ARIC Manuscript Proposal #3510

1.a. Full Title: Associations between Proteomic Biomarkers and Risk of Abdominal Aortic Aneurysm

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

2. Writing Group:
Writing group members: Weihong Tang, James S. Pankow, Kunihiro Matsushita, Faye Norby, Ron C. Hoogeveen, Weihua Guan, Pamela L. Lutsey, Ching-Ping Hong, Ryan Demmer, Eric Boerwinkle, Joe Coresh, Aaron R. Folsom. Others are welcome

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

First author: Weihong Tang
Address: Division of Epidemiology and Community Health
University of Minnesota
Phone: 612-626-9140 Fax: 612-624-0315
E-mail: tang0097@umn.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: Aaron Folsom
Address: same as above
Phone: 612-626-8862 Fax: 612-624-0315
E-mail: folso001@umn.edu

3. Timeline: Finish by March 2020

4. Rationale:
Abdominal aortic aneurysm (AAA) is an important vascular disease in older adults and rupture of AAA is associated with a high mortality rate. The etiology of AAA is complex and not well understood. Among established cardiovascular risk factors, age, male sex, smoking, hypertension, and dyslipidemia increase the risk of AAA. The key pathological feature of AAA is progressive degradation and remodeling of extracellular matrix in the aortic wall.
Factors contributing to AAA pathogenesis include activation and perpetuation of inflammation, abnormal response of the innate or adaptive immune systems, up-regulation of matrix metalloproteinases (MMPs) and other proteinases, and impaired compensatory repair of extracellular matrix. In ARIC, we identified middle-age protein biomarkers for AAA, including circulating protein biomarkers of inflammation (CRP, fibrinogen, and IL-6), thrombin generation (D-dimer), vascular stiffness (N-terminal pro-brain natriuretic peptide), cardiac injury (troponin T), lipoprotein(a), and MMP-9. Participants with 4-6 elevated biomarkers had 10-fold higher risk than those with no biomarker elevations.

The SOMAscan v 4.0 that is being measured in ARIC includes 4,931 human proteins, belonging to a broad range of biological subgroups, of which cytokines, proteases, and protease inhibitors, among others, are highly relevant to the pathogenesis of AAA. Recently, researchers have applied SOMAscan to several diseases and successfully identified both established and novel biomarker signatures, including for muscular dystrophy and cardiovascular disease (CVD). Several small, cross-sectional studies of 4 to 20 participants using agnostic, proteomic scans for AAA in serum or plasma by mass-spectrometry reported previously known and some new biomarkers associated with AAA. However, the small sample sizes and use of liberal statistical thresholds preclude a clear interpretation of these findings. To date, there is no report of large, prospective population-based proteomic study of AAA. The substantial protein coverage in the SOMAscan v 4.0 and large number of AAAs in ARIC provide an excellent opportunity for discovery of protein risk markers for AAA.

5. Main Hypothesis/Study Questions:

We hypothesize that some proteomic biomarkers measured by the SOMAscan v 4.0 in ARIC will be associated positively with incidence of AAA over the next 20 years, and that we will be able to identify these protein biomarkers by both a targeted and agnostic approaches.

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).

