis.

To evaluate the robustness of the findings based on the predefined ±2 kg threshold for weight-loss maintenance, additional analyses were conducted using alternative thresholds of ± 1.5 kg, ± 2.5 kg, and ± 3 kg. T-tests or Wilcoxon tests were performed to statistically compare the metabolite effects between the ± 2 kg threshold and each alternative threshold. Consistency of effect trends (betas) for FDR-significant metabolites across all examined thresholds (± 1.5 kg, ± 2 kg, ± 2.5 kg, and ± 3 kg) was also assessed.

Furthermore, to evaluate the impact of the functional form of the CFR variable, we conducted a sensitivity analysis in which CFR was treated as a continuous variable (numerically reflecting increasing carbohydrate content), rather than as an ordinal variable. The same linear regression models and covariate adjustments were applied. To assess concordance between modeling approaches, we compared the beta coefficients from the continuous and ordinal models across all metabolites and computed the Spearman rank correlation. We also evaluated the direction and statistical significance of effect estimates for consistency. The results from the two approaches were highly consistent (rho = 0.991, $P < 2.2e\text{-}16$ for beta estimates across all metabolites). Specifically, of the 142 metabolites that were FDR-significant in the continuous CFR model, 139 (97.9%) were also significant in the ordinal model and exhibited concordant directional trends. Three metabolites were uniquely significant in the continuous model, and nine were uniquely significant in the ordinal model. No discordant directions were observed for any metabolites that were FDR-significant in either analysis (Supplementary Data 10).

Mediation analyses

We conducted mediation analyses to test whether changes in specific metabolites mediated the association between dietary fat-to-carbohydrate ratio and a set of cardiometabolic risk biomarkers, including HDL, LDL, non-HDL cholesterol, total cholesterol, lipoprotein(a), the lipoprotein insulin resistance index (LP-IR), 1,5-anhydroglucitol, high-sensitivity C-reactive protein, diastolic blood pressure, and systolic blood pressure. These biomarkers were assessed

as part of the (FS)2 study and have been described in detail in prior publications[8,49].

Mediation analysis was conducted using a two-step regression modeling approach: (1) a mediator model, in which the association between dietary pattern and each metabolite was estimated using linear regression; and (2) an outcome model, in which the association between dietary pattern, metabolite levels, and each cardiometabolic biomarker was estimated using a second linear regression model. Covariates included in both models were the same as those used in the main analysis, as described above. To obtain robust and interpretable results, we restricted mediation analyses to diet-responsive metabolites identified in the (FS)2 and validated in the independent Popular Diets Study. To further reduce collinearity and the multiple testing burden, we excluded metabolites with high pairwise correlation (correlation coefficient >0.6), assessed using Pearson or Spearman methods as appropriate, retaining the metabolite with the lower coefficient of variation in each correlated pair. Mediation analysis was performed for each remaining metabolite-outcome pair using a quasi-Bayesian approximation with 1000 Monte Carlo simulations to estimate the indirect effect (Average Causal Mediation Effect, ACME), direct effect, and total effect of the dietary pattern on cardiometabolic biomarkers; 95% confidence intervals were derived for each effect estimate; and the mediation proportion was calculated as the proportion of the total effect mediated by each metabolite. The two-step models were implemented using the mediate() function from the mediation R package, according to its default settings.

All statistical tests in the present study were two-tailed, with a significance threshold set at $P<0.05$, and correction for multiple comparisons was performed using the FDR ($P<0.05$). Normality of continuous variables was assessed using the Shapiro-Wilk test. Data not under restricted access (due to consent restrictions) that support the findings of this study are available within the article, its supplementary files, and the source data. De-identified individual-level data are available to qualified researchers under restricted access upon request to the corresponding authors, subject to Institutional Review Board approval and data use agreements (see Data availability statement). Analyses were performed using R software version 4.3.2 (R packages: dplyr (v1.1.4), ggplot2 (v3.5.0), tidyverse (v2.0.0), igraph (v2.0.3), caret (v6.0.94), fgsea (v1.28.0), mixOmics (v6.26.0), circlize (v0.4.16), stats (v4.3.2), ComplexHeatmap (v2.18.0), glmnet (v4.1.8), mice (v3.16.0), pROC (v1.18.5), RColorBrewer (v1.1.3), mediation (v4.5.1), reshape2 (v1.4.4)).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The full trial protocol, demographic characteristics of participants, and cardiometabolic outcomes from the (FS)2 trial are publicly available through Open Science Framework (https://osf.io/rvbuy/). All data supporting the findings described in this manuscript are available in the article and in the Supplementary Information and Supplementary Data. De-identified metabolomics data are available under controlled access to qualified investigators for non-commercial research purposes. The raw data are protected and not publicly available due to consent restrictions. Access to metabolomics data can be requested by contacting the corresponding authors (joel.hirschhorn@childrens.harvard.edu; cara.ebbeling@childrens.harvard.edu). Requests will be reviewed in accordance with IRB approval and applicable ethical and legal requirements. Eligible investigators will be required to sign a Data Use Agreement (DUA). Upon execution of the DUA, requests will be fulfilled within 4 weeks, and data will remain available for the period specified in the DUA (at least one year). Trials were registered at ClinicalTrials.gov: (FS)2, NCT02068885, registered on 21 February 2014 (https://clinicaltrials.gov/study/NCT02068885); and the Popular Diets Study (validation study), NCT00315354, registered on 18 April 2006 (https://clinicaltrials.gov/study/NCT00315354). Source data are provided in this paper.

