ARIC Manuscript Proposal #3519

PC Reviewed: 12/10/19	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: *APOL1* risk variants and proteomics in a community based population

b. Abbreviated Title (Length 26 characters): *APOL1* and proteomics

2. Writing Group:

Writing group members: Teresa K. Chen, Bing Yu, Jingsha Chen, Josef Coresh, Christie Ballantyne, Eric Boerwinkle, Adrienne Tin, Dan Arking, Katalin Susztak, Anna Kottgen, Morgan E. Grams, *others welcome* (order TBD).

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

First author: Teresa K. Chen
Address: 1800 E. Monument Street, Suite 416
Baltimore, Maryland 2187
Phone: 410-955-5268 Fax: 410-550-7950
E-mail: tchen39@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: Morgan E. Grams
Address: 2024 East Monument Street, Suite 2-600
Baltimore, Maryland 21287
Phone: 410-287-1827 Fax: 410-955-0485
E-mail: mgrams2@jhmi.edu

3. Timeline: Analyses will begin once the manuscript proposal has been approved. We anticipate that the manuscript will be written and submitted to the ARIC Publications Committee within one year of the manuscript proposal being approved.

4. Rationale: The *APOL1* high-risk genotypes, present in 13% of black Americans, have been associated with incident and progressive chronic kidney disease (CKD).[1-5] In the African American Study of Kidney Disease and Hypertension (AASK), having two copies of the *APOL1* risk variants (high-risk; G1/G1, G1/G2, or G2/G2) was associated with a 1.88-fold higher risk of CKD progression, defined as a doubling of serum creatinine or end-stage renal disease (ESRD), compared to having one or no copies (low-risk; G1/G0, G2/G0, G0/G0). However, 40% of

AASK participants with two copies of the *APOL1* risk variants did not experience CKD progression over a median follow-up of 9 years.[4] Similarly, Grams *et al.* previously reported in ARIC that black participants with the *APOL1* high-risk status had a significantly higher risk of incident ESRD and faster rate of eGFR decline compared to black participants with the *APOL1* low-risk status yet there was significant overlap in the distributions of average yearly eGFR slopes among the two *APOL1* risk groups.[3] Given that not all individuals with the *APOL1* high-risk status develop kidney disease during their lifetime, a "second hit" is likely necessary. To date, *in vitro* studies have demonstrated that activation of the innate immunity pathway increases APOL1 expression in podocytes and endothelial cells[6] and that higher levels of APOL1 protein, particularly G1 and G2, are associated with increased cytotoxicity in human embryonic kidney cells.[7] In mouse models, transgenic podocyte expression of the *APOL1* risk variants has also been associated with albuminuria and azotemia.[7] Taken together, these studies suggest that proteins play an integral role in the pathogenesis of *APOL1*-associated kidney disease. On the other hand, the association of *APOL1* high-risk status with mortality is less clear, with some studies suggesting no association[3,8] and other studies suggesting a lower risk of mortality.[9,10] We propose to leverage proteomics data from ARIC Visits 3 and 5 to better understand the complex interplay of *APOL1* risk variants, the plasma proteome, and kidney disease.

5. Main Hypothesis/Study Questions:

Our overarching hypothesis is that proteomic pathways will provide insight on how the *APOL1* risk variants lead to kidney disease.

Aim 1: To identify proteins and pathways with different levels in persons with the *APOL1* high-risk alleles compared to the low-risk alleles in African-American ARIC participants.

Aim 2: To determine whether higher levels of APOL1 protein and other *APOL1* high-risk genotype-associated proteins are associated with incident CKD, ESRD, and mortality in both African-American and white ARIC participants.

Aim 3: To perform a genome-wide association study (GWAS) detecting genetic markers of higher APOL1 protein (and other *APOL1* high-risk genotype-associated proteins) levels separately in African-American and white ARIC participants. We will also explore the risk associations of identified genotypes with CKD in external datasets using Mendelian randomization methods.

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 will conduct analyses of the ARIC cohort, treating Visit 3 (1993-1995) as the baseline visit.

Study Population: The study population will consist of African-American ARIC participants with *APOL1* genotyping and proteomics data from Visits 3 and/or 5. In sensitivity analyses, we will also consider white ARIC participants as a comparison group.