We propose to conduct a prospective study of proteomic risk markers measured in SOMAscan v4.0 at visits 2 and 3 with incident, clinical AAA in ARIC. By event follow-up to 2012-16, we have identified a total of 645 clinical AAAs since visit 1 (partly contributed by the ARIC AAA ancillary study AS 2009.18). We will soon add, using identical methods, hospital AAA cases for 2017 (and 2018 if the data arrive early enough). At visit 5 (2011-12), an abdominal ultrasound examination identified additional 75 asymptomatic AAAs. With visit 2 as baseline and after exclusion of prevalent AAAs diagnosed before visit 2, there are 13,627 ARIC participants at risk of developing AAA, among whom we have identified 552 incident, clinical AAAs through 2016 and an additional 69 ultrasound AAAs. The analysis will start with visit 3 proteins (as baseline) and extend to visit 2 when the data become available. In a sensitivity analysis, we will analyze the average protein level of visits 2 and 3, using the visit 3 as the baseline for follow-up. The exposure is each of the 4,931 proteins measured by SOMAscan v4.0, with the exclusion of proteins that have large CV (e.g. > 20%), poor reproducibility between the blind duplicate pairs, non-specific binding, are non-human proteins, or extreme/unrealistic outliers. We will employ
Cox proportional hazards regression to investigate the associations of AAA with each protein, as in previous ARIC AAA publications. In the regression model, the dependent variable will be the time to first clinical AAA diagnosis with independent variables being the individual proteins, and potential confounders measured at ARIC visits 2 and 3. We will consider as possible confounders established risk factors for AAA, including age, gender, race, smoking, total cholesterol, HDL cholesterol, hypertension, and diabetes, etc, as well as measures of kidney function (e.g. estimated glomerular filtration rate or cystatin C). For top/significant proteins, we will assess the proportional hazards assumption by testing for interaction of each independent variable with survival time, corrected for the number of tests performed. In the event of significant violations, time-dependent Cox regression models will be used.

We will use the following strategy to prioritize the proteins:

a. Priority 1— targeted analysis of novel proteins encoded by genes reported in a large AAA genome-wide association study (GWAS), with gene expression data supporting their functional relevance to AAA: IL6R, SORT1, BCAR3, NOTCH2, FGF9, and PLTP. The cutpoint for statistical significance will be based on the alpha of 0.008 (0.05/6).

b. Priority 2— targeted analysis of SOMAscan proteins that have been associated with smoking in the published Framingham Heart Study (n=40 proteins). The analysis will be stratified by smoking status (ie never-smokers vs ever-smokers reported at baseline visit). The purpose is to identify smoking biomarkers that predict AAA risk as smoking is the strongest risk factor for AAA. The cutpoint for statistical significance will be based on the alpha of 0.001 (0.05/40).

c. Priority 3— agnostic discovery among the remaining 4,885 proteins, which we will screen for associations with AAA risk. The cutpoint for statistical significance will be based on the alpha of 1x10^-5.

For the proteins that are identified for clinical AAA, we will conduct a secondary analysis to evaluate their associations with the asymptomatic ultrasound AAAs ascertained from the visit 5 exam (n=69 ultrasound AAAs). In our previous study, some of the protein biomarkers identified for clinical AAA also showed consistent and significant associations with ultrasound-detected asymptomatic AAA (e.g., fibrinogen and MMP-9). This provides a more complete picture on protein biomarker-AAA associations.

We will develop prediction models to evaluate the prediction of AAA risk by known risk factors, including the previously detected blood biomarkers, and any new proteins identified in this proposal. We will use area under the ROC curve to evaluate if protein biomarkers improve risk prediction beyond traditional risk factors.

We will also conduct the following sensitivity analysis for clinical AAAs in ARIC: 1) remove AAAs who were diagnosed within 10 years from the baseline; 2) conduct the analysis separately in ever-smokers and never-smokers reported at baseline visit; 3) check if top proteins have known associations with medications (ie, medications for treating hypertension, high cholesterol, or diabetes). If so, we will consider additional adjustment or restriction in sensitivity analyses.
We will seek replication of significant findings from ARIC in external cohorts that have available the SOMAscan and AAA data (ie, HUNT3 and SCCS, in the process of applying for NIH funding)

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/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)?
    1505. Risk Factors for Abdominal Aortic Aneurysm (Tang)
    1505A. Hemostatic Factors and Aortic Aneurysm Incidence (Folsom)

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 (AS 2009.18: “Identifying Genetic and Epidemiological Risk Factors for Abdominal Aortic Aneurysm”, R01HL103695, PI Weihong Tang)

    _____ 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.

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: