The American Journal of Clinical Nutrition xxx (xxxx) xxx

The American Journal of CLINICAL NUTRITION

journal homepage: https://ajcn.nutrition.org/

Original Research Article

Association of the EAT-Lancet diet, serial measures of serum proteome and gut microbiome, and cardiometabolic health: a prospective study of Chinese middle-aged and elderly adults

Kui Deng [1,2,3,4,†], Luqi Shen [1,3,4,†], Zhangzhi Xue [1,4,5,†], Bang-yan Li [2,†], Jun Tang [1,3,4], Hui Zhao [1,3,4], Fengzhe Xu [1,3,4], Zelei Miao [1,3,4,6], Xue Cai [1,4,5], Wei Hu [2], Yuanqing Fu [1,3,4,6], Zengliang Jiang [1,3,4], Xinxiu Liang [1,3,4], Congmei Xiao [1,3,4], Menglei Shuai [1,3,4], Wanglong Gou [1,3,4], Liang Yue [1,4,5], Yuting Xie [1,4,5], Ting-yu Sun [2], Tiannan Guo [1,4,5,6], Yu-ming Chen [2,*], Ju-Sheng Zheng [1,3,4,6,**]

[1] *Westlake Center for Intelligent Proteomics, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, China;* [2] *Guangdong Provincial Key Laboratory of Food, Nutrition and Health, Department of Epidemiology, School of Public Health, Sun Yat-sen University, Guangzhou, China;* [3] *Research Center for Industries of the Future, School of Life Sciences, Westlake University, Hangzhou, China;* [4] *Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, China;* [5] *Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China;* [6] *Affiliated Hangzhou First People's Hospital, School of Medicine, Westlake University, Hangzhou, China*

A B S T R A C T

Background: The EAT-Lancet diet was reported to be mutually beneficial for the human cardiometabolic system and planetary health. However, mechanistic evidence linking the EAT-Lancet diet and human cardiometabolic health is lacking.

Objectives: We aimed to investigate the role of blood proteins in the association between the EAT-Lancet diet and cardiometabolic health and explore the underlying gut microbiota–blood protein interplay.

Methods: Our study was based on a prospective cohort including 3742 Chinese participants enrolled from 2008–2013 with serum proteome data repeatedly measured ≤3 times ($N_{\text{proteome}} = 7514$) and 1195 with gut metagenomic data measured ≤2 times over 9 y ($N_{\text{microbiota}} = 1695$). Least absolute shrinkage and selection operator and multivariable linear regression were used to explore the associations of the EAT-Lancet diet (assessed by semi-quantitative food frequency questionnaire) with serum proteins and gut microbes. Linear mixed-effect model and logistic regression were used to examine the associations of selected proteins with 11 cardiometabolic risk factors and 4 cardiometabolic diseases, respectively. Mediation analysis was used to identify potential mediation effects. Multiple comparisons were adjusted using the Benjamini-Hochberg method.

Results: The mean (standard deviation) age of enrolled participants was 58.4 (6.1) y (31.6% men). The EAT-Lancet diet was prospectively associated with 4 core proteins, including α-2-macroglobulin (A2M) (pooled β: 0.12; 95% confidence interval [CI]: 0.05, 0.2), retinol-binding protein 4 (pooled β: −0.14; 95% CI: −0.24, −0.04), TBC1 domain family member 31 (pooled β: −0.11; 95% CI: −0.22, 0), and adenylate kinase 4 (pooled β: −0.19; 95% CI: −0.3, −0.08). The identified proteins were prospectively associated with cardiometabolic diseases (pooled odds ratio ranged from 0.8–1.18) and risk factors (pooled β ranged from −0.1 to 0.12), mediating the association between the EAT-Lancet diet and blood triglycerides. We then identified 5 gut microbial biomarkers of the EAT-Lancet diet, and discovered a potential gut microbiota–blood protein interplay (EAT-Lancet diet→*Rothia mucilaginosa*→A2M) underlying the EAT-Lancet diet–cardiometabolic health association.

Conclusions: Our study presents key molecular evidence to support the role of EAT-Lancet diet adherence in promoting cardiometabolic health.

Abbreviations: A2M, α-2-macroglobulin; AK4, adenylate kinase 4; aMed, Alternate Mediterranean diet; BH, Benjamini-Hochberg; CI, confidence interval; DASH, Dietary Approaches to Stop Hypertension; DBP, diastolic blood pressure; FDR, false discovery rate; FFQ, food frequency questionnaire; HbA1c, glycated hemoglobin; LASSO, least absolute shrinkage and selection operator; MET, total metabolic equivalent of task; MetS, metabolic syndrome; MS, mass spectrometry; OR, odds ratio; PDI, plant-based diet index; RBP4, retinol-binding protein 4; SBP, systolic blood pressure; T2D, type 2 diabetes; SWATH, sequential window acquisition of all theoretical mass spectra; TBC1D31, TBC1 domain family member 31; TC, total cholesterol; TG, triglycerides.

* Corresponding author.

** Corresponding author.

E-mail addresses: chenyum@mail.sysu.edu.cn (Y.-m. Chen), zhengjusheng@westlake.edu.cn (J.-S. Zheng).

† KD, LS, ZX, and BL contributed equally.

https://doi.org/10.1016/j.ajcnut.2024.10.011

Received 20 May 2024; Received in revised form 20 September 2024; Accepted 16 October 2024; Available online xxxx

Keywords: EAT-Lancet diet, proteomics, gut microbiota, cardiometabolic health, multiomics integration

Introduction

The EAT-Lancet diet is a recently defined global healthy reference dietary pattern promoting human health and environmental sustainability [1–3]. This dietary pattern is characterized by emphasized intakes of fruits, vegetables, unsaturated fat, and nuts, and limited intakes of dairy products, saturated fat, and animal-source foods. EAT-Lancet diet adherence has been beneficially associated with cardiometabolic diseases, including ischemic heart disease [4–6], diabetes [4,7–9], stroke [10], and obesity [11]. While several physiological mechanisms for EAT-Lancet diet components have been demonstrated, such as their effects on established cardiovascular disease risk factors [12–14], the biomarkers of the EAT-Lancet diet and the underlying mechanisms are less clear.

The molecular mechanisms underlying the EAT-Lancet diet–cardiometabolic health association could potentially be unveiled through multiomics integration analysis. Proteomics is a promising approach for identifying the biomarkers of diets [15–18], providing a functional link between habitual diets and cardiometabolic health. Existing human studies linking diet and the blood proteome were mainly based on cross-sectional data, and little is known about the relationship between diet-related proteins and cardiometabolic diseases [19–22]. Similarly, while several EAT-Lancet diet components have been associated with the gut microbiome [13,23,24], whether and how EAT-Lancet diet adherence may influence the gut microbiome has not been fully elucidated [25]. Moreover, several recent human studies demonstrated the contributions of the gut microbiome to blood proteins involved in metabolism and inflammation [26,27], indicating the potential existence of a gut microbiome–blood protein interplay. Therefore, we hypothesized that blood proteins and gut microbes may be involved in the pathway from the EAT-Lancet diet to cardiometabolic health, which may help interpret the health benefit of this dietary pattern.

In this study, our main aim was to investigate the longitudinal associations between the EAT-Lancet diet, serum proteome, and cardiometabolic diseases and risk factors in a large prospective cohort of Chinese middle-aged and elderly participants. Our secondary aim was to examine the associations between the EAT-Lancet diet and gut microbiota and explore the potential gut microbiota–blood protein interplay underlying the association between the EAT-Lancet diet and cardiometabolic health.

Methods

Study design and population

The overall study design is displayed in Figure 1. This study was based on the Guangzhou Nutrition and Health Study, which consists of 4048 Chinese participants aged 40 to 75 y and living in the urban area of Guangzhou, China, for ≥5 y. We recruited the participants between 2008 and 2013 and followed them every 3 y. Participants with dietary information and without missing covariates were included in this study ($N = 3991$). Among them, 7620 fasting serum samples from 3796 participants were collected at baseline (2008–2013), the 2014–2017 follow-up visit, or the 2018–2019 follow-up visit. We excluded participants with cancer at baseline ($n = 15$) or extreme levels of dietary total energy intake (<800 kcal or >4000 kcal for men; <500 kcal or >3500 kcal for women) ($n = 39$), which resulted in 3742 participants with 7514 serum samples in our formal analysis. We further divided the study into the discovery and validation sets because they were measured at different time points with independent sample preparation, mass spectrometry (MS) acquisition, and data preprocessing [28]. The discovery set included 1897 participants with 4532 serum samples (baseline: 1745 samples; 2014–2017 follow-up visit: 1656 samples; 2018–2019 follow-up visit: 1131 samples); the validation set contained 1845 participants with 2982 serum samples (baseline: 1589 samples; 2014–2017 follow-up visit: 845 samples; 2018–2019 follow-up visit: 548 samples). The flowchart of study participants involved in proteomics analysis is shown in Supplemental Figure 1.

Because some participants were lost to follow-up and some did not provide stool samples during their stay at our study center, a total of 1844 stool samples from 1339 participants were available at the follow-up visits (2014–2019). After excluding participants without dietary information ($n = 126$), antibiotics usage within 2 wk ($n = 8$), or extreme levels of dietary total energy intake (<800 kcal or >4000 kcal for men; <500 kcal or >3500 kcal for women) ($n = 10$), 1195 participants with 1695 stool samples remained in this study (2014–2017 follow-up visit: 735 samples; 2018–2019 follow-up visit: 960 samples), of which 500 participants had paired stool samples and 695 had a single stool sample. There were 1094 participants with both stool and serum samples collected at the same time point, which were used for the gut microbiome–blood proteome integration analysis.

Written informed consent was obtained from all enrolled participants prior to sample collection. This study was approved by the Ethics Committee of the School of Public Health at Sun Yat-sen University and Westlake University.

Dietary assessment and EAT-Lancet score calculation

We obtained the information on habitual dietary intake in the past 1 y based on a validated semi-quantitative food frequency questionnaire (FFQ) with 79 food items at baseline (2008–2013) during face-to-face questionnaire interviews [23,29,30]. We classified the food items into 14 dietary components: whole grains; tubers and starchy vegetables; vegetables; fruits; dairy foods; beef, lamb, and pork; poultry; eggs; fish and seafood; dry beans, lentils, peas; soy foods; nuts; added fats (unsaturated/saturated fat ratio); and added sugars, which are the components of the EAT-Lancet diet [1]. We then created the EAT-Lancet score according to a prior study [4], which was based on the references set by the EAT-Lancet Commission [1]. One point was assigned to participants who adhered to the dietary components; otherwise, no points were assigned (the scoring criteria are shown in Supplemental Table 1). The EAT-Lancet score was calculated by summing the points for individual components and ranged from 0 (nonadherence) to 14 (perfect adherence).

