ARIC Manuscript Proposal #3497

PC Reviewed: 1/14/20 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: Proteomic Profiling and Kidney Function in the Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): Proteomics and Kidney Function

2. Writing Group:

Writing group members: Zhi Yu, Adrienne Tin, Jingsha Chen, Anna Kottgen, Pascal Schlosser, Christie Ballantyne, Ron Hoogeveen, Bing Yu, [Eric Boerwinkle invited], Dan Arking, Morgan Grams, Josef Coresh (last author): others welcome

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

First author: **Zhi Yu, BM, MS**
Address: Johns Hopkins Bloomberg School of Public Health and Welch Center for Prevention, Epidemiology, and Clinical Research
2024 E. Monument St.,
Suite 2-600
Baltimore, MD 21287

Phone: 617-583-3861 Fax:
E-mail: zyu33@jhmi.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: Johns Hopkins Bloomberg School of Public Health and Welch Center for Prevention, Epidemiology, and Clinical Research
2024 E. Monument St.,
Suite 2-600
Baltimore, MD 21287

Phone: 410-995-0495 Fax: 410-955-0476
E-mail: coresh@jhu.edu

3. Timeline: 6 months; manuscript submission in spring 2020

4. Rationale:

Chronic kidney disease (CKD) is defined as abnormalities of kidney structure or function, present for >3 months, with implications for health[1]. It affects >10% adults in the U.S. and across the world[2, 3]. The clinical impact of CKD is substantial: major complications include end stage kidney disease (ESKD), cardiovascular disease (CVD), and mortality[4]. As a chronic, complex, and heterogeneous disease, CKD is influenced by multiple genomic and environmental factors. CKD is estimated to have high heritability (30–75%)[5-7]. Genome-wide association studies (GWAS) have identified genetic susceptibility locus for estimated glomerular filtration rate (eGFR) and albuminuria (urine albumin-to creatinine ratio [ACR])[8, 9], which are measurements of kidney function and change in these markers can be valid surrogate end points for clinical trials in early stage of CKD[10]. However, the genetic basis of reduced kidney function is becoming clearer but the proteomic expression and mechanisms are largely unknown.

Proteins often express gene functions and their variation in the human body. Alterations of protein structure or abundance can lead to changes in phenotype. Previous studies have found that the concentrations of multiple proteins in blood were significantly associated with reduced kidney function or kidney disease risks[11-14]. However evidence from large-scale studies that systematically evaluating the relationships between global plasma proteins and genetically reduced kidney function need to be generated.

Therefore, we propose to relate polygenic risk scores (PRS) to plasma proteins to discover proteins whose levels are associated with high polygenic risk of reduced kidney function, measured as low eGFR and high albuminuria, within the Atherosclerosis Risks in Communities (ARIC) Cohort. We will also identify pathways enriched with these proteins (separately for eGFR and ACR; and will consider if high and low PRS are mirror images of one another or not) by pathway analysis. Our study results will shed light on the potential mechanism of genetically reduced kidney function on the observed patterns of plasma proteins.

5. Main Hypothesis/Study Questions:

We hypothesize that genetic susceptibility for reduced kidney function aggregated in a PRS will be related to altered blood concentrations of a range of proteins in specific pathways (e.g. inflammation and fibrosis), potentially identifying new risk factors and targets for therapy.

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).

Inclusion criteria:
(i) Attended visit 3 or visit 5.
(ii) Have SOMAscan protein measurements available from plasma sample collected at visit 3 or visit 5.
(iii) Have valid genetic measurements and genome-wide imputation.

Exclusion criteria:

(i) Non-white and non-black participants.
(ii) Non-white participants in Washington County and Minnesota.

Exposure variables:
PRSs of (i) low eGFR and (ii) high albuminuria calculated using GWAS results from the CKDGen Consortium (Wuttke et al and Teumer et al). We will build (i) global PRSs and (ii) PRSs generated using index SNPs only for low eGFR and high albuminuria separately. All PRSs will be race specific (white / black).

