ARIC Manuscript Proposal #3718

PC Reviewed: 10/13/20	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Serum Metabolomic Markers of Protein and Kidney Disease Risk

b. Abbreviated Title (Length 26 characters): Dietary protein metabolomics

2. Writing Group:
Writing group members:

Casey M. Rebholz, Jingsha Chen, Bing Yu, Lyn M. Steffen, Eric Boerwinkle, Morgan E. Grams, Josef Coresh

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

First author: Casey M. Rebholz
Address: 2024 East Monument Street, Suite 2-500
Baltimore, Maryland 21287

Phone: 410-502-2359
E-mail: crebhol1@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: Josef Coresh
Address: 2024 East Monument Street, Suite 2-600
Baltimore, Maryland 21287

Phone: 410-955-0495 Fax: 410-955-0476
E-mail: coresh@jhu.edu

3. Timeline: Analyses will begin after the manuscript proposal is approved. We anticipate that a first draft of the manuscript will be available within approximately one year of manuscript proposal approval.

4. Rationale:
Diet is an important modifiable risk factor for cardiovascular disease, kidney disease, and other chronic diseases.[1,2] Clinical recommendations for kidney disease patients recommend restriction of protein.[3,4] In the ARIC study, we have previously reported that total protein and animal protein were not associated with kidney disease risk, but that higher intake of vegetable protein was significantly associated with a reduced risk of incident chronic kidney disease (HR for

quartile 4 vs. 1: 0.72, 95% CI: 0.61, 0.85; p-trend across quartiles <0.001).[5] In addition, we have shown that, in the ARIC study, whereas higher consumption of red and processed meat was associated with a higher risk of incident kidney disease, higher consumption of fish, low-fat dairy products, nuts, and legumes were significantly associated with a lower risk of developing kidney disease.[6] Further research is necessary to examine the metabolic disturbances associated with overall dietary protein and specific sources of protein which have variable associations with kidney disease risk.

Metabolomics allows for the comprehensive characterization of small metabolic compounds in biological specimens (serum).[7] The metabolome is responsive to dietary intake and therefore is a useful method for detecting biomarkers of dietary intake of protein and metabolic pathways leading to kidney disease that are potentially modifiable by diet.[8] The untargeted and unbiased metabolomic approach maximizes the potential for discovery of novel markers of dietary intake and could provide insights about metabolic pathways underlying the diet-kidney disease relationship.

5. Main Hypothesis/Study Questions:

Hypothesis #1: We hypothesize that we will be able to identify known and novel metabolites associated with overall dietary protein and specific food sources of protein. We hypothesize that there will be metabolites that are uniquely similarly associated with specific protein sources.

Hypothesis #2: We hypothesize that some of the dietary protein-related metabolites will be associated with kidney disease risk.

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: cross-sectional analysis (hypothesis #1) of metabolomics and measures of dietary intake of protein, which both were assessed at study visit 1 (1987-1989) and prospective analysis (hypothesis #2) of diet-related metabolites and risk of incident kidney disease through the latest follow-up period

Eligibility Criteria: Approximately 4,000 African-American and Caucasian ARIC study participants with metabolomic profiling data from visit 1 serum specimens (ancillary study #2014.20 and 2008.16; two "batches")

Exposures & Outcomes: For hypothesis #1, the exposure will be dietary protein and outcome will be metabolites. For hypothesis #2, the exposure will be diet-related metabolites and the outcome will be incident kidney disease.

> **Dietary Protein:** Dietary intake was assessed at study visit using an interview-administered, in-person, 66-item, semi-quantitative, food frequency questionnaire 1, which was modified from an instrument developed by Willett et al.[9] We will assess the following measures of dietary intake of protein, with the same classifications as our previous study[6]:

1. Total protein
2. Animal protein
3. Vegetable protein
4. Red meat
5. Processed meat
6. Fish
7. Low-fat dairy products
8. Nuts
9. Legumes
10. Poultry
11. Eggs
12. High-fat dairy products

Metabolomics: Metabolites were measured from stored fasting serum samples by Metabolon, Inc. (Durham, North Carolina) using an untargeted, ultra-performance liquid chromatography tandem mass spectrometry approach. This analysis will focus on the approximately 600-800 named metabolites that were identified in the two ARIC samples.

Kidney Disease: We will study incident chronic kidney disease (CKD) and incident end-stage renal disease (ESRD). Incident CKD is defined by at least one of the following four criteria: 1) development of reduced kidney function (eGFR <60 mL/min/1.73 m^2) accompanied by 25% eGFR decline at any subsequent study visit relative to baseline, 2) International Classification of Diseases (ICD)-9/10 code for a hospitalization related to CKD stage 3+ identified through active surveillance of the ARIC cohort, 3) ICD 9/10 code for a death related to CKD stage 3+ identified through linkage to the National Death Index, and 4) end-stage renal disease identified by linkage to the US Renal Data System (USRDS) registry.[10] Incident ESRD cases will be identified by linkage to the USRDS registry.[11]

Other Variables of Interest: In multivariable linear regression models, we will consider adjusting for the following variables: age, sex, race, center, body mass index, total energy intake, and estimated glomerular filtration rate (eGFR).