Measurement of cardiometabolic health and covariates

Fasting venous blood samples were collected at baseline and follow-up visits. Insulin was measured by electrochemiluminescence immunoassay using a Cobas 8000/e602 immunoanalyzer (Roche Diagnostics). Glycated hemoglobin (HbA1c) was measured by HPLC using the Bole D-10 Hemoglobin A1c Program on a Bole D-10 Hemoglobin Testing System. Glucose, total cholesterol (TC), triglycerides (TG), HDL cholesterol, and LDL cholesterol were measured by

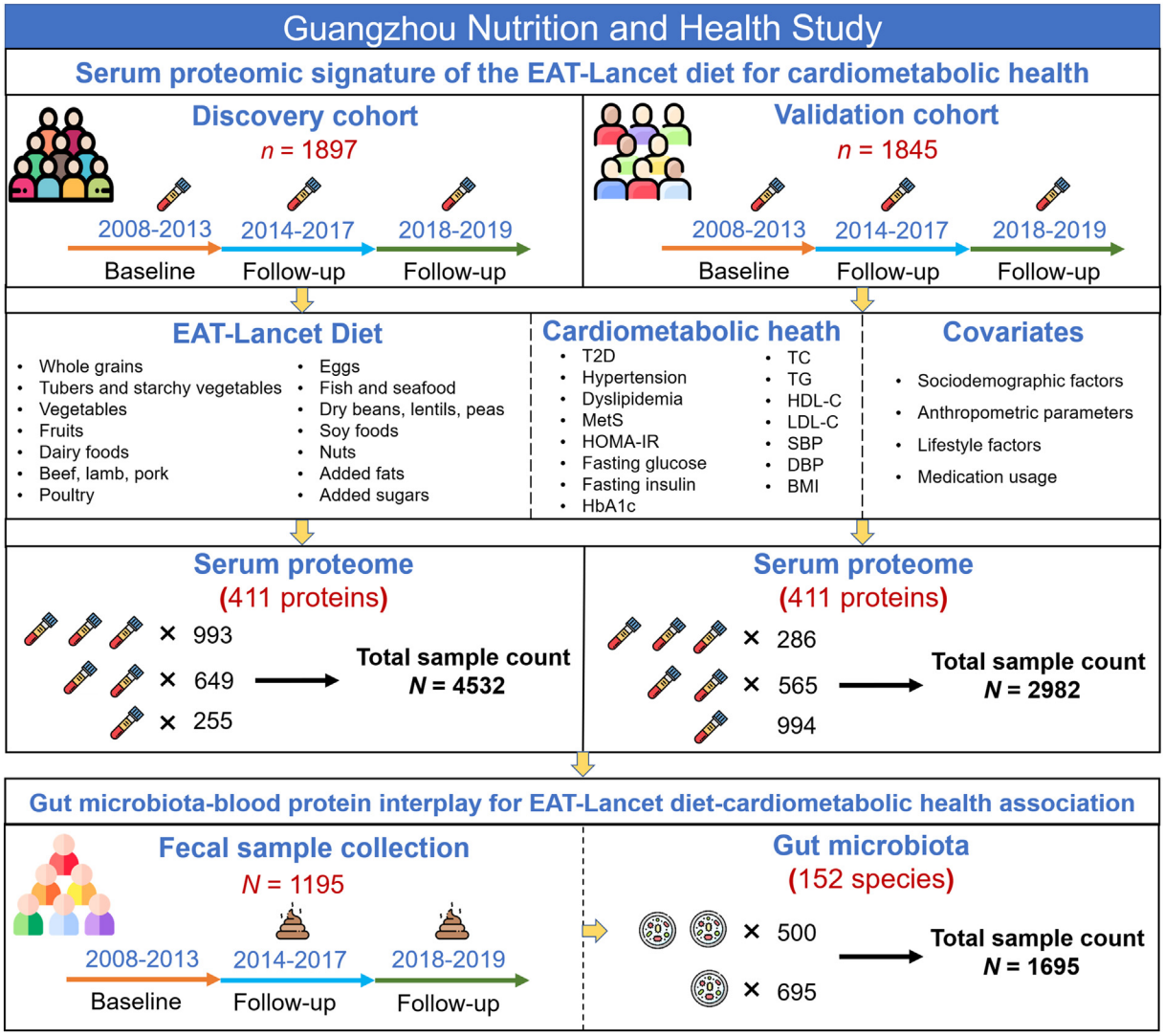


FIGURE 1. Study design. To identify serum proteomic signature of the EAT-Lancet diet for promoting cardiometabolic health and investigate the gut microbiota–blood protein interplay underlying the association between the EAT-Lancet diet and cardiometabolic health, we profiled serum proteomes and gut microbiomes from the Guangzhou Nutrition and Health Study, which is an ongoing prospective cohort involving 4048 middle-aged and elderly Chinese participants. Dietary information was collected at the baseline (2008–2013) using a FFQ. During the period 2008–2019, blood samples were collected at ≤3 time points per individual, and stool samples were collected at ≤2 time points. We obtained serum proteomics data by data-independent acquisition mass spectrometry and gut metagenomic data by shotgun metagenomic sequencing. Participants with proteomics data were divided into the discovery and validation sets based on the batch of sample preparation, mass spectrometry acquisition, and data preprocessing. Cardiometabolic risk factors including glucose homeostasis, lipid metabolism, blood pressure, and BMI were measured using standard methods. Cardiometabolic diseases were defined by recognized standards. DBP, diastolic blood pressure; HOMA-IR, homeostasis model assessment of insulin resistance; HbA1c, glycated hemoglobin; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; MetS, metabolic syndrome; SBP, systolic blood pressure; T2D, type 2 diabetes; TC, total cholesterol; TG, triglycerides.

colorimetric methods using a Cobas 8000 c702 automated analyzer (Roche Diagnostics). HOMA-IR was calculated based on fasting insulin and fasting glucose [31]. Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured using an arm digital sphygmomanometer (HEM-7011, OMRON Corporation). Anthropometric parameters, including height and weight, were measured by trained nurses on site, and BMI was calculated as weight in kilograms divided by height in meters squared.

Type 2 diabetes (T2D) was defined as fasting glucose ≥7.0 mmol/L (126 mg/dL), HbA1c ≥6.5% (48 mmol/mol), or self-reported use of medications for T2D [32]. Hypertension was diagnosed as SBP ≥140 mmHg, DBP ≥90 mmHg, or current antihypertensive medication use [33]. Dyslipidemia was based on TC ≥6.2 mmol/L, TG ≥2.3 mmol/L, LDL cholesterol ≥4.1 mmol/L, HDL cholesterol <1.0 mmol/L, or lipid-lowing drug intake [34]. Participants diagnosed as metabolic syndrome (MetS) met 3 of the 5 following criteria: *1*) waist circumference >90 cm (male) or >85 cm (female); *2*) fasting glucose ≥6.1 mmol/L (110 mg/dL) or previously diagnosed with T2D; *3*) TG ≥1.7 mmol/L (150 mg/dL); *4*) HDL cholesterol <1.04 mmol/L (40 mg/dL); and *5*) SBP/DBP ≥130/85 mmHg or previously diagnosed with hypertension [34].

Sociodemographic information was collected using a structured questionnaire during face-to-face interviews. The factors included age (y), sex (male or female), education (middle school or lower, high

school or professional college, or university), income (low, middle, or high), lifestyle factors (smoking status [current smoker or noncurrent smoker], alcohol status [current alcohol drinker or noncurrent alcohol drinker], and physical activity [total metabolic equivalent of task {MET}]), health status (T2D status [yes or no] and hypertension status [yes or no]), and medication usage (dyslipidemia medication usage [yes or no]). Physical activity was assessed as MET-h/d based on a validated physical activity questionnaire with 19 items [35].

Serum proteomics profiling and preprocessing

Blood samples from enrolled participants were collected on site on the examination day after overnight fasting, temporarily stored on ice, and then transported to the research laboratory. Serum was prepared using a standardized protocol and stored at −80°C in 0.5 mL aliquots within 4 h. Detailed information on peptide extraction, MS analysis, MS data preprocessing, quality control, and missing value imputation (missing value patterns are shown in Supplemental Figure 2) can be found in our recently published article [28] and the Supplementary Materials. The peptides were analyzed by sequential window acquisition of all theoretical mass spectra (SWATH)-MS using a TripleTOF 5600 system (SCIEX) coupled to a NanoLC 400 System (Eksigent) [36] and then analyzed using DIA-NN (1.8) software against a serum spectral library from the *Homo sapiens* Swiss-Prot database [36,37]. Samples in the discovery set were analyzed by SWATH-MS at 3 time points (1323, 1779, and 1430 samples, respectively; hereafter called inner sequencing batch). We obtained 411 proteins that overlapped between the discovery and validation sets for formal analysis.

Shotgun metagenomic sequencing and preprocessing

Stool samples were collected from each participant on site on the examination day, temporarily stored on ice, and then transported to the research laboratory and stored at −80°C within 4 h. Detailed information on microbial DNA isolation, shotgun metagenome sequencing, and bioinformatics processing of the raw metagenomic data can be found in our published articles [38–40] and the Supplementary Materials. We included 152 bacterial species with a minimum detectable relative abundance of 0.01% in ≥10% of the samples. The relative abundance of the genome data for specific microbial species was extracted from the stratified gene family data of the metagenomic samples obtained by HUMAnN2 (version 2.8.1) with default settings [41], which was based on the UniRef database [42]. We only included microbial genes that were detected in >10% of samples for analysis. We performed centered log-ratio transformation (zero values were replaced with 1×10^{-5} for species data and 1×10^{-9} for gene family data) and *z*-score normalization (zero-mean and unit-variance) for the relative abundance of microbial species and gene family data.

Statistical analysis

To evaluate the differences between the EAT-Lancet score and other standard diet quality scores, we assessed Spearman correlations of the EAT-Lancet score with the Dietary Approaches to Stop Hypertension (DASH) score [43,44], Alternate Mediterranean diet (aMed) score [45], and plant-based diet index (PDI) [46]. We categorized the EAT-Lancet score into 3 groups based on tertiles (highest tertile: 11–14; middle tertile: 10; lowest tertile: 1–9). In the discovery set, we used a 2-step strategy to select EAT-Lancet diet-related serum proteins. First, we used least absolute shrinkage and selection operator (LASSO) regression to identify the proteins associated with the baseline EAT-Lancet score, in which only follow-up proteomics data were included and transformed into *z*-scores. For participants with follow-up proteomics data measured at both 2014–2017 and 2018–2019 visits (993 participants), we only included the proteomics data from the 2018–2019 follow-up visit in the LASSO regression to capture the long-term effects of the EAT-Lancet diet on serum proteins. LASSO regression was implemented using R package glmnet (version 4.1-3) with the binomial link function for binary dependent variables (highest tertile/lowest tertile of the EAT-Lancet score) [47]. In the second step of our strategy, we further evaluated the prospective associations between baseline EAT-Lancet score (highest compared with lowest tertile) and future protein levels identified in the first step using multivariable linear regression in the discovery set, adjusted for potential covariates including age, sex, BMI, smoking status, alcohol status, education, income, physical activity, total energy intake, T2D status, hypertension status, dyslipidemia medication usage, time interval, inner sequencing batch, and corresponding baseline protein abundance. Similar to LASSO regression, only the follow-up proteomics data from the 2018–2019 visit were included in multivariable linear regression for participants with proteomics data measured at both the 2014–2017 and 2018–2019 follow-up visits. Proteins with skewed distribution were log-transformed, and all proteins were standardized into *z*-scores. We then used the same models to assess the linear trends of the above associations based on per-tertile difference in the EAT-Lancet score. We used the Benjamini-Hochberg (BH) method to control for the false discovery rate (FDR) caused by multiple testing.