Generation of global PRSs will be performed as follows: (i) download the GWAS summary statistics for low eGFR and high albuminuria from the CKDGen Consortium, (ii) filter out non-genome-wide-significant SNPs, (iii) remove SNPs that were not available in ARIC or did not match the ARIC data, (iv) conduct LD clumping using PLINK with r^2 cutoff = 0.1 and use the remaining SNPs as risk alleles, and (v) calculate the weighted sum of the number of risk alleles possessed by an individual, in which the weight will be taken as the effect estimate associated with each SNP. For PRSs generated using index SNPs only, the risk alleles will be the index SNPs identified by the CKDGen Consortium, and the score will also be calculated as the weighted sum of the number of risk alleles. In addition, for both (i) global PRSs and (ii) PRSs generated using index SNPs only, we will use LDpred as secondary method to calculate the scores[15].

Outcome variables:
Around 4000 proteins per person measured using plasma collected at visit 3 (1993-1995) and 5 (2011-2013). The proteins were measured using a Slow Off-rate Modified Aptamer (SOMAmer)-based capture array (SomaLogic, Inc, Boulder, Colorado), which uses modified DNA bases rather than antibodies.

Analytic Plan:
(i) Plasma proteins associated with PRSs of low eGFR and high albuminuria: we will use linear regression models to examine the extent to which PRSs are associated with each protein (natural log transformed) at visit 3 and visit 5 separately. Analyses will be adjusted for age at sample acquisition and sex. Bonferroni corrected p-values will be used as significance threshold. We will compare the results between visit 3 and visit 5 results. If we observe major differences, we will also examine the PRSs in relation to changes in protein levels over time.

(ii) Protein pathway enrichment: we will perform enrichment analysis based on KEGG PATHWAY to identify pathways with over representation of the PRS-associated proteins[16]. Since KEGG has a selective list of proteins and pathways, we will use WikiPathways as supplemental resources.

7.a. Will the data be used for non-CVD analysis in this manuscript? __X__ Yes ____ 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? __X__ 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? __X__ Yes ____ 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"? __X__ 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

___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)?

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 (list number* 2017.27_)
"Proteomic longitudinal ARIC study: SOMAscan of multiple visits"

___ 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 https://www2.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.

Get it.

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. KDIGO CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl.* 2013;3:1–150.
2. Eckardt KU, Coresh J, Devuyst O, et al. Evolving importance of kidney disease: from subspecialty to global health burden. *Lancet (London, England).* 2013;382:158-169.
3. Baumeister SE, Boger CA, Kramer BK, et al. Effect of chronic kidney disease and comorbid conditions on health care costs: A 10-year observational study in a general population. *American journal of nephrology.* 2010;31:222-229.
4. Whelton PK, Klag MJ. Hypertension as a risk factor for renal disease. Review of clinical and epidemiological evidence. *Hypertension (Dallas, Tex. : 1979).* 1989;13:I19-27.
5. Satko SG, Freedman BI. The familial clustering of renal disease and related phenotypes. *The Medical clinics of North America.* 2005;89:447-456.
6. O'Seaghdha CM, Fox CS. Genome-wide association studies of chronic kidney disease: what have we learned? *Nature reviews. Nephrology.* 2011;8:89-99.
7. Regele F, Jelencsics K, Shiffman D, et al. Genome-wide studies to identify risk factors for kidney disease with a focus on patients with diabetes. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2015;30 Suppl 4:iv26-34.
8. Wuttke M, Li Y, Li M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nature genetics.* 2019;51:957-972.
9. Teumer A, Li Y, Ghasemi S, et al. Genome-wide association meta-analyses and fine-mapping elucidate pathways influencing albuminuria. *Nature communications.* 2019;10:4130.
10. Levey AS, Gansevoort RT, Coresh J, et al. Change in Albuminuria and GFR as End Points for Clinical Trials in Early Stages of CKD: A Scientific Workshop Sponsored by the National Kidney Foundation in Collaboration With the US Food and Drug Administration and European Medicines Agency. *American journal of kidney diseases : the official journal of the National Kidney Foundation.* 2019.
11. Siwy J, Zurbig P, Argiles A, et al. Noninvasive diagnosis of chronic kidney diseases using urinary proteome analysis. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2017;32:2079-2089.
12. Good DM, Zurbig P, Argiles A, et al. Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Molecular & cellular proteomics : MCP.* 2010;9:2424-2437.
13. Schanstra JP, Zurbig P, Alkhalaf A, et al. Diagnosis and Prediction of CKD Progression by Assessment of Urinary Peptides. *Journal of the American Society of Nephrology : JASN.* 2015;26:1999-2010.
14. Pontillo C, Jacobs L, Staessen JA, et al. A urinary proteome-based classifier for the early detection of decline in glomerular filtration. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2017;32:1510-1516.
15. Vilhjalmsson BJ, Yang J, Finucane HK, et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *American journal of human genetics.* 2015;97:576-592.
16. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols.* 2009;4:44-57.