Statistical Analysis: We will use multivariable linear regression models to estimate the cross-sectional association between dietary protein (exposure) and metabolites (outcome) and we will use Cox proportional hazards regression to estimate the prospective association between diet-related metabolites and incident kidney disease. Effect estimates will be calculated per one unit higher in dietary protein for the cross-sectional analysis and per one standard deviation higher in the metabolites for the prospective analysis. Metabolites will be log-transformed for analysis. We will adjust for the following covariates in the multivariable regression model: age, sex, race, center, body mass index, total energy intake, and eGFR. For the prospective analysis, we will additionally adjust for hypertension, history of cardiovascular disease, and diabetes. Analyses will be conducted by batch (1st batch: discovery, 2nd batch: replication). We will use C statistics to evaluate the ability of the candidate biomarkers (metabolites found to be significantly associated with dietary protein) improve the prediction of highest quartile vs. lower 3 quartiles of dietary protein intake.

Anticipated Methodologic Limitations or Challenges: Given the large number of metabolites, there is a high likelihood of detecting a false positive association. We will adjust the significance threshold by the Bonferroni method (dividing by the number of metabolites) to account for multiple comparisons (0.05/number of metabolites).[12] For hypothesis #1, all metabolites will be analyzed. For hypothesis #2, only metabolites that are associated with diet at the Bonferroni threshold will then be investigated in relation to kidney disease risk.

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes __X__ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used? ____ Yes ____ No
(This 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 __X__ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 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/search.php

___X___ 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)?

#2034: The human metabolome is associated with dietary intake among African Americans in the Atherosclerosis Risk in Communities Study (lead author: Yan Zheng)

> The manuscript based on this proposal has already been published [Zheng Z, Yu B, Alexander D, Steffen LM, Boerwinkle E. Human metabolome associates with dietary intake habits among African Americans. Am J Epidemiol 2014;179(12):1424-1433.] It was focused on food groups and food items, whereas the present manuscript proposal is focused on dietary patterns. In addition, it included data on African Americans only, whereas the present manuscript proposal will include data on both African Americans and Caucasians.

#1882: A longitudinal study of metabolomics and kidney function among African Americans in the Atherosclerosis Risk in Communities (ARIC) study (lead author: Bing Yu)

The manuscript based on this proposal has already been published [Yu B, Zheng Y, Nettleton JA, Alexander D, Coresh J, Boerwinkle E. Serum metabolomic profiling and incident CKD among African Americans. Clin J Am Soc Neprhol 2014;9(8):1410-1417]. Similar to the other manuscript proposal, this analysis only included data on African Americans. The published paper is more comprehensive than the proposed analysis in that the authors reported on all available metabolites whereas this proposal will analyze only diet-related metabolites in association with kidney disease risk.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __ X __ Yes __ __ No

11.b. If yes, is the proposal

_X__ A. primarily the result of an ancillary study

2014.20: Genomics, Metabolomics, and Cardiovascular Disease (PI: Eric Boerwinkle)

2008.16: Metabolomics & Heart Failure: A Novel Approach to Biomarker Discovery (PI: Jennifer Nettleton)

___ 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 at http://www.cscc.unc.edu/aric/forms/

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. Mozaffarian D, Appel LJ, Van Horn L. Components of a cardioprotective diet: new insights. *Circulation.* 2011;123(24):2870-2891.
2. Micha R, Penalvo JL, Cudhea F, Imamura F, Rehm CD, Mozaffarian D. Association Between Dietary Factors and Mortality From Heart Disease, Stroke, and Type 2 Diabetes in the United States. *JAMA : the journal of the American Medical Association.* 2017;317(9):912-924.

3. National Kidney Foundation. KDOQI Clinical Practice Guideline for Diabetes and CKD: 2012 Update. *American journal of kidney diseases : the official journal of the National Kidney Foundation.* 2012;60(5):850-886.
4. Kidney Disease: Improving Global Outcomes (KDIGO). KDIGO clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney International Supplements.* 2013;3(1):1-150.
5. Rebholz CM, Coresh J, Grams ME, et al. Dietary Acid Load and Incident Chronic Kidney Disease: Results from the ARIC Study. *American journal of nephrology.* 2015;42(6):427-435.
6. Haring B, Selvin E, Liang M, et al. Dietary Protein Sources and Risk for Incident Chronic Kidney Disease: Results From the Atherosclerosis Risk in Communities (ARIC) Study. *J Ren Nutr.* 2017;27(4):233-242.
7. Tzoulaki I, Ebbels TM, Valdes A, Elliott P, Ioannidis JP. Design and analysis of metabolomics studies in epidemiologic research: a primer on -omic technologies. *American journal of epidemiology.* 2014;180(2):129-139.
8. Guasch-Ferre M, Bhupathiraju SN, Hu FB. Use of Metabolomics in Improving Assessment of Dietary Intake. *Clinical chemistry.* 2018;64(1):82-98.
9. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *American journal of epidemiology.* 1985;122(1):51-65.
10. Grams ME, Rebholz CM, McMahon B, et al. Identification of incident CKD stage 3 in research studies. *Am J Kidney Dis.* 2014;64(2):214-221.
11. Rebholz CM, Coresh J, Ballew SH, et al. Kidney Failure and ESRD in the Atherosclerosis Risk in Communities (ARIC) Study: Comparing Ascertainment of Treated and Untreated Kidney Failure in a Cohort Study. *American journal of kidney diseases : the official journal of the National Kidney Foundation.* 2015;66(2):231-239.
12. Curtin F, Schulz P. Multiple correlations and Bonferroni's correction. *Biological psychiatry.* 1998;44(8):775-777.