Subsequently, we constructed the Lancet-protein index using the unweighted method (only regression coefficient direction was used due to its robustness to different datasets) based on the proteins identified in the above multivariable linear regression to capture the overall effect of the EAT-Lancet diet on serum proteins. We used the following formula to calculate the Lancet-protein index: Lancet-protein index= $\sum\pm$ normalized abundance of proteomic biomarkers, where ± signs depended on the signs of regression coefficients in multivariable linear regression, and the abundances of proteins and Lancet-protein index were normalized into *z*-scores. To test the reliability of the Lancet-protein index, we used multivariable linear regression to examine the association between baseline EAT-Lancet score (highest compared with lowest tertile) and future Lancet-protein index, adjusted for the same covariates as the above analysis of individual proteins except for the corresponding baseline Lancet-protein index. Linear trend analysis was performed for the above association using the same model based on per-tertile difference in the EAT-Lancet score. We examined the potential interactions of baseline EAT-Lancet score (highest compared with lowest tertile) with age (≥60 y compared with <60 y) and sex on levels of the identified proteins at follow-up and Lancet-protein index and performed subgroup analysis if there were significant interactions ($FDR_{interaction} < 0.05$).

We further performed sensitivity analysis for the associations of baseline EAT-Lancet score (highest compared with lowest tertile) with follow-up levels of the identified proteins and Lancet-protein index using a linear mixed-effect model including all repeated-measured proteomics data and accounting for within-person correlation, adjusting for the same covariates as above, wherein the corresponding baseline protein abundance or Lancet-protein index was adjusted. A linear trend test was also performed using a linear mixed-effect model based on per-tertile difference in the EAT-Lancet score. The linear mixed-effect model was implemented using the R package lme4 (version 1.1-27.1) [48].

We then explored the associations of EAT-Lancet diet-related proteins and Lancet-protein index with cardiometabolic risk factors (fasting glucose, fasting insulin, HOMA-IR, HbA1c, TC, TG, HDL

cholesterol, LDL cholesterol, SBP, DBP, and BMI) using a linear mixed-effect model accounting for serial measures and within-person correlation, adjusted for age, sex, BMI, smoking status, alcohol status, education, income, physical activity, total energy intake, and inner sequencing batch (BMI was not adjusted when it served as the dependent variable) among all participants with proteomics data. The concentrations of fasting glucose, fasting insulin, HOMA-IR, and TG were log-transformed. We also investigated the associations of baseline EAT-Lancet diet-related proteins and Lancet-protein index with follow-up cardiometabolic disease incidence (T2D, hypertension, MetS, and dyslipidemia) using logistic regression, adjusted for the same covariates as the above analysis. In this analysis, participants with cardiometabolic diseases diagnosed at the baseline were removed. Multiple comparisons were controlled for using the BH method across the full set of outcomes and separately for cardiometabolic risk factors and diseases. We examined the potential interactions of EAT-Lancet diet-related proteins and Lancet-protein index with age (≥60 y compared with <60 y) and sex on cardiometabolic risk factors and diseases and performed subgroup analysis if significant interactions ($FDR_{interaction} < 0.05$) existed. The abundance of EAT-Lancet diet-related proteins, Lancet-protein index, and cardiometabolic risk factors were standardized into *z*-scores.

Furthermore, we performed mediation analysis to evaluate whether proteins and Lancet-protein index could mediate the associations between the EAT-Lancet diet and cardiometabolic health. To ensure temporal ordering of exposure, mediator, and outcome, the baseline EAT-Lancet score (highest compared with lowest tertile) served as the exposure, proteins and Lancet-protein index measured at the 2014–2017 follow-up visit served as the mediators, and cardiometabolic risk factors measured at the 2018–2019 follow-up visit served as the outcomes. Because there were few incident cardiometabolic diseases at the 2018–2019 follow-up visit (the incident cases for T2D, hypertension, MetS, and dyslipidemia were 90, 56, 85, and 58, respectively, in the discovery set and 10, 10, 6, and 8, respectively, in the validation set), the mediation analysis for cardiometabolic diseases was not performed. Mediation analysis were conducted using R package mediation (version 4.5.0) [49].

The results from the above analyses were replicated in the validation set. The effect estimates from the discovery and validation sets were pooled by fixed-effect meta-analysis using the R package metafor (version 3.0-2) [50,51].

We then explored the potential gut microbiota–blood protein interplay underlying the EAT-Lancet diet–cardiometabolic health association. Detailed statistical analyses for the investigation of the gut microbiota–blood protein interplay can be found in the Supplemental Methods. Briefly, we used the same 2-step strategy mentioned above (LASSO regression and multivariable linear regression) to identify gut microbes associated with the EAT-Lancet diet. Then, we used multivariable linear regression to assess the cross-sectional associations between the identified gut microbes and serum proteins. Among the identified gut microbe–blood protein associations, we investigated whether gut microbes could mediate the associations between the EAT-Lancet diet and serum proteins using mediation analysis. For identified microbial species that may mediate the EAT-Lancet diet-protein associations, we further investigated the specific genes of the species that may link the EAT-Lancet diet and serum proteins. We used the abovementioned 2-step strategy (LASSO regression and multivariable linear regression) to select EAT-Lancet diet-related microbial genes. The associations between the identified EAT-Lancet diet-related genes and serum proteins were examined by multivariable linear regression.

In the above multivariable linear regression and linear mixed-effect model analyses, the normality and homoscedasticity assumptions were assessed using the normal quantile plot of the residuals and the scatterplot of residuals compared with predicted values, respectively, and there were no violations of normality and homoscedasticity assumptions in our analyses.

All statistical analyses were performed using R software (version R-4.1.1). FDR < 0.05 or $P < 0.05$ was considered statistically significant.

Results

Participant characteristics

The mean (SD) age of our study participants was 58.4 (6.1) years (31.6% men) at enrollment (Table 1). Higher adherence to the EAT-Lancet

TABLE 1
Baseline characteristics of study participants based on the EAT-Lancet score.

	All ($N = 3751$)	Lower tertile ($N = 1859$)	Middle tertile ($N = 1173$)	Upper tertile ($N = 719$)	*P*
Age, y	58.4 (6.1)	58.3 (6.1)	58.5 (6.0)	58.5 (6.1)	0.483
Male, *n* (%)	1184 (31.6)	586 (31.5)	378 (32.2)	220 (30.6)	0.760
BMI, kg/m^2	23.3 (3.1)	23.3 (3.1)	23.3 (3.0)	23.5 (3.1)	0.243
Current smoker, *n* (%)	620 (16.5)	334 (18.0)	180 (15.3)	106 (14.7)	0.060
Current alcohol drinker, *n* (%)	260 (6.9)	139 (7.5)	71 (6.1)	50 (7.0)	0.323
Education, *n* (%)					0.135
Middle school or lower	1134 (30.2)	538 (28.9)	387 (33.0)	209 (29.1)	
High school or professional college	1689 (45.0)	848 (45.6)	503 (42.9)	338 (47.0)	
University	928 (24.7)	473 (25.4)	283 (24.1)	172 (23.9)	
Income level, *n* (%)					0.598
Low (≤1500 ¥/month)	1033 (27.5)	527 (28.3)	324 (27.6)	182 (25.3)	
Middle (1501–3000 ¥/month)	2163 (57.7)	1054 (56.7)	678 (57.8)	431 (59.9)	
High (>3000 ¥/month)	555 (14.8)	278 (15.0)	171 (14.6)	106 (14.7)	
Physical activity, MET	41.1 (14.8)	40.6 (14.4)	41.7 (15.5)	41.4 (14.4)	0.098
Total energy intake, kcal/d	1765.0 (508.3)	1757.4 (501.2)	1741.3 (485.4)	1823.1 (557.0)	0.002
Type 2 diabetes, *n* (%)	197 (5.3)	93 (5.0)	69 (5.9)	35 (4.9)	0.501
Hypertension, *n* (%)	1072 (28.6)	522 (28.1)	338 (28.8)	212 (29.5)	0.760
Dyslipidemia, *n* (%)	1730 (46.1)	839 (45.1)	558 (47.6)	333 (46.3)	0.420
Metabolic syndrome, *n* (%)	584 (15.6)	280 (15.1)	198 (16.9)	106 (14.7)	0.321

Data are presented as mean (standard deviation) for continuous variables or *n* (%) for categorical variables.
Abbreviation: MET, metabolic equivalent of task.