Code availability

The code and documentation for PAIRUP-MS, which was previously developed and published and applied in the present paper, are publicly available on GitHub (https://github.com/yuhanhsu/PAIRUP-MS). No new custom or unique computational code was developed for this study. All packages used are publicly available and have been cited in the manuscript.

References

1. Micha, R. et al. Etiologic effects and optimal intakes of foods and nutrients for risk of cardiovascular diseases and diabetes: Systematic reviews and meta-analyses from the Nutrition and Chronic Diseases Expert Group (NutriCoDE). *PLoS ONE* **12**, e0175149 (2017).
2. Schwingshackl, L. et al. Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies. *Eur. J. Epidemiol.* **32**, 363–375 (2017).
3. Sacks, F. M. et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N. Engl. J. Med.* **360**, 859–873 (2009).
4. GBD 2017 Diet Collaborators Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **393**, 1958–1972 (2019).
5. Hu, Y. et al. Low-Carbohydrate Diet Scores and Mortality Among Adults With Incident Type 2 Diabetes. *Diabetes Care* **46**, 874–884 (2023).
6. Lopez-Garcia, E. et al. The Mediterranean-style dietary pattern and mortality among men and women with cardiovascular disease. *Am. J. Clin. Nutr.* **99**, 172–180 (2014).
7. Bhupathiraju, S. N. et al. Glycemic index, glycemic load, and risk of type 2 diabetes: results from 3 large US cohorts and an updated meta-analysis. *Am. J. Clin. Nutr.* **100**, 218–232 (2014).
8. Ebbeling, C. B. et al. Effects of a low carbohydrate diet on energy expenditure during weight loss maintenance: randomized trial. *BMJ* **363**, k4583 (2018).
9. Ludwig, D. S. et al. The carbohydrate-insulin model: a physiological perspective on the obesity pandemic. *Am. J. Clin. Nutr.* **114**, 1873–1885 (2021).
10. Hengist, A. et al. Ketogenic diet but not free-sugar restriction alters glucose tolerance, lipid metabolism, peripheral tissue phenotype, and gut microbiome: RCT. *Cell Rep. Med.* **5**, 101667 (2024).
11. Vogelzangs, N. et al. Metabolic profiling of tissue-specific insulin resistance in human obesity: results from the Diogenes study and the Maastricht Study. *Int. J. Obes.* **44**, 1376–1386 (2020).
12. Meyer, A. et al. Plasma metabolites and lipids predict insulin sensitivity improvement in obese, nondiabetic individuals after a 2-phase dietary intervention. *Am. J. Clin. Nutr.* **108**, 13–23 (2018).
13. Guasch-Ferré, M., Bhupathiraju, S. N. & Hu, F. B. Use of metabolomics in improving assessment of dietary intake. *Clin. Chem.* **64**, 82–98 (2018).
14. Ulaszewska, M. M. et al. Nutrimetabolomics: an integrative action for metabolomic analyses in human nutritional studies. *Mol. Nutr. Food Res.* **63**, e1800384 (2019).
15. Posma, J. M. et al. Nutriome-metabolome relationships provide insights into dietary intake and metabolism. *Nat. Food* **1**, 426–436 (2020).
16. Esko, T. et al. Metabolomic profiles as reliable biomarkers of dietary composition. *Am. J. Clin. Nutr.* **105**, 547–554 (2017).

17. Ebbeling, C. B. et al. Effects of dietary composition on energy expenditure during weight-loss maintenance. *JAMA* **307**, 2627–2634 (2012).
18. Morze, J. et al. Metabolomics and type 2 diabetes risk: an updated systematic review and meta-analysis of prospective cohort studies. *Diabetes Care* **45**, 1013–1024 (2022).
19. Ebbeling, C. B. et al. A randomized study of dietary composition during weight-loss maintenance: Rationale, study design, intervention, and assessment. *Contemp. Clin. Trials* **65**, 76–86 (2018).
20. Hsu, Y.-H. H. et al. PAIRUP-MS: Pathway analysis and imputation to relate unknowns in profiles from mass spectrometry-based metabolite data. *PLoS Comput. Biol.* **15**, e1006734 (2019).
21. Adams, S. H. et al. Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid beta-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. *J. Nutr.* **139**, 1073–1081 (2009).
22. Audzeyenka, I. et al. β-Aminoisobutyric acid (L-BAIBA) is a novel regulator of mitochondrial biogenesis and respiratory function in human podocytes. *Sci. Rep.* **13**, 766 (2023).
23. Sandoval-Salazar, C., Ramírez-Emiliano, J., Trejo-Bahena, A., Oviedo-Solís, C. I. & Solís-Ortiz, M. S. A high-fat diet decreases GABA concentration in the frontal cortex and hippocampus of rats. *Biol. Res.* **49**, 15 (2016).
24. Wang, Q., Ren, L., Wan, Y. & Prud'homme, G. J. GABAergic regulation of pancreatic islet cells: Physiology and antidiabetic effects. *J. Cell Physiol.* **234**, 14432–14444 (2019).
25. Ma, X. et al. ACE2 modulates glucose homeostasis through GABA signaling during metabolic stress. *J. Endocrinol.* **246**, 223–236 (2020).
26. Geng, J., Ni, Q., Sun, W., Li, L. & Feng, X. The links between gut microbiota and obesity and obesity related diseases. *Biomed. Pharmacother.* **147**, 112678 (2022).
27. Tuomainen, M. et al. Associations of serum indolepropionic acid, a gut microbiota metabolite, with type 2 diabetes and low-grade inflammation in high-risk individuals. *Nutr. Diabetes* **8**, 35 (2018).
28. Cipryan, L. et al. A lipidomic and metabolomic signature of a very low-carbohydrate high-fat diet and high-intensity interval training: an additional analysis of a randomized controlled clinical trial. *Metabolomics* **20**, 10 (2023).
29. Dambrova, M. et al. Acylcarnitines: nomenclature, biomarkers, therapeutic potential, drug targets, and clinical trials. *Pharmacol. Rev.* **74**, 506–551 (2022).
30. Strand, E. et al. Serum acylcarnitines and risk of cardiovascular death and acute myocardial infarction in patients with stable angina pectoris. *J. Am. Heart Assoc.* **6**, e003620 (2017).
31. An, J. et al. Hepatic expression of malonyl-CoA decarboxylase reverses muscle, liver and whole-animal insulin resistance. *Nat. Med.* **10**, 268–274 (2004).
32. Lu, J. et al. High-coverage targeted lipidomics reveals novel serum lipid predictors and lipid pathway dysregulation antecedent to type 2 diabetes onset in normoglycemic Chinese adults. *Diabetes Care* **42**, 2117–2126 (2019).
33. Rhee, E. P. et al. Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *J. Clin. Invest.* **121**, 1402–1411 (2011).
34. Hornburg, D. et al. Dynamic lipidome alterations associated with human health, disease and ageing. *Nat. Metab.* **5**, 1578–1594 (2023).
35. Wang, F. et al. Plasma metabolomic profiles associated with mortality and longevity in a prospective analysis of 13,512 individuals. *Nat. Commun.* **14**, 5744 (2023).
36. van der Veen, J. N. et al. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochim. Biophys. Acta Biomembr.* **1859**, 1558–1572 (2017).
37. Zhong, J. et al. Specific triacylglycerol, diacylglycerol, and lysophosphatidylcholine species for the prediction of type 2 diabetes: a ~16-year prospective study in Chinese. *Cardiovasc. Diabeto.l* **21**, 234 (2022).
38. Meikle, P. J. & Summers, S. A. Sphingolipids and phospholipids in insulin resistance and related metabolic disorders. *Nat. Rev. Endocrinol.* **13**, 79–91 (2017).
39. Dean, J. M. & Lodhi, I. J. Structural and functional roles of ether lipids. *Protein Cell* **9**, 196–206 (2018).
40. Pietiläinen, K. H. et al. Acquired obesity is associated with changes in the serum lipidomic profile independent of genetic effects–a monozygotic twin study. *PLoS ONE* **2**, e218 (2007).
41. Chang, W., Hatch, G. M., Wang, Y., Yu, F. & Wang, M. The relationship between phospholipids and insulin resistance: From clinical to experimental studies. *J. Cell Mol. Med.* **23**, 702–710 (2019).
42. Holland, W. L. & Summers, S. A. Sphingolipids, insulin resistance, and metabolic disease: new insights from in vivo manipulation of sphingolipid metabolism. *Endocr. Rev.* **29**, 381–402 (2008).
43. Meikle, P. J. et al. Plasma lipid profiling shows similar associations with prediabetes and type 2 diabetes. *PLoS ONE* **8**, e74341 (2013).
44. Barber, M. N. et al. Plasma lysophosphatidylcholine levels are reduced in obesity and type 2 diabetes. *PLoS ONE* **7**, e41456 (2012).
45. Wang, W. et al. Dietary patterns and cardiometabolic health: Clinical evidence and mechanism. *MedComm* **4**, e212 (2023).
46. Goldenberg, J. Z. et al. Efficacy and safety of low and very low carbohydrate diets for type 2 diabetes remission: systematic review and meta-analysis of published and unpublished randomized trial data. *BMJ* **372**, m4743 (2021).
47. Ludwig, D. S., Dickinson, S. L., Henschel, B., Ebbeling, C. B. & Allison, D. B. Do lower-carbohydrate diets increase total energy expenditure? an updated and reanalyzed meta-analysis of 29 controlled-feeding studies. *J. Nutr.* **151**, 482–490 (2021).
48. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
49. Ebbeling, C. B. et al. Effects of a low-carbohydrate diet on insulin-resistant dyslipoproteinemia-a randomized controlled feeding trial. *Am. J. Clin. Nutr.* **115**, 154–162 (2022).