Exposure: For Aim 1, the primary exposure will be *APOL1* genotype status. We will utilize a recessive genetic model with the *APOL1* high-risk status defined as having 2 risk alleles and the low-risk status defined as having 0 or 1 risk allele. In sensitivity analysis, we will also consider additive models. For Aim 2, the exposure of interest will be APOL1 protein (and other *APOL1* high-risk genotype-associated proteins) levels of APOL1 at Visit 3 and in secondary analysis we will explore protein levels at Visit 5. For Aim 3, the exposures will be genetic markers in the genome.

Outcomes: For Aim 1, the outcome will be proteomics data previously measured from Visits 3 and Visit 5 samples using the SomaScan Platform (SomaLogic, Inc). For Aim 2, the outcomes will be: 1) incident CKD; 2) incident ESRD; and 3) mortality. We will also consider rate of eGFR decline as an exploratory outcome. Creatinine-based and/or cystatin-based Chronic Kidney Disease Epidemiology (CKD-EPI) Collaboration equations will be used to estimate GFR.[11,12] Consistent with prior ARIC publications, incident CKD will be defined as having any one of the following: 1) an eGFR <60 ml/min/1.73 m^2 at follow-up (accompanied by a ≥25% eGFR decline relative to baseline; 2) CKD-related hospitalization or death based on the International Classification of Diseases (ICD) 9 or 10 codes; or 3) ESRD as identified by the US Renal Data System (USRDS) registry.[13-15] For exploratory analyses of eGFR decline, at the time of ESRD onset, eGFR will be imputed as 15 ml/min/1.73 m^2. Mortality will be determined from active surveillance techniques, including linkage to the National Death Index for the latter outcome.[3] For Aim 3, outcomes will be APOL1 protein (and other *APOL1* high-risk genotype-associated proteins) levels at Visit 3 and in secondary analysis we will explore protein levels at Visit 5.

Statistical Analysis: We will use descriptive statistics, including means, medians, and proportions to compare baseline characteristics by *APOL1* genotype status at Visit 3. Formal testing will be performed using student's t-test or Wilcoxon rank-sum test for continuous variables and chi-squared for categorical variables. We anticipate that the distributions of proteins will be skewed: we plan to transform (e.g., log base-2) to achieve a more normal distribution. For Aim 1 (n~2,700), linear regression models will be used to study the associations of *APOL1* risk status with proteins in the proteome. Bonferroni correction will be used to account for multiple comparisons. Model 1 will be unadjusted; Model 2 will adjust for age, sex, and center; Model 3 will further adjust for eGFR at Visit 3; and Model 4 will further adjust for cholesterol, HDLc, diabetes, systolic blood pressure, anti-hypertension medication at Visit 3. All analyses will be performed adjusting and not adjusting for European ancestry. When using Visit 5 as a baseline, Model 3-4 will also adjust for urine ACR. For Aim 2 (n~2,700), Cox proportional hazards models will be constructed to study the associations of APOL1 protein at Visit 3 and change in APOL1 protein from Visit 3 to Visit 5 with: 1) incident CKD; 2) incident ESRD; and 3) mortality. We will adjust for the same covariates as in Aim 1. To examine the association of APOL1 protein with subsequent eGFR decline, we will fit linear mixed-effects models with random intercepts and random slopes, adjusting for the same covariates as above except eGFR, which will be the independent variable. We will assess for effect modification by age, sex, and diabetes through the use of stratified analyses and inclusion of an interaction term with APOL1 protein level. We will perform these analyses overall and within African-American participants only. For Aim 3 (n~11,500), we will perform a GWAS to identify genetic markers

that are associated with higher APOL1 protein levels separately within African-American and white ARIC participants. Given the low prevalence of *APOL1* alleles among individuals of European ancestry, all white participants will be imputed as having the *APOL1* low-risk genotypes.[3,16] Analyses for Aims 2 and 3 will be repeated for other *APOL1* high-risk genotype-associated proteins identified in Aim 1.

Limitations: We acknowledge that our proposed study has a few limitations. First, we are only using data from Visit 3 onwards. Still, duration of follow-up for the current proposed study is still long (up to 17 years). Second, the SOMA is an aptamer based platform and the accuracy of protein identification is not always known. Third, power is limited for Mendelian randomization analyses, which requires us to evaluate associations in other, larger datasets such as CKD-Gen and the UK Biobank. Fourth, the renal toxicity of *APOL1* risk variants may be related to the APOL1 protein that is synthesized locally in the kidney rather than circulating plasma levels.[17,18]

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? __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/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)?