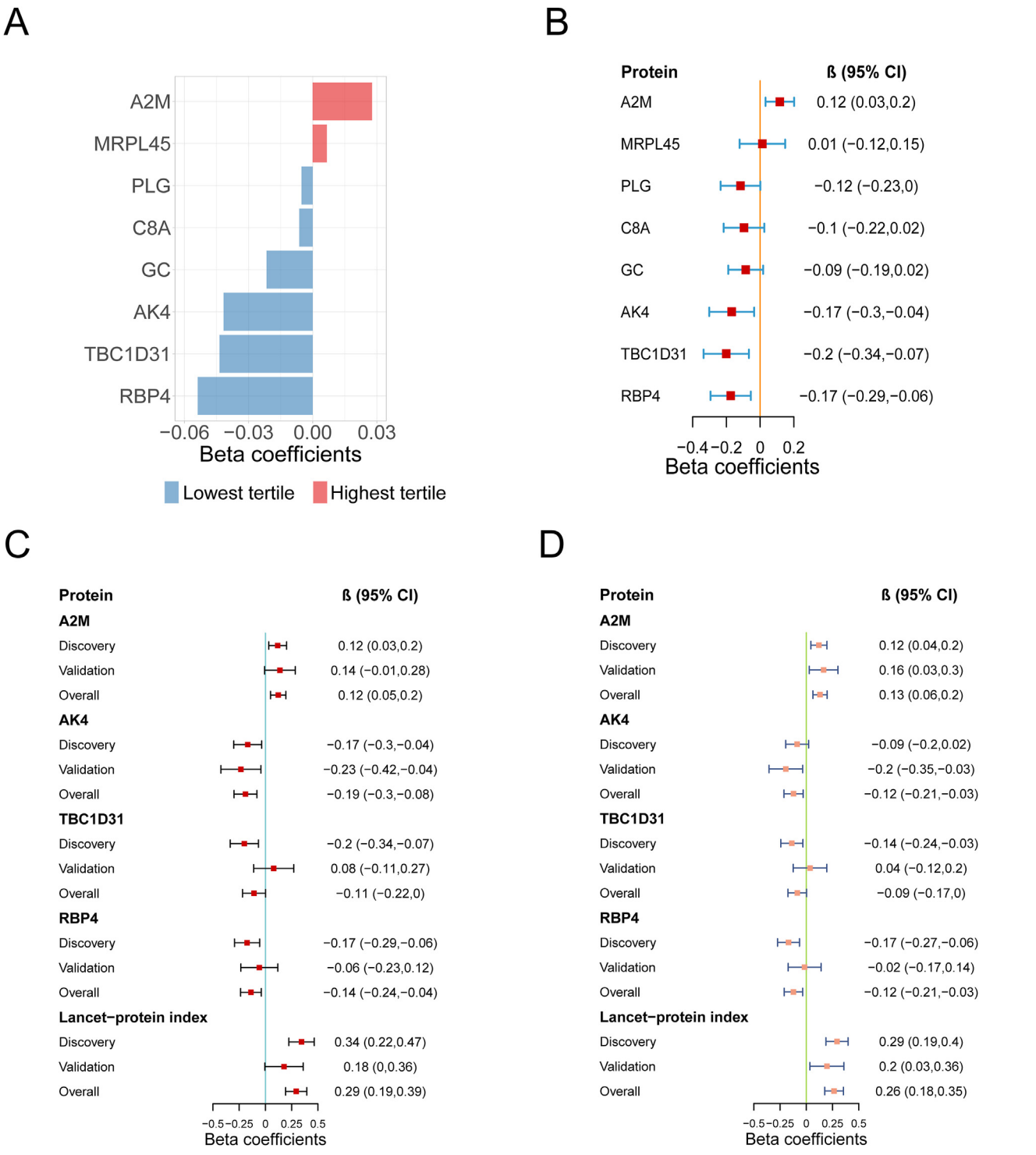
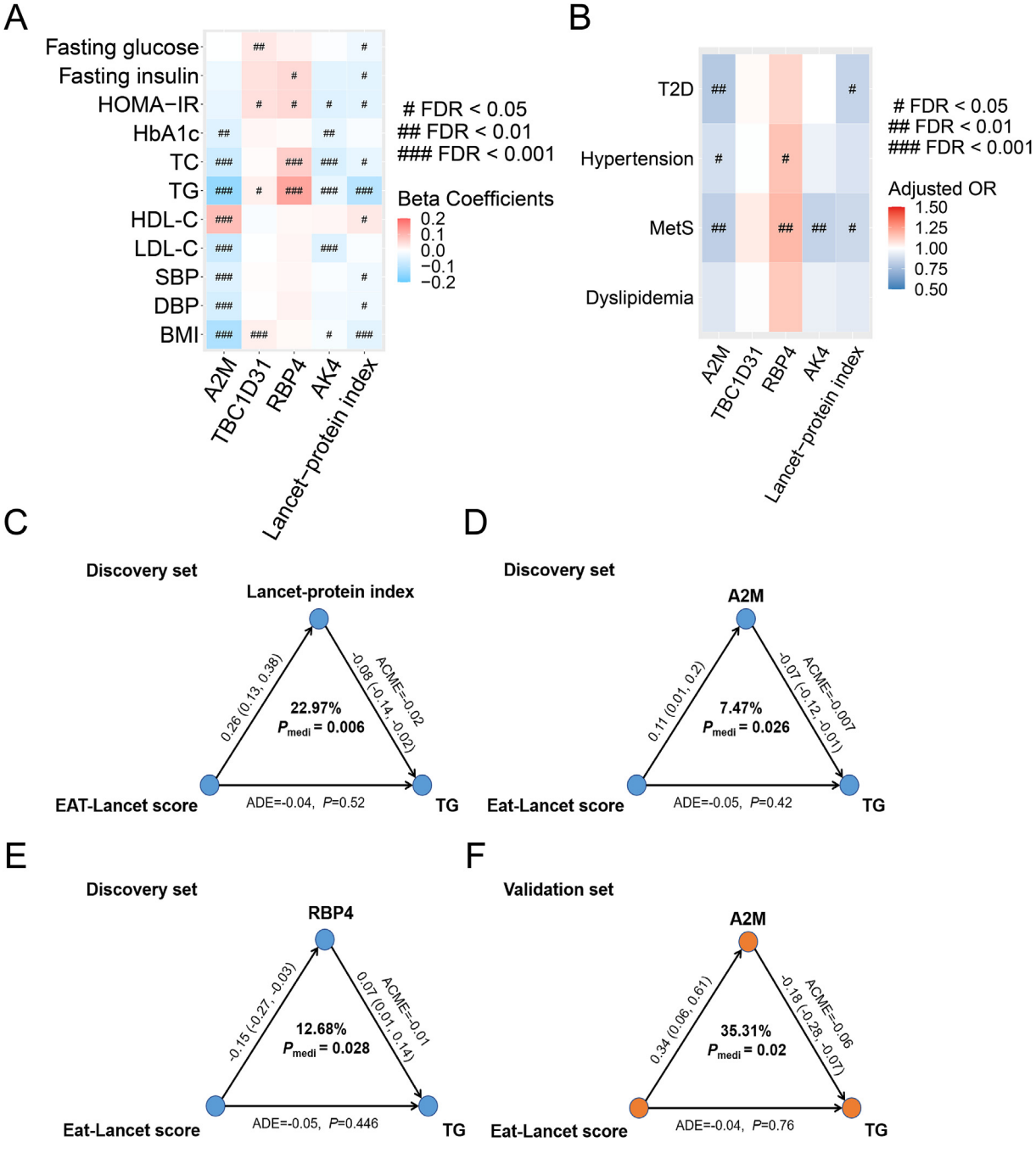


FIGURE 2. Prospective associations between the EAT-Lancet diet and serum proteins. (A) Associations between EAT-Lancet score and serum proteins based on LASSO regression in the discovery set. (B) Prospective associations between EAT-Lancet score and serum proteins based on multivariable linear regression, adjusted for potential confounders (baseline protein abundances were included as confounders) in the discovery set. In A–B, the analyses were based on 1060 participants in the discovery set ($N_{\text{highest tertile}} = 302$; $N_{\text{lowest tertile}} = 758$). (C) Results of meta-analysis for the prospective associations of EAT-Lancet score with identified serum proteins and Lancet-protein index from the discovery and validation sets. The associations were estimated by multivariable linear regression in the discovery and validation sets, adjusted for potential confounders (baseline protein abundances or Lancet-protein index were included as confounders). Fixed-effect meta-analysis was used to integrate the results from the discovery and validation sets. The analyses conducted in the validation set were based on 537 participants ($N_{\text{highest tertile}} = 153$; $N_{\text{lowest tertile}} = 384$). (D) Sensitivity analysis by linear mixed-effect model including all repeated-measured proteomics data for the associations of EAT-Lancet score with identified serum proteins and Lancet-protein index in the discovery and validation sets, adjusted for potential confounders (baseline protein abundances or Lancet-protein index were included as confounders). Fixed-effect meta-analysis was used to integrate the results from the discovery and validation sets. The analyses were based on 1734 ($N_{\text{highest tertile}} = 485$; $N_{\text{lowest tertile}} = 1{,}249$) and 739 ($N_{\text{highest tertile}} = 217$; $N_{\text{lowest tertile}} = 522$) serial measures of proteomics data in the discovery and validation set, respectively. All proteins and Lancet-protein index were transformed into *z*-scores. CI, confidence interval; FDR, false discovery rate; LASSO, least absolute shrinkage and selection operator.

diet (measured by EAT-Lancet score) was positively associated with beneficial components of the EAT-Lancet diet, including fruits, nuts, and added fats (higher unsaturated/saturated fat ratio) and negatively associated with animal-based foods, soy foods, and added sugars (Supplemental Table 2, Supplemental Figure 3). There were generally moderate correlations among the components of the EAT-Lancet diet (from −0.41 to 0.26; Supplemental Figure 3). EAT-Lancet score was modestly correlated with DASH score ($r = 0.16$, $P < 0.001$) and PDI ($r = 0.18$, $P < 0.001$) but not correlated with aMed score ($r = 0.03$, $P = 0.155$). From baseline to the 2014–2017 follow-up visit (583 participants with dietary information at both time points), most participants had stable adherence to the EAT-Lancet diet, with 73.8% having EAT-Lancet score changes within 1 point and 92.5% having EAT-Lancet score changes within 2 points. From baseline to the 2018–2019 follow-up visit (950 participants with dietary information at both time points), EAT-Lancet diet adherence was still stable, with 64.2% of EAT-Lancet score changes within 1 point and 89.7% of EAT-Lancet score changes within 2 points.

Adherence to the EAT-Lancet diet was prospectively associated with specific serum proteins

In the discovery set, we identified 4 proteins that were associated with the baseline EAT-Lancet score, with α-2-macroglobulin (A2M) level being significantly higher and retinol-binding protein 4 (RBP4), TBC1 domain family member 31 (TBC1D31) and adenylate kinase 4 (AK4) levels being significantly lower in the highest tertile of EAT-Lancet score compared with the lowest tertile of EAT-Lancet score (FDR < 0.05;



(caption on next page)

Figure 2A, B). Higher EAT-Lancet scores were prospectively associated with higher levels of Lancet-protein index (constructed based on the 4 identified EAT-Lancet diet-related proteins; see Methods) (β: 0.34; 95% confidence interval [CI]: 0.22, 0.47; $P < 0.001$) (Figure 2C). Linear trend analyses showed consistent results (Supplemental Figure 4). We did not find significant interactions of EAT-Lancet diet adherence with age (≥60 y compared with <60 y) or sex on the levels of the identified proteins or the Lancet-protein index (all $FDR_{interaction} > 0.05$).

We further replicated the above-identified associations in the validation set (Figure 2C) and found significant associations between the baseline EAT-Lancet score (highest compared with lowest tertile) and follow-up levels of AK4 (β: −0.23; 95% CI: −0.42, −0.04: $P = 0.017$) and A2M (β: 0.14; 95% CI: −0.01, 0.28; $P = 0.063$) and the Lancet-protein index (β: 0.18; 95% CI: 0, 0.36; $P = 0.056$). The results of meta-analysis from the discovery and validation sets showed that adherence to the EAT-Lancet diet was significantly associated with higher levels of A2M (pooled β: 0.12; 95% CI: 0.05, 0.20) and Lancet-protein index (pooled β: 0.29; 95% CI: 0.19, 0.39) and lower levels of AK4 (pooled β: −0.19; 95% CI: −0.30, −0.08) (Figure 2C). Linear trend tests showed similar results to the above associations (Supplementary Figure 4).

Sensitivity analysis by linear mixed-effect model including all repeated-measured proteomics data showed similar results, and the associations of baseline EAT-Lancet score (highest compared with lowest tertile) with future AK4 (β: −0.2; 95% CI: −0.35, −0.03; $P = 0.019$) and A2M levels (β: 0.16; 95% CI: 0.03, 0.3; $P = 0.021$) and Lancet-protein index (β: 0.2; 95% CI: 0.03, 0.36; $P = 0.019$) were significant in the validation set (Figure 2D). Similar results were found in linear trend analyses (Supplementary Figure 4).