Acknowledgements

We thank the study investigators and participants of the (FS)2 for their time and effort. We thank Jakub Morze, the first author of the meta-analysis Metabolomics and Type 2 Diabetes, for generously providing data for analysis in this study. We would like to thank the Consortium of METabolomics Studies (COMETS) for promoting collaborations among the investigators. M.Y.M. was supported by NIH grant HL171855. N.ET. and C.W. were supported by the SciLifeLab & Wallenberg Data-Driven Life Science Program (grant: KAW 2020.0239). Analyses for this paper were supported by the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK125273), and additional support was provided by R01DK075787. Implementation of the Framingham State Food Study was supported by grants from Nutrition Science Initiative (made possible by gifts from Arnold Ventures and Robert Lloyd Corkin Charitable Foundation), New Balance Foundation, Many Voices Foundation, and Blue Cross Blue Shield. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation and writing of the manuscript.

Author contributions

C.B.E. and J.N.H. designed and coordinated the study; D.S.L. and C.B.E. designed and implemented the (FS)2; A.M.A. analyzed data and designed the figures with input from E.B., Y.H., O.A.Z., and J.N.H.; A.M.A. wrote the first draft; A.M.A., C.B.C., C.B.E., and J.N.H. wrote the paper; C.B.C. conducted the metabolomics measurements; J.LS, C.B.E., and J.N.H. supervised the project. All authors (A.M.A., E.B., Y.H., O.A.Z., N.ET., M.Y.M., S.N.B., R.S.K., C.W., J.L.-S., C.B.C., D.S.L., C.B.E., and J.N.H.)

contributed to the discussion and interpretation of results, critically reviewed the paper, and agreed to submit the paper for publication.

Competing interests

D.S.L. received royalties for books on obesity and nutrition that recommend a low-glycemic load diet. J.L.-S. is a scientific advisor to TruDiagnostic and Precion Inc. The rest of the authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-68353-z.

Correspondence and requests for materials should be addressed to Cara B. Ebbeling or Joel N. Hirschhorn.

Peer review information *Nature Communications* thanks Armand Valsesia, Stefania Noerman and Amir Asiaee for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Angeliki M. Angelidi[1,2,3], Eric Bartell[1,2,4], Yisong Huang[1,2], Oana A. Zeleznik[5], Núria Estanyol-Torres[6], Michael Y. Mi[7], Shilpa N. Bhupathiraju[5], Rachel S. Kelly[5], Clemens Wittenbecher[6], Jessica Lasky-Su[5], Clary B. Clish[8], David S. Ludwig[3,9,10], Cara B. Ebbeling[3,9] ✉ & Joel N. Hirschhorn[1,2,4] ✉

[1]Division of Endocrinology, Boston Children's Hospital, Boston, MA, USA. [2]Program in Metabolism, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [3]Department of Pediatrics, Harvard Medical School, Boston, MA, USA. [4]Departments of Pediatrics and Genetics, Harvard Medical School, Boston, MA, USA. [5]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [6]Department of Life Sciences, SciLifeLab, Chalmers University of Technology, Gothenburg, Sweden. [7]Division of Cardiovascular Medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA. [8]Metabolomics Platform, Broad Institute, Cambridge, MA, USA. [9]New Balance Foundation Obesity Prevention Center, Division of Endocrinology, Boston Children's Hospital, Boston, MA, USA. [10]Steno Diabetes Center Copenhagen, Herlev, Denmark.
✉e-mail: cara.ebbeling@childrens.harvard.edu; joel.hirschhorn@childrens.harvard.edu