ARIC Manuscript Proposal #3753

PC Reviewed: 12/8/20	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Circulating Proteome and Plasma Amyloid-ß Biomarkers

b. Abbreviated Title (Length 26 characters): proteomics of plasma amyloid-ß

2. Writing Group:

Writing group members: Adrienne Tin, Kevin Sullivan, Rajesh Talluri, Bing Yu, Keenan Walker, Jeanette Simino, B. Gwen Windham, Michael Griswold, Jan Bressler, Hao Mei, Myriam Fornage, Joe Coresh, Eric Boerwinkle, Tom Mosley

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: **Adrienne Tin**
Address: 2500 N State Street
Jackson, MS 39216

Phone: 601-496-9600 Fax:
E-mail: atin@umc.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: **Tom Mosley**
Address: 2500 N State Street
Jackson, MS 39216

Phone: 601-984-4467 Fax:
E-mail: tmosley@umc.edu

3. Timeline: Analysis will start immediately after aricpub approval with a draft manuscript for circulation in 12 months.

4. Rationale:

Blood levels of amyloid beta (Aß) peptides, namely $Aß_{1-40}$, $Aß_{1-42}$, and their ratio, have been proposed as biomarkers for the early detection of Alzheimer's disease (AD) pathology, particularly cerebral amyloidosis.[1-3] Although $Aß_{1-40}$ is the most abundant Aß peptide, the accumulation of $Aß_{1-42}$ fibrils in the brain are a neuropathological hallmark of AD.[4] Higher

plasma levels of Aß$_{1-42}$ and Aß$_{1-42}$ / Aß$_{1-40}$ ratio have been associated with a lower risk for AD or dementia[5,6], including work in ARIC (MP #3489), potentially reflecting increased clearance of Aß plaques from the central nervous system.

The determinants of Aß peptides in plasma are not well understood. These peptides are cleaved from the amyloid-beta precursor protein, whose encoding gene, *APP*, is expressed in a wide range of tissues.[7,8] These Aß peptides have been found in peripheral tissues, including arteries, platelets, skeletal