Serum proteomic biomarkers of the EAT-Lancet diet were associated with cardiometabolic health

A2M, which was positively associated with the EAT-Lancet diet, was beneficially associated with many cardiometabolic risk factors, such as HbA1c, TC, TG, HDL cholesterol, LDL cholesterol, SBP, DBP, and BMI (FDR < 0.05; Figure 3A and Supplemental Figure 5A, B). Similarly, both TBC1D31 and RBP4 were associated with several cardiometabolic risk factors (FDR < 0.05; Figure 3A and Supplementary Figure 5A, B). Additionally, the Lancet-protein index was positively associated with HDL cholesterol and negatively associated with fasting glucose, fasting insulin, HOMA-IR, TC, TG, SBP, DBP, and BMI (FDR < 0.05; Figure 3A and Supplementary Figure 5C). We did not observe significant interactions of EAT-Lancet diet-related proteins or Lancet-protein index with age (≥60 y compared with <60 y) or sex on any cardiometabolic traits (all $FDR_{interaction} > 0.05$).

For incident cardiometabolic diseases, 1 SD increase in A2M was associated with 20% lower risk of T2D (pooled odds ratio [OR]: 0.8; 95% CI: 0.69, 0.92), 12% lower risk of hypertension (pooled OR: 0.88; 95% CI: 0.79, 0.98), and 18% lower risk of MetS (pooled OR: 0.82; 95% CI: 0.73, 0.92) (FDR < 0.05; Figure 3B and Supplemental Figure 6A). One SD increase in RBP4 was associated with 15% higher risk of hypertension (pooled OR: 1.15; 95% CI: 1.04, 1.28) and 18% higher risk of MetS (pooled OR: 1.18; 95% CI: 1.05, 1.31) (FDR < 0.05; Figure 3B and Supplementary Figure 6A). Finally, 1 SD increase in Lancet-protein index was associated with 16% lower risk of T2D (pooled OR: 0.84; 95% CI: 0.74, 0.96) and 14% lower risk of MetS (pooled OR: 0.86; 95% CI: 0.77, 0.96) (FDR < 0.05; Figure 3B and Supplementary Figure 6B). No significant interactions were observed for EAT-Lancet diet-related proteins or Lancet-protein index with age (≥60 y compared with <60 y) or sex on any cardiometabolic diseases (all $FDR_{interaction} > 0.05$).

Serum proteins mediated the association between the EAT-Lancet diet and blood TG

The results of mediation analysis showed that the Lancet-protein index (22.97%, $P = 0.006$, Figure 3C in the discovery set), A2M level (7.47%, $P = 0.026$, Figure 3D in the discovery set; 35.31%, $P = 0.02$, Figure 3F in the validation set), and RBP4 level (12.68%, $P = 0.028$, Figure 3E in the discovery set) could mediate the association between baseline EAT-Lancet score and future TG level.

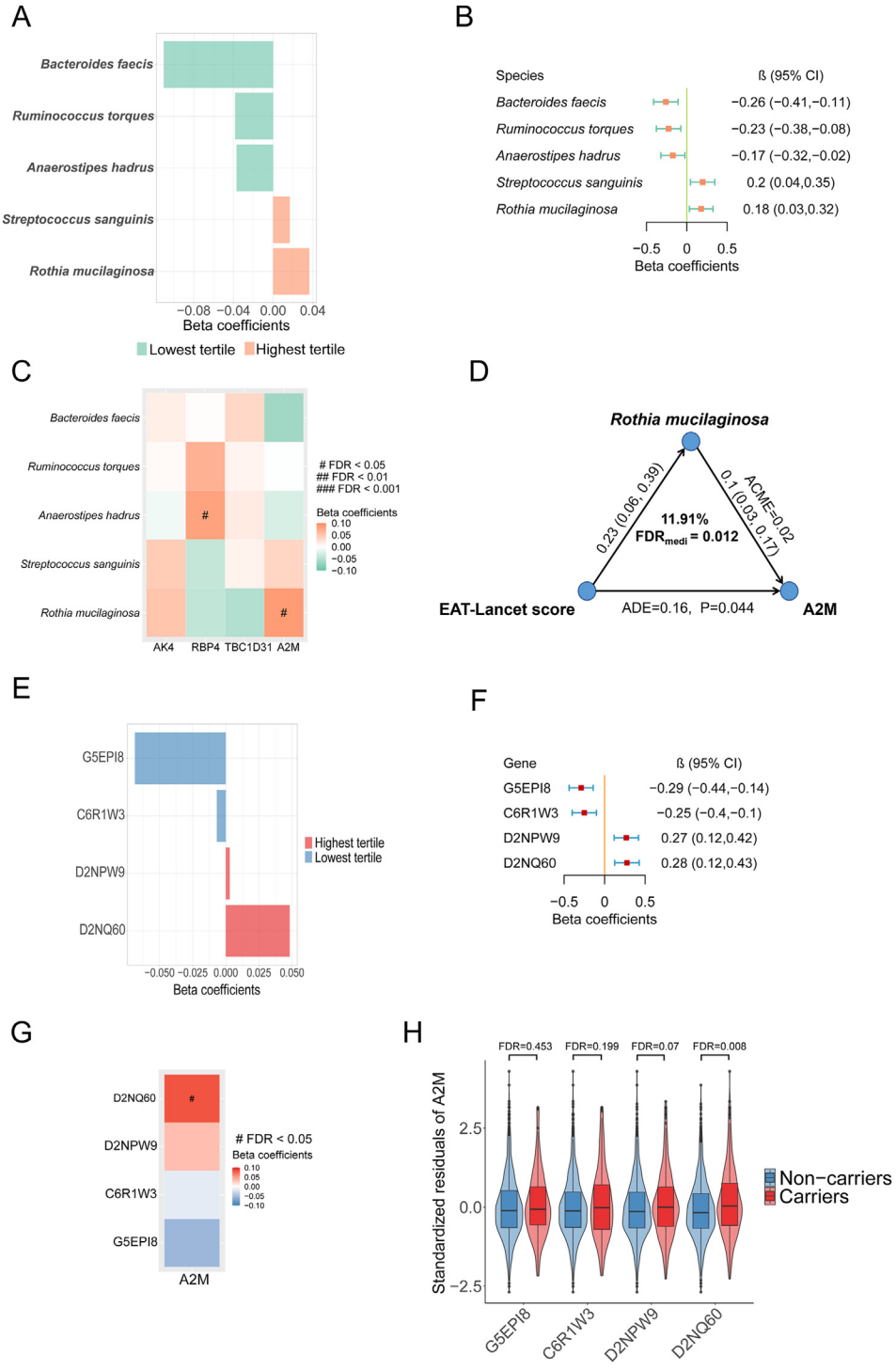
Potential role of gut microbiota in linking the EAT-Lancet diet and its related serum proteins

We identified *Rothia mucilaginosa* and *Streptococcus sanguinis* as positively associated with the baseline EAT-Lancet score and *Bacteroides faecis*, *Ruminococcus torques*, and *Anaerostipes hadrus* as negatively associated with the baseline EAT-Lancet score (FDR < 0.05; Figure 4A, B). We did not find significant interactions of EAT-Lancet score with age (≥60 y compared with <60 y) or sex on selected gut microbes (all $FDR_{interaction} > 0.05$). Sensitivity analysis by linear mixed-effect model also showed that the EAT-Lancet score was

FIGURE 3. Associations of EAT-Lancet diet-related proteins with cardiometabolic health. (A) Results of meta-analysis for the associations of EAT-Lancet diet-related proteins and Lancet-protein index with cardiometabolic risk factors from the discovery and validation sets. Associations were estimated by linear mixed-effect models, adjusted for potential confounders. The concentrations of fasting glucose, fasting insulin, HOMA-IR, and TG were log-transformed. The analyses in the discovery and validation sets were based on 4459 and 2788 serial measures of fasting glucose, 1699 and 478 serial measures of fasting insulin, 1698 and 478 serial measures of HOMA-IR, 2740 and 1046 serial measures of HbA1c, 4137 and 2305 serial measures of total cholesterol, 4460 and 2744 serial measures of TG, 4460 and 2614 serial measures of HDL cholesterol, 4460 and 2614 serial measures of LDL cholesterol, 4530 and 2843 serial measures of SBP, 4530 and 2843 serial measures of DBP, and 4531 and 2843 serial measures of BMI, respectively. (B) Results of meta-analysis for the prospective associations of baseline EAT-Lancet diet-related proteins and Lancet-protein index with incident cardiometabolic diseases from the discovery and validation sets. The associations were estimated by logistic regression, adjusted for potential confounders. Participants with cardiometabolic diseases diagnosed at baseline were excluded in this analysis. Fixed-effect meta-analysis was used to integrate the results from the discovery and validation sets. The analyses in the discovery and validation sets were based on 1670 (incident cases: 331) and 1494 (incident cases: 27) participants for T2D, 1234 (incident cases: 403) and 1133 (incident cases: 225) participants for hypertension, 1458 (incident cases: 428) and 1352 (incident cases: 113) participants for MetS, and 900 (incident cases: 523) and 887 (incident cases: 327) participants for dyslipidemia. (C) Mediation effect of the Lancet-protein index on the prospective association between EAT-Lancet score and TG in the discovery set ($N = 855$). (D) Mediation effect of A2M on the prospective association between EAT-Lancet score and TG in the discovery set ($N = 855$). (E) Mediation effect of RBP4 on the prospective association between EAT-Lancet score and TG in the discovery set ($N = 855$). (F) Mediation effect of A2M on the prospective association between EAT-Lancet score and TG in the validation set ($N = 186$). The potential mediation effect was detected by mediation analysis. The abundance of EAT-Lancet diet-related proteins, Lancet-protein index, and cardiometabolic risk factors were standardized into *z*-scores. ACME, average causal mediation effects; ADE, average direct effects; DBP, diastolic blood pressure; FDR, false discovery rate; HbA1c, glycated hemoglobin; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostasis model assessment of insulin resistance; LDL-C, low-density lipoprotein cholesterol; MetS, metabolic syndrome; OR, odds ratio; SBP, systolic blood pressure; T2D, type 2 diabetes; TC, total cholesterol; TG, triglycerides.

positively associated with *R. mucilaginosa* (β: 0.15; 95% CI: 0.02, 0.28; $P = 0.027$) and negatively associated with *B. faecis* (β: −0.17 95% CI: −0.31, −0.03], P=0.018; Supplemental Figure 7). The prevalences of *B. faecis*, *A. hadrus*, *S. sanguinis*, and *R. mucilaginosa* were significantly different between the highest and lowest tertiles of EAT-Lancet score ($P < 0.05$; Supplemental Figure 8).

