ARIC Manuscript Proposal #4091

PC Reviewed: 8/9/22	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Associations of circulating proteins with lipoprotein profiles: Proteomic analyses from the OmniHeart Randomized Trial and the Atherosclerosis Risk in Communities.

b. Abbreviated Title (Length 26 characters): Proteomics, diet, lipids

2. Writing Group:

Hyunju Kim
Alice H. Lichtenstein
Peter Ganz
Edgar R. Miller, III
Josef Coresh
Lawrence J. Appel
Casey M. Rebholz

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __HK__ **[please confirm with your initials electronically or in writing]**

First author: Hyunju Kim
Address: 2024 E. Monument St, Suite 2-500
Baltimore MD 21205

Phone: 202-384-9291 Fax: N/A
E-mail: hkim25@jhu.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).
Name: Casey M. Rebholz
Address: 2024 E. Monument St, Suite 2-519
Baltimore MD 21287

Phone: 410-502-2359 Fax: N/A
E-mail: crebhol1@jhu.edu

3. Timeline:

We have completed data analyses and written a draft manuscript that includes the proteomic findings for the OmniHeart trial. To strengthen the value of the findings we would like to use

data from the ARIC Study to conduct validation only analysis involving10 proteins of interest associated with lipid outcomes identified from the OmniHeart trial analyses and add these results to the existing manuscript as soon as possible.

4. Rationale:

Healthy dietary patterns have been recommended to decrease cardiovascular disease (CVD) risk [1,2]. Within healthy dietary patterns, research has shown that manipulation of the proportion of macronutrients can reduce CVD risk [3]. Despite the evidence relating healthy dietary patterns and macronutrient intake, key gaps in knowledge persist. The molecular mechanisms through which macronutrients, within the framework of healthy dietary patterns, are associated with CVD is poorly understood.

Large-scale proteomic profiling offers an opportunity to address these gaps. Plasma proteins has diverse biological functions, including transporting nutrients and hormones, serving as enzymes and regulatory molecules, transmitting and receiving signals, and mediating immune responses [4–6]. These functions are relevant to digestion, absorption, transport, and metabolism of food. Thus, characterizing protein signatures of dietary patterns can expand our understanding of biological mechanisms underpinning diet/disease relationships.

Previous research on protein signatures of dietary patterns identified using established diet indices or principal component analysis relied on self-reported dietary data, which has the limitation of potential systematic biases and measurement error [7–10]. Some of these prior studies reported that proteins associated with dietary patterns were involved in lipid metabolism [8–10], but no study directly linked protein signatures of dietary patterns with serum lipoprotein concentrations. Further, most studies did not validate significant associations in an independent population or in a different study design, making it challenging to distinguish true and spurious associations.

The Optimal Macronutrient Intake Trial to Prevent Heart Disease (OmniHeart) was a randomized isocaloric feeding study consisting of 3 healthful dietary patterns which differed in the relative proportion of macronutrients (carbohydrate, protein, unsaturated fat) [3]. Compared to the carbohydrate-rich dietary pattern, the protein-rich dietary pattern reduced low-density lipoprotein (LDL)-cholesterol (C), high-density lipoprotein (HDL)-C, triglycerides, and non-HDL-C. Compared to the carbohydrate-rich dietary pattern, the unsaturated fat-rich dietary pattern increased HDL-C, reduced triglycerides, and non-HDL-C, and had no significant effect on LDL-C concentrations. Compared to the unsaturated fat-rich dietary pattern, the protein-rich dietary pattern reduced HDL-C and triglyceride concentrations, and had no significant effect on LDL-C and non-HDL-C concentrations.

Using proteomics data generated from plasma specimens collected during the OmniHeart trial, we aimed to determine whether any proteins or protein signatures mediated the association between dietary patterns and serum lipoprotein concentrations. Then, we aimed to *validate* the associations between diet-related proteins and lipoprotein outcomes in a large observational study (Atherosclerosis Risk in Communities [ARIC] Study).

5. Main Hypothesis/Study Questions:

Aim 1: To identify proteins mediating the association between dietary patterns and lipoprotein concentrations in the OmniHeart trial (completed already)

Aim 2: To validate the associations between diet-related proteins and lipid outcomes from the OmniHeart trial in the ARIC Study.

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Study Design:

We completed data analysis and manuscript preparation of aim 1. This proposal will describe validation analyses for aim 2 only. For aim 2, we will conduct cross-sectional analysis of 10 diet-related proteins and 4 lipid outcomes using data from ARIC visit 2.

Inclusion/Exclusion Criteria:
We will include all participants who attended ARIC visit 2. Participants will be excluded based on the following criteria: 1) non-black or non-white participants, blacks in Minneapolis, Minnesota, or blacks in Washington County, Maryland due to small numbers, 2) participants with incomplete information on 10 diet-related proteins measured by SomaLogic, 3) participants with missing lipid outcomes, or 4) participants with missing covariates.

Exposure:

We will use $\log_2$-transformed proteins significantly associated with dietary patterns and lipoprotein concentrations in the OmniHeart trial. These 10 diet-related proteins are

1. apolipoprotein M
2. afamin
3. collagen alpha-3(VI) chain
4. chitinase-3-like protein 1
5. inhibin beta A chain
6. palmitoleoyl-protein carboxylesterase NOTUM
7. cathelicidin antimicrobial peptide
8. guanylate-binding protein 2
9. COP9 signalosome complex subunit 7b
10. sodium-coupled monocarboxylate transporter 1

Outcome:

We will use 4 serum lipid outcomes:

1. HDL-C
2. Triglycerides
3. Non-HDL-C
4. Ratio of total cholesterol to HDL

We excluded LDL-C and total cholesterol, because we found no significant association between diet-related proteins and LDL-C, and the effect of the diet interventions varied for HDL-C and non-HDL-C concentrations [3].