1. KDIGO CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl.* 2013;3:1–150.
2. Eckardt KU, Coresh J, Devuyst O, et al. Evolving importance of kidney disease: from subspecialty to global health burden. *Lancet (London, England).* 2013;382:158-169.
3. Baumeister SE, Boger CA, Kramer BK, et al. Effect of chronic kidney disease and comorbid conditions on health care costs: A 10-year observational study in a general population. *American journal of nephrology.* 2010;31:222-229.
4. Whelton PK, Klag MJ. Hypertension as a risk factor for renal disease. Review of clinical and epidemiological evidence. *Hypertension (Dallas, Tex. : 1979).* 1989;13:I19-27.

5. Satko SG, Freedman BI. The familial clustering of renal disease and related phenotypes. *The Medical clinics of North America.* 2005;89:447-456.
6. O'Seaghdha CM, Fox CS. Genome-wide association studies of chronic kidney disease: what have we learned? *Nature reviews. Nephrology.* 2011;8:89-99.
7. Regele F, Jelencsics K, Shiffman D, et al. Genome-wide studies to identify risk factors for kidney disease with a focus on patients with diabetes. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2015;30 Suppl 4:iv26-34.
8. Wuttke M, Li Y, Li M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nature genetics.* 2019;51:957-972.
9. Teumer A, Li Y, Ghasemi S, et al. Genome-wide association meta-analyses and fine-mapping elucidate pathways influencing albuminuria. *Nature communications.* 2019;10:4130.
10. Levey AS, Gansevoort RT, Coresh J, et al. Change in Albuminuria and GFR as End Points for Clinical Trials in Early Stages of CKD: A Scientific Workshop Sponsored by the National Kidney Foundation in Collaboration With the US Food and Drug Administration and European Medicines Agency. *American journal of kidney diseases : the official journal of the National Kidney Foundation.* 2019.
11. Siwy J, Zurbig P, Argiles A, et al. Noninvasive diagnosis of chronic kidney diseases using urinary proteome analysis. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2017;32:2079-2089.
12. Good DM, Zurbig P, Argiles A, et al. Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Molecular & cellular proteomics : MCP.* 2010;9:2424-2437.
13. Schanstra JP, Zurbig P, Alkhalaf A, et al. Diagnosis and Prediction of CKD Progression by Assessment of Urinary Peptides. *Journal of the American Society of Nephrology : JASN.* 2015;26:1999-2010.
14. Pontillo C, Jacobs L, Staessen JA, et al. A urinary proteome-based classifier for the early detection of decline in glomerular filtration. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2017;32:1510-1516.
15. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols.* 2009;4:44-57.