We also found that *R. mucilaginosa* was positively associated with A2M level (β: 0.09; 95% CI: 0.03, 0.15), and *A. hadrus* was positively associated with RBP4 level (β: 0.08; 95% CI: 0.03, 0.14) (FDR < 0.05; Figure 4C). The distributions of A2M and RBP4 were significantly higher in carriers of *R. mucilaginosa* and *A. hadrus* than in noncarriers of these 2 species, respectively ($P < 0.05$; Supplemental Figure 9). For EAT-Lancet diet–*R. mucilaginosa*–A2M and EAT-Lancet diet–*A. hadrus*–RBP4 associations, the results of mediation analysis showed that *R. mucilaginosa* may mediate the association between the EAT-Lancet score and A2M level (11.91%, FDR = 0.012; Figure 4D).



(caption on next page)

We then explored specific genes (2312 in total) of the microbe *R. mucilaginosa* that may play roles in the association between the EAT-Lancet diet and serum A2M. Adherence to the EAT-Lancet diet was associated with lower expression of *G5EPI8* (β: −0.29; 95% CI: −0.44, −0.14) and *C6R1W3* (β: −0.25; 95% CI: −0.4, −0.1), and higher expression of *D2NPW9* (β: 0.27; 95% CI: 0.12, 0.42) and *D2NQ60* (β: 0.28; 95% CI: 0.12, 0.43) in this microbe (FDR < 0.05; Figure 4E, F). Prevalence of *D2NPW9* and *D2NQ60* was higher in participants in the highest tertile of the EAT-Lancet score than those in the lowest tertile of the EAT-Lancet score ($P < 0.001$; Supplemental Figure 10). Among them, *D2NQ60* was positively associated with serum protein A2M (FDR < 0.05; Figure 4G), and A2M levels were significantly higher in carriers of *D2NQ60* than in the noncarriers of *D2NQ60* (FDR = 0.008; Figure 4H). Thus, these data suggest that gene *D2NQ60* in *R. mucilaginosa* may play an important role in linking the association between the EAT-Lancet diet and serum protein A2M level.

Discussion

In this large-scale prospective cohort study, we identified serum proteomic biomarkers of the EAT-Lancet diet that were associated with cardiometabolic diseases and risk factors and may mediate the association between the EAT-Lancet diet and TG level. In addition, we identified gut microbial biomarkers of the EAT-Lancet diet and discovered a potential gut microbiota–blood protein interplay (EAT-Lancet diet→*R. mucilaginosa*→A2M association) underlying the EAT-Lancet diet–cardiometabolic health association.

A healthy dietary pattern is essential for human health. We showed that the EAT-Lancet score was only moderately correlated with standard diet quality scores (e.g., DASH score, PDI, and aMed scores; $r <$ 0.2). The EAT-Lancet score differs from other standard diet scores in that it considers the dimensions of both environmental sustainability and human health in its scoring system. Thus, the EAT-Lancet diet emphasizes the intake of whole grains, vegetables, fruits, nuts, and unsaturated oils, consists of a low to moderate intake of fish and seafood and poultry, and considerably limits the intake of red meats, added sugars, and tubers and starchy vegetables [1]. To our knowledge, our study was the first to investigate multiomics biomarkers (serum proteins and gut microbes) of the EAT-Lancet diet and their relationship with cardiometabolic health. Unlike previous studies identifying biomarkers of dietary patterns based on cross-sectional data [22,52,53], we explored the prospective associations between the EAT-Lancet diet and serum proteome, providing high-level causal evidence for EAT-Lancet diet-protein associations.

The protein A2M, as a protease inhibitor and cytokine transporter, could inhibit inflammatory cytokines [54]. A2M might play a mediating role in the association between nutritional status and disease prevention in the Thai population [55]. Another study found that A2M was downregulated in rats fed a high-fat diet (60% kcal fat, 20% kcal carbohydrate, and 20% kcal protein) compared with rats fed a normal feed pellets diet [56,57]. These studies provided preliminary evidence for the positive association between a healthy diet and A2M, consistent with our results. As a broad-spectrum protease inhibitor, A2M can remove 500 types of toxins (proteases) that are the main cause of many complex diseases and plays a role in anti-aging mechanisms [58,59]. Studies have also shown that A2M can inhibit inflammatory pathways involving IL-1β and NF-κB [54,60]. Our study supported evidence that A2M is beneficially associated with human cardiometabolic health and may mediate the association of the EAT-Lancet diet with blood lipids.

RBP4, an adipokine, has been reported to be positively associated with obesity, insulin resistance, T2D, MetS, and dyslipidemia [61–63], consistent with our findings. Several other studies demonstrated that diet-induced weight loss, bariatric surgery, or exercise decrease RBP4 levels [64–66]. Yang et al. [61] suggested that lowering RBP4 levels could be a new strategy for treating T2D. Our study showed that the EAT-Lancet diet adherence was associated with a lower level of RBP4, which may be a promising dietary intervention target for improving cardiometabolic health.

The gut microbiome exerts its role in the physiological effects of diet [67]. We showed that *R. mucilaginosa* may mediate the association between the EAT-Lancet diet and A2M level. *R. mucilaginosa*, an anti-inflammatory bacterium, is inversely associated with proinflammatory makers IL-8 and IL-1β by inhibiting the NF-κB pathway [68], the same pathway through which A2M plays its anti-inflammatory role [54]. Therefore, we speculate that *R. mucilaginosa* and A2M may jointly modulate the association between the EAT-Lancet diet and cardiometabolic health through inflammatory pathways. *D2NQ60*, a gene in *R. mucilaginosa*, was positively associated with both the EAT-Lancet diet and A2M. D2NQ60, namely superoxide dismutase, may regulate oxidative stress, lipid metabolism, and inflammation [69–72]. The microbial gene *D2NQ60* in *R. mucilaginosa* may serve as a target to modulate A2M and improve cardiometabolic health.

Our study has some limitations. First, as this study was had an observational study design, we cannot fully avoid the influence of residual confounders on our results, although we have adjusted for several important confounding factors in our analyses. Second, because matched proteomics and gut microbiome data are rare in the field at this stage, we did not have other datasets to validate the *R. mucilaginosa*–A2M association. However, our proposed gut microbiota–blood protein interplay

FIGURE 4. Role of gut microbiota in linking the associations between the EAT-Lancet diet and serum proteins. (A) Associations between EAT-Lancet score and gut microbes based on LASSO regression. (B) Associations between EAT-Lancet score and gut microbes based on multivariable linear regression, adjusted for potential confounders. In A–B, the analyses were based on 821 participants in the discovery set ($N_{\text{highest tertile}} = 238$; $N_{\text{lowest tertile}} = 583$). (C) Cross-sectional associations between EAT-Lancet diet-related gut microbes and serum proteins based on multivariable linear regression, adjusted for potential confounders ($N = 1094$). (D) Mediation effect of *Rothia mucilaginosa* on the association between EAT-Lancet score and serum protein A2M ($N = 750$). (E) Associations between EAT-Lancet score and genes in *R. mucilaginosa* based on LASSO regression. (F) Associations between EAT-Lancet score and genes in *R. mucilaginosa* based on multivariable linear regression, adjusted for potential confounders. In E–F, the analyses were based on 821 participants in the discovery set ($N_{\text{highest tertile}} = 238$; $N_{\text{lowest tertile}} = 583$). (G) Cross-sectional associations between EAT-Lancet diet-related genes and serum protein A2M based on multivariable linear regression, adjusted for potential confounders ($N = 1094$). (H) Distributions of standardized residuals of serum A2M between the carriers and noncarriers of identified microbial genes of *R. mucilaginosa* ($N = 1094$). The residuals of serum A2M were obtained using multivariable linear regression adjusted for potential confounders. Differences in standardized residuals (z-scores) of serum A2M between the carriers and noncarriers of genes were tested by Wilcoxon rank-sum test. In the above analyses, the relative abundances of gut microbiota data and genome data were centered log-ratio transformed. The data of gut microbes, genes, and proteins were transformed into z-scores. ACME, average causal mediation effects; ADE, average direct effects; CI, confidence interval; FDR, false discovery rate; LASSO, least absolute shrinkage and selection operator.

provides an important mechanistic hypothesis underlying the EAT-Lancet diet–cardiometabolic health association, warranting future validations in other large prospective cohort studies or experimental studies. Third, our cohort only surveyed antibiotics use within 2 wk at the time of stool sample collection, and it may take longer to reconstitute the gut microbiome after antibiotics treatment. Meanwhile, our cohort only included middle-aged and elderly Chinese adults living in urban Guangzhou, China. The generalizability of our findings on serum proteins and gut microbes should be further validated in external large cohorts of racially and geographically diverse populations surveyed for long-term antibiotics use (>2 mo). Finally, given the dietary assessment in this study was based on self-reported FFQs with a relatively small number of food items, which is prone to recall and social desirability bias, the EAT-Lancet score constructed based on cutoff points of absolute amounts of dietary components may have non-negligible measurement errors.

In summary, we identified serum proteins underlying the beneficial association of the EAT-Lancet diet with cardiometabolic health based on a longitudinal human cohort. Additionally, we identified gut microbial biomarkers of the EAT-Lancet diet, and proposed a potential gut microbiota–blood protein interplay (EAT-Lancet diet→*R. mucilaginosa*→A2M association) that may link the EAT-Lancet diet and cardiometabolic health. The identified biomarkers may serve as dietary intervention targets for reducing the risk of cardiometabolic disorders in the future.

Acknowledgments

We thank all study participants of the Guangzhou Nutrition and Health Study. We thank Westlake University High-Performance Computing Center for assistance in data storage and computation. We thank the support from Westlake Education Foundation and Westlake Intelligent Biomarker Discovery Lab at the Westlake Laboratory of Life Sciences and Biomedicine.

Author contributions

The authors' responsibilities were as follows—J-SZ, Y-mC, TG, KD: study conceptualization and design; KD, LS, JT: data analysis; ZX, BL, XC, WH, LY, YX, TS: data collection; HZ, FX, ZM, XL, MS: data curation; KD, J-SZ: wrote the manuscript; LS, YF, ZJ, CX, WG: contributed to the discussion; J-SZ, Y-mC, TG, KD: writing, reviewing, and editing the manuscript; J-SZ, Y-mC, TG: are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis; and all authors: read, revised, and approved the final manuscript.

Conflict of interest

TG is a shareholder of Westlake Omics Inc. All other authors report no conflicts of interest.

Funding

This study was funded by the National Natural Science Foundation of China (82204161, 82073546, 92374112, 82073529, 82103826), Zhejiang Provincial Natural Science Foundation of China (LQ21H260002), Research Program (No. 202208012) of Westlake Laboratory of Life Sciences and Biomedicine, Zhejiang Ten-thousand Talents Program (2019R52039), and the 5010 Program for Clinical Research (2007032) of the Sun Yat-sen University (Guangzhou, China).