Manuscript #: 2370
Description: This published manuscript studied the association of race and *APOL1* risk status with various clinical outcomes including incident hypertension, incident diabetes, incident cardiovascular disease, incident ESRD, hospitalizations during follow-up, all-cause mortality, and eGFR decline from Visits 1 to 5.

Manuscript #: 3324
Whole Genome Sequence and Proteomics for Gene Discovery in the Atherosclerosis Risk in Communities (ARIC) Study
Description: This group will evaluate genetic determinants of the proteome using whole genome sequencing data.

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

___ A. primarily the result of an ancillary study (list number*_____)
_X_ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* __2011.03_(Selvin for funding on visit 6 labs, Matsushita for funding of visit 3 labs))

*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.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes __**X**__ No.

References

1. Tzur S, Rosset S, Shemer R, et al. Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Hum Genet.* 2010;128(3):345-350.
2. Genovese G, Friedman DJ, Ross MD, et al. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science.* 2010;329(5993):841-845.
3. Grams ME, Rebholz CM, Chen Y, et al. Race, APOL1 Risk, and eGFR Decline in the General Population. *J Am Soc Nephrol.* 2016;27(9):2842-2850.
4. Parsa A, Kao WH, Xie D, et al. APOL1 risk variants, race, and progression of chronic kidney disease. *N Engl J Med.* 2013;369(23):2183-2196.
5. Foster MC, Coresh J, Fornage M, et al. APOL1 variants associate with increased risk of CKD among African Americans. *J Am Soc Nephrol.* 2013;24(9):1484-1491.
6. Nichols B, Jog P, Lee JH, et al. Innate immunity pathways regulate the nephropathy gene Apolipoprotein L1. *Kidney Int.* 2015;87(2):332-342.
7. Beckerman P, Bi-Karchin J, Park AS, et al. Transgenic expression of human APOL1 risk variants in podocytes induces kidney disease in mice. *Nat Med.* 2017;23(4):429-438.
8. Grams ME, Surapaneni A, Ballew SH, et al. APOL1 Kidney Risk Variants and Cardiovascular Disease: An Individual Participant Data Meta-Analysis. *J Am Soc Nephrol.* 2019;30(10):2027-2036.
9. Ma L, Langefeld CD, Comeau ME, et al. APOL1 renal-risk genotypes associate with longer hemodialysis survival in prevalent nondiabetic African American patients with end-stage renal disease. *Kidney Int.* 2016;90(2):389-395.
10. Gutierrez OM, Irvin MR, Zakai NA, et al. APOL1 Nephropathy Risk Alleles and Mortality in African American Adults: A Cohort Study. *Am J Kidney Dis.* 2019.
11. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med.* 2009;150(9):604-612.
12. Inker LA, Schmid CH, Tighiouart H, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. *N Engl J Med.* 2012;367(1):20-29.
13. Rebholz CM, Selvin E, Liang M, et al. Plasma galectin-3 levels are associated with the risk of incident chronic kidney disease. *Kidney Int.* 2018;93(1):252-259.
14. Sumida K, Kwak L, Grams ME, et al. Lung Function and Incident Kidney Disease: The Atherosclerosis Risk in Communities (ARIC) Study. *Am J Kidney Dis.* 2017;70(5):675-685.
15. Tin A, Scharpf R, Estrella MM, et al. The Loss of GSTM1 Associates with Kidney Failure and Heart Failure. *J Am Soc Nephrol.* 2017;28(11):3345-3352.
16. O'Seaghdha CM, Parekh RS, Hwang SJ, et al. The MYH9/APOL1 region and chronic kidney disease in European-Americans. *Hum Mol Genet.* 2011;20(12):2450-2456.
17. Kozlitina J, Zhou H, Brown PN, et al. Plasma Levels of Risk-Variant APOL1 Do Not Associate with Renal Disease in a Population-Based Cohort. *J Am Soc Nephrol.* 2016;27(10):3204-3219.
18. Bruggeman LA, O'Toole JF, Ross MD, et al. Plasma apolipoprotein L1 levels do not correlate with CKD. *J Am Soc Nephrol.* 2014;25(3):634-644.