Statistical Analysis

We will use multivariable linear regression models to validate the associations between diet-related proteins and lipoprotein outcomes from the OmniHeart trial. In the ARIC study, we will adjust for age, sex, race-center, smoking status, body mass index, and estimated glomerular filtration rate based on creatinine. We will consider associations to be validated at the Bonferroni threshold. In the OmniHeart trial, there were 13 significant diet-lipoprotein associations. Therefore, the Bonferroni threshold used in the ARIC study will be 3.85×10^{-3} (0.05/13).

Anticipated Methodologic Limitations or Challenges:

We were not able to validate the associations between diet and significant proteins in the ARIC study. It is difficult to model the unique design of the OmniHeart trial in which intervention arms differed in the level of macronutrients intake within the context of healthy dietary patterns. However, validation of the associations between diet-related proteins and lipoprotein concentrations increases confidence that the associations observed in the OmniHeart trial are true associations.

In the OmniHeart trial and the ARIC study, only relative quantification of proteins is available, and there are no data on absolute protein concentrations.

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes __√__ No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No
(The file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ____ Yes __√__ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 must be used to exclude those with value RES_DNA = "No use/storage DNA"? ____ Yes ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.
ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/aric/mantrack/maintain/search/dtSearch.html

___√___ Yes _____ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

#3755: Protein biomarkers of the DASH diet and Hypertension Risk in the ARIC Study (lead author: Shutong Du)

> This proposal aims to identify protein biomarkers of the DASH diet and hypertension risk. This proposal focuses on the DASH diet, but we propose to conduct validation analyses of only the diet-related proteins that were associated with lipid outcomes in the OmniHeart Trial. Further, the OmniHeart diets are modified versions of the DASH diet and is different from the DASH diet in that the OmniHeart diets focused on macronutrient composition.

#3509: Proteomics of apolipoprotein E genetic polymorphisms: The Atherosclerosis Risk in Communities Study (lead author: Pamela L. Lutsey)

> This proposal aims to identify plasma proteins associated with apolipoprotein E genetic polymorphisms. Our proposal is significantly different from Dr. Lutsey's proposal in that we aim to conduct validation analyses of 10 proteins that had statistically significant associations between diet and lipoprotein concentrations in the OmniHeart trial. Further, our proposal does not examine apolipoprotein E genetic polymorphisms.

There are no existing manuscript proposals on proteins that are related to diet and lipoprotein concentrations.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? _√___ Yes ____ No

11.b. If yes, is the proposal

_√ _ A. primarily the result of an ancillary study (list number* 2019.26, 2017.27)

___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____ _____ _____)

*ancillary studies are listed by number https://sites.cscc.unc.edu/aric/approved-ancillary-studies

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms.

http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

References:

1. Lichtenstein AH, Appel LJ, Vadiveloo M, Hu FB, Kris-Etherton PM, Rebholz CM, et al. 2021 Dietary Guidance to Improve Cardiovascular Health: A Scientific Statement From the American Heart Association. Circulation. 2021;144:e472–87.

2. US Department of Agriculture, US Department of Health and Human Services, Health and Human Services. Dietary Guidelines for Americans, 2020-2025 [Internet]. Washington (DC): US Government Printing Office; 2020. Available from: dietaryguidelines.gov

3. Appel LJ, Sacks FM, Carey VJ, Obarzanek E, Swain JF, Miller ER, et al. Effects of protein, monounsaturated fat, and carbohydrate intake on blood pressure and serum lipids: results of the OmniHeart Randomized trial. JAMA. 2005;294:2455.

4. Anderson NL, Anderson NG. The Human Plasma Proteome: History, Character, and Diagnostic Prospects*. Molecular & Cellular Proteomics. 2002;1:845–67.

5. Lee SE, Schulze K, West KP. Rainer Gross award lecture 2018: the childhood plasma proteome: discovering its applications in public health nutrition. Food Nutr Bull. SAGE Publications Inc; 2019;40:144–50.

6. Wang J, Li D, Dangott LJ, Wu G. Proteomics and Its Role in Nutrition Research. The Journal of Nutrition. 2006;136:1759–62.

7. Walker ME, Song RJ, Xu X, Gerszten RE, Ngo D, Clish CB, et al. Proteomic and metabolomic correlates of healthy dietary patterns: the Framingham heart study. Nutrients. 2020;12:1476.

8. Kim Y, Lu S, Ho JE, Hwang S-J, Yao C, Huan T, et al. Proteins as mediators of the association between diet quality and incident cardiovascular disease and all-cause mortality: the Framingham heart study. J Am Heart Assoc. 2021;10:e021245.

9. García-Bailo B, Brenner DR, Nielsen D, Lee H-J, Domanski D, Kuzyk M, et al. Dietary patterns and ethnicity are associated with distinct plasma proteomic groups. Am J Clin Nutr. 2012;95:352–61.

10. Warensjö Lemming E, Byberg L, Stattin K, Ahmad S, Lind L, Elmståhl S, et al. Dietary pattern specific protein biomarkers for cardiovascular disease: a cross-sectional study in 2 independent cohorts. J Am Heart Assoc. 2019;8:e011860.