Data availability

The raw serum proteomics data are available in the iProX database (https://www.iprox.cn/page/home.html) at accession numbers PXD039236, PXD039231, and PXD038253. The raw metagenomic sequencing data are available in the Genome Sequence Archive (GSA) (https://ngdc.cncb.ac.cn/gsa/) at accession numbers CRA008796 and CRA010223. Other datasets generated and/or analyzed during this study are available upon reasonable request by bona fide researchers for specified scientific purposes via contacting the corresponding authors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajcnut.2024.10.011.

References

[1] W. Willett, J. Rockström, B. Loken, M. Springmann, T. Lang, S. Vermeulen, et al., Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems, Lancet 393 (10170) (2019) 447–492, https://doi.org/10.1016/S0140-6736(18)31788-4.

[2] J.E. Laine, I. Huybrechts, M.J. Gunter, P. Ferrari, E. Weiderpass, K. Tsilidis, et al., Co-benefits from sustainable dietary shifts for population and environmental health: an assessment from a large European cohort study, Lancet Planet. Health 5 (11) (2021) e786–e796, https://doi.org/10.1016/S2542-5196(21)00250-3.

[3] M. Springmann, L. Spajic, M.A. Clark, J. Poore, A. Herforth, P. Webb, et al., The healthiness and sustainability of national and global food based dietary guidelines: modelling study, BMJ 370 (2020) m2322, https://doi.org/10.1136/bmj.m2322.

[4] A. Knuppel, K. Papier, T.J. Key, R.C. Travis, EAT-Lancet score and major health outcomes: the EPIC-Oxford study, Lancet 394 (10194) (2019) 213–214, https://doi.org/10.1016/s0140-6736(19)31236-x.

[5] S. Zhang, J. Dukuzimana, A. Stubbendorff, U. Ericson, Y. Borné, E. Sonestedt, Adherence to the EAT-Lancet diet and risk of coronary events in the Malmö Diet and Cancer cohort study, Am. J. Clin. Nutr. 117 (5) (2023) 903–909, https://doi.org/10.1016/j.ajcnut.2023.02.018.

[6] S. Zhang, A. Stubbendorff, U. Ericson, P. Wändell, K. Niu, L. Qi, et al., The EAT-Lancet diet, genetic susceptibility and risk of atrial fibrillation in a population-based cohort, BMC Med 21 (1) (2023) 280, https://doi.org/10.1186/s12916-023-02985-6.

[7] C. Xu, Z. Cao, H. Yang, Y. Hou, X. Wang, Y. Wang, Association between the EAT-Lancet diet pattern and risk of type 2 diabetes: a prospective cohort study, Front. Nutr. 8 (2021) 784018, https://doi.org/10.3389/fnut.2021.784018.

[8] F. Langmann, D.B. Ibsen, A. Tjønneland, A. Olsen, K. Overvad, C.C. Dahm, Adherence to the EAT-Lancet diet is associated with a lower risk of type 2 diabetes: the Danish Diet, Cancer and Health cohort, Eur. J. Nutr. 62 (3) (2023) 1493–1502, https://doi.org/10.1007/s00394-023-03090-3.

[9] S. Zhang, A. Stubbendorff, K. Olsson, U. Ericson, K. Niu, L. Qi, et al., Adherence to the EAT-Lancet diet, genetic susceptibility, and risk of type 2 diabetes in Swedish adults, Metabolism 141 (2023) 155401, https://doi.org/10.1016/j.metabol.2023.155401.

[10] D.B. Ibsen, A.H. Christiansen, A. Olsen, A. Tjønneland, K. Overvad, A. Wolk, et al., Adherence to the EAT-Lancet diet and risk of stroke and stroke subtypes: a cohort study, Stroke 53 (1) (2022) 154–163, https://doi.org/10.1161/STROKEAHA.121.036738.

[11] L.T. Cacau, I.M. Benseñor, A.C. Goulart, L.O. Cardoso, P.A. Lotufo, L.A. Moreno, et al., Adherence to the Planetary Health Diet Index and obesity indicators in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil), Nutrients 13 (11) (2021) 3691, https://doi.org/10.3390/nu13113691.

[12] Y. Wang, M. Fan, H. Qian, H. Ying, Y. Li, L. Wang, Whole grains-derived functional ingredients against hyperglycemia: targeting hepatic glucose metabolism, Crit. Rev. Food Sci. Nutr. 64 (20) (2024) 7268–7289, https://doi.org/10.1080/10408398.2023.2183382.

[13] M. Shuai, L.S.Y. Zuo, Z. Miao, W. Gou, F. Xu, Z. Jiang, et al., Multi-omics analyses reveal relationships among dairy consumption, gut microbiota and cardiometabolic health, EBioMedicine 66 (2021) 103284, https://doi.org/10.1016/j.ebiom.2021.103284.

[14] J.A. Buffa, K.A. Romano, M.F. Copeland, D.B. Cody, W. Zhu, R. Galvez, et al., The microbial *gbu* gene cluster links cardiovascular disease risk associated with red meat consumption to microbiota L-carnitine catabolism, Nat. Microbiol. 7 (1) (2022) 73–86, https://doi.org/10.1038/s41564-021-01010-x.

[15] J. Wang, D. Li, L.J. Dangott, G. Wu, Proteomics and its role in nutrition research, J. Nutr. 136 (7) (2006) 1759–1762, https://doi.org/10.1093/jn/136.7.1759.
[16] V. Ganesh, N.S. Hettiarachchy, Nutriproteomics: a promising tool to link diet and diseases in nutritional research, Biochim. Biophys. Acta. 1824 (10) (2012) 1107–1117, https://doi.org/10.1016/j.bbapap.2012.06.006.
[17] M. Kussmann, A. Panchaud, M. Affolter, Proteomics in nutrition: status quo and outlook for biomarkers and bioactives, J. Proteome Res. 9 (10) (2010) 4876–4887, https://doi.org/10.1021/pr1004339.
[18] D. Fuchs, I. Winkelmann, I.T. Johnson, E. Mariman, U. Wenzel, H. Daniel, Proteomics in nutrition research: principles, technologies and applications, Br. J. Nutr. 94 (3) (2005) 302–314, https://doi.org/10.1079/bjn20051458.
[19] M.E. Walker, R.J. Song, X. Xu, R.E. Gerszten, D. Ngo, C.B. Clish, et al., Proteomic and metabolomic correlates of healthy dietary patterns: the Framingham Heart Study, Nutrients 12 (5) (2020) 1476, https://doi.org/10.3390/nu12051476.
[20] B. García-Bailo, D.R. Brenner, D. Nielsen, H.J. Lee, D. Domanski, M. Kuzyk, et al., Dietary patterns and ethnicity are associated with distinct plasma proteomic groups, Am. J. Clin. Nutr. 95 (2) (2012) 352–361, https://doi.org/10.3945/ajcn.111.022657.
[21] E. Warensjö Lemming, L. Byberg, K. Stattin, S. Ahmad, L. Lind, S. Elmståhl, et al., Dietary pattern specific protein biomarkers for cardiovascular disease: a cross-sectional study in 2 independent cohorts, J. Am. Heart Assoc. 8 (11) (2019) e011860, https://doi.org/10.1161/JAHA.118.011860.
[22] S. Zhang, I. Marken, A. Stubbendorff, U. Ericson, L. Qi, E. Sonestedt, et al., The EAT-Lancet diet index, plasma proteins, and risk of heart failure in a population-based cohort, JACC Heart Fail 12 (7) (2024) 1197–1208, https://doi.org/10.1016/j.jchf.2024.02.017.
[23] Z. Jiang, T.Y. Sun, Y. He, W. Gou, L.S.Y. Zuo, Y. Fu, et al., Dietary fruit and vegetable intake, gut microbiota, and type 2 diabetes: results from two large human cohort studies, BMC Med 18 (1) (2020) 371, https://doi.org/10.1186/s12916-020-01842-0.
[24] W. Ma, L.H. Nguyen, M. Song, D.D. Wang, E.A. Franzosa, Y. Cao, et al., Dietary fiber intake, the gut microbiome, and chronic systemic inflammation in a cohort of adult men, Genome Med 13 (1) (2021) 102, https://doi.org/10.1186/s13073-021-00921-y.
[25] J. Rehner, G.P. Schmartz, T. Kramer, V. Keller, A. Keller, S.L. Becker, The effect of a planetary health diet on the human gut microbiome: a descriptive analysis, Nutrients 15 (8) (2023) 1924, https://doi.org/10.3390/nu15081924.
[26] D.V. Zhernakova, T.H. Le, A. Kurilshikov, B. Atanasovska, M.J. Bonder, S. Sanna, et al., Individual variations in cardiovascular-disease-related protein levels are driven by genetics and gut microbiome, Nat. Genet. 50 (11) (2018) 1524–1532, https://doi.org/10.1038/s41588-018-0224-7.
[27] W. Gou, Y. Fu, L. Yue, G.D. Chen, X. Cai, M. Shuai, et al., Gut microbiota, inflammation, and molecular signatures of host response to infection, J. Genet. Genomics. 48 (9) (2021) 792–802, https://doi.org/10.1016/j.jgg.2021.04.002.
[28] X. Cai, Z. Xue, F.F. Zeng, J. Tang, L. Yue, B. Wang, et al., Population serum proteomics uncovers a prognostic protein classifier for metabolic syndrome, Cell Rep. Med. 4 (9) (2023) 101172, https://doi.org/10.1016/j.xcrm.2023.101172.
[29] C.X. Zhang, S.C. Ho, Validity and reproducibility of a food frequency questionnaire among Chinese women in Guangdong province, Asia Pac, J. Clin. Nutr. 18 (2) (2009) 240–250.
[30] Z. Miao, J.S. Lin, Y. Mao, G.D. Chen, F.F. Zeng, H.L. Dong, et al., Erythrocyte n-6 polyunsaturated fatty acids, gut microbiota, and incident type 2 diabetes: a prospective cohort study, Diabetes Care 43 (10) (2020) 2435–2443, https://doi.org/10.2337/dc20-0631.
[31] A.J.G. Hanley, K. Williams, M.P. Stern, S.M. Haffner, Homeostasis model assessment of insulin resistance in relation to the incidence of cardiovascular disease: the San Antonio Heart Study, Diabetes Care 25 (7) (2002) 1177–1184, https://doi.org/10.2337/diacare.25.7.1177.
[32] American Diabetes Association, Diagnosis and classification of diabetes mellitus, Diabetes Care 37 (Suppl 1) (2014) S81–S90, https://doi.org/10.2337/dc14-S081.
[33] J. Chalmers, S. MacMahon, G. Mancia, J. Whitworth, L. Beilin, L. Hansson, et al., 1999 World Health Organization-International Society of Hypertension Guidelines for the management of hypertension, Guidelines sub-committee of the World Health Organization, Clin. Exp. Hypertens. 21 (5–6) (1999) 1009–1060, https://doi.org/10.3109/10641969909061028.
[34] Joint committee for guideline revision, 2016 Chinese guidelines for the management of dyslipidemia in adults, J. Geriatr. Cardiol. 15 (1) (2018) 1–29, https://doi.org/10.11909/j.issn.1671-5411.2018.01.011.
[35] B. Liu, J. Woo, N. Tang, K. Ng, R. Ip, A. Yu, Assessment of total energy expenditure in a Chinese population by a physical activity questionnaire: examination of validity, Int. J. Food Sci. Nutr. 52 (3) (2001) 269–282, https://doi.org/10.1080/09637480120044138.
[36] Y. Zhang, X. Cai, W. Ge, D. Wang, G. Zhu, L. Qian, et al., Potential use of serum proteomics for monitoring COVID-19 progression to complement RT-PCR detection, J. Proteome Res. 21 (1) (2022) 90–100, https://doi.org/10.1021/acs.jproteome.1c00525.
[37] V. Demichev, C.B. Messner, S.I. Vernardis, K.S. Lilley, M. Ralser, DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput, Nat. Methods. 17 (1) (2020) 41–44, https://doi.org/10.1038/s41592-019-0638-x.
[38] K. Deng, M. Shuai, Z. Zhang, Z. Jiang, Y. Fu, L. Shen, et al., Temporal relationship among adiposity, gut microbiota, and insulin resistance in a longitudinal human cohort, BMC Med 20 (1) (2022) 171, https://doi.org/10.1186/s12916-022-02376-3.
[39] M. Shuai, Y. Fu, H.L. Zhong, W. Gou, Z. Jiang, Y. Liang, et al., Mapping the human gut mycobiome in middle-aged and elderly adults: multiomics insights and implications for host metabolic health, Gut 71 (9) (2022) 1812–1820, https://doi.org/10.1136/gutjnl-2021-326298.
[40] M. Shuai, G. Zhang, F.F. Zeng, Y. Fu, X. Liang, L. Yuan, et al., Human gut antibiotic resistome and progression of diabetes, Adv. Sci. (Weinh) 9 (11) (2022) e2104965, https://doi.org/10.1002/advs.202104965.
[41] S. Abubucker, N. Segata, J. Goll, A.M. Schubert, J. Izard, B.L. Cantarel, et al., Metabolic reconstruction for metagenomic data and its application to the human microbiome, PLoS Comput. Biol. 8 (6) (2012) e1002358, https://doi.org/10.1371/journal.pcbi.1002358.
[42] B.E. Suzek, Y. Wang, H. Huang, P.B. McGarvey, C.H. Wu, UniProt Consortium, UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches, Bioinformatics 31 (6) (2015) 926–932, https://doi.org/10.1093/bioinformatics/btu739.
[43] P.B. Mellen, S.K. Gao, M.Z. Vitolins, D.C. Goff Jr., Deteriorating dietary habits among adults with hypertension: DASH dietary accordance, NHANES 1988-1994 and 1999-2004, Arch. Intern. Med. 168 (3) (2008) 308–314, https://doi.org/10.1001/archinternmed.2007.119.
[44] T.T. Fung, S.E. Chiuve, M.L. McCullough, K.M. Rexrode, G. Logroscino, F.B. Hu, Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women, Arch. Intern. Med. 168 (7) (2008) 713–720, https://doi.org/10.1001/archinte.168.7.713.
[45] T.T. Fung, M.L. McCullough, P. Newby, J.E. Manson, J.B. Meigs, N. Rifai, et al., Diet-quality scores and plasma concentrations of markers of inflammation and endothelial dysfunction, Am. J. Clin. Nutr. 82 (1) (2005) 163–173, https://doi.org/10.1093/ajcn.82.1.163.
[46] A. Satija, S.N. Bhupathiraju, D. Spiegelman, S.E. Chiuve, J.E. Manson, W. Willett, et al., Healthful and unhealthful plant-based diets and the risk of coronary heart disease in U.S. adults, J. Am. Coll. Cardiol. 70 (4) (2017) 411–422, https://doi.org/10.1016/j.jacc.2017.05.047.
[47] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, J. Stat. Softw. 33 (1) (2010) 1–22, https://doi.org/10.18637/jss.v033.i01.
[48] D.M. Bates, lme4: mixed-effects modeling with R, Springer, New York, 2010.
[49] D. Tingley, T. Yamamoto, K. Hirose, L. Keele, K. Imai, Mediation: R package for causal mediation analysis, J. Stat. Softw. 59 (5) (2014) 1–38, https://doi.org/10.18637/jss.v059.i05.
[50] W. Viechtbauer, Conducting meta-analyses in R with the metafor package, J. Stat. Softw. 36 (3) (2010) 1–48, https://doi.org/10.18637/jss.v036.i03.
[51] W.G. Cochran, The combination of estimates from different experiments, Biometrics 10 (1) (1954) 101–129, https://doi.org/10.2307/3001666.
[52] G.C. Chen, J.C. Chai, J. Xing, J.Y. Moon, Z. Shan, B. Yu, et al., Healthful eating patterns, serum metabolite profile and risk of diabetes in a population-based prospective study of US Hispanics/Latinos, Diabetologia 65 (7) (2022) 1133–1144, https://doi.org/10.1007/s00125-022-05690-w.
[53] J. Li, M. Guasch-Ferré, W. Chung, M. Ruiz-Canela, E. Toledo, D. Corella, et al., The Mediterranean diet, plasma metabolome, and cardiovascular disease risk, Eur. Heart J. 41 (28) (2020) 2645–2656, https://doi.org/10.1093/eurheartj/ehaa209.
[54] C. Sun, C. Cao, T. Zhao, H. Guo, B.C. Fleming, B. Owens, et al., A2M inhibits inflammatory mediators of chondrocytes by blocking IL-1β/NF-κB pathway, J. Orthop. Res. 41 (1) (2023) 241–248, https://doi.org/10.1002/jor.25348.
[55] R. Tungtrongchitr, P. Pongpaew, N. Vudhivai, S. Changbumrung, A. Tungtrongchitr, B. Phonrat, et al., Relationship between alpha-2-macroglobulin, anthropometric parameters and lipid profiles in Thai overweight and obese in Bangkok, Nutr. Res. 23 (9) (2003) 1143–1152, https://doi.org/10.1016/S0271-5317(02)00529-8.
[56] M. Wilkinson, P. Sinclair, L. Dellatorre-Teixeira, P. Swan, E. Brennan, B. Moran, et al., The molecular effects of a high fat diet on endometrial tumour biology, Life (Basel) 10 (9) (2020) 188, https://doi.org/10.3390/life10090188.

[57] M. Wilkinson, B. Moran, P. Sinclair, P. Swan, E. Brennan, N. Docherty, et al., A2m–a novel diet related endometrial cancer biomarker, Gynecol. Oncol. 159 (Suppl 1) (2020) 38–39, https://doi.org/10.1016/j.ygyno.2020.06.082.
[58] M.M.H. Khan, Is this the protein, alpha-2-macroglobulin (A2M) is keeping us healthy, World J. Pharm. Res. 9 (3) (2020) 108–115, https://doi.org/10.20959/wjpr20203-16894.
[59] MM Khan, Elevated broad spectrum protease inhibitor, alpha-2-macroglobulin (A2M) in long lived naked mole rat (NMR) may open a new window to develop a medicine for our healthy life, Int. J. Med. Sci. Health Res 2 (3) (2018) 299–308.
[60] M. Zhu, B. Zhao, L. Wei, S. Wang, Alpha-2-macroglobulin, a native and powerful proteinase inhibitor, prevents cartilage degeneration disease by inhibiting majority of catabolic enzymes and cytokines, Curr. Mol. Biol. Rep. 7 (1) (2021) 1–7, https://doi.org/10.1007/s40610-020-00142-z.
[61] Q. Yang, T.E. Graham, N. Mody, F. Preitner, O.D. Peroni, J.M. Zabolotny, et al., Serum retinol binding protein 4 contributes to insulin resistance in obesity and type 2 diabetes, Nature 436 (7049) (2005) 356–362, https://doi.org/10.1038/nature03711.
[62] M. Fasshauer, M. Blüher, Adipokines in health and disease, Trends Pharmacol. Sci. 36 (7) (2015) 461–470, https://doi.org/10.1016/j.tips.2015.04.014.
[63] R.C.R. Meex, M.J. Watt, Hepatokines: linking nonalcoholic fatty liver disease and insulin resistance, Nat. Rev. Endocrinol. 13 (9) (2017) 509–520, https://doi.org/10.1038/nrendo.2017.56.
[64] T.E. Graham, Q. Yang, M. Blüher, A. Hammarstedt, T.P. Ciaraldi, R.R. Henry, et al., Retinol-binding protein 4 and insulin resistance in lean, obese, and diabetic subjects, N. Engl. J. Med. 354 (24) (2006) 2552–2563, https://doi.org/10.1056/NEJMoa054862.
[65] D.G. Haider, K. Schindler, G. Prager, A. Bohdjalian, A. Luger, M. Wolzt, et al., Serum retinol-binding protein 4 is reduced after weight loss in morbidly obese subjects, J. Clin. Endocrinol. Metab. 92 (3) (2007) 1168–1171, https://doi.org/10.1210/jc.2006-1839.
[66] J.W. Lee, H.R. Lee, J.Y. Shim, J.A. Im, D.C. Lee, Abdominal visceral fat reduction is associated with favorable changes of serum retinol binding protein-4 in nondiabetic subjects, Endocr. J. 55 (5) (2008) 811–818, https://doi.org/10.1507/endocrj.k08e-030.
[67] A.M. Armet, E.C. Deehan, A.F. O'Sullivan, J.F. Mota, C.J. Field, C.M. Prado, et al., Rethinking healthy eating in light of the gut microbiome, Cell Host Microbe 30 (6) (2022) 764–785, https://doi.org/10.1016/j.chom.2022.04.016.
[68] C. Rigauts, J. Aizawa, S.L. Taylor, G.B. Rogers, M. Govaerts, P. Cos, et al., *Rothia mucilaginosa* is an anti-inflammatory bacterium in the respiratory tract of patients with chronic lung disease, Eur. Respir. J. 59 (5) (2022) 2101293, https://doi.org/10.1183/13993003.01293-2021.
[69] M.N. Islam, A. Rauf, F.I. Fahad, T.B. Emran, S. Mitra, A. Olatunde, et al., Superoxide dismutase: an updated review on its health benefits and industrial applications, Crit. Rev. Food Sci. Nutr. 62 (26) (2022) 7282–7300, https://doi.org/10.1080/10408398.2021.1913400.
[70] H. Younus, Therapeutic potentials of superoxide dismutase, Int. J. Health Sci. (Qassim). 12 (3) (2018) 88–93.
[71] T. Fukai, M. Ushio-Fukai, Superoxide dismutases: role in redox signaling, vascular function, and diseases, Antioxid. Redox Signal. 15 (6) (2011) 1583–1606, https://doi.org/10.1089/ars.2011.3999.
[72] K. Yasui, A. Baba, Therapeutic potential of superoxide dismutase (SOD) for resolution of inflammation, Inflamm. Res. 55 (9) (2006) 359–363, https://doi.org/10.1007/s00011-006-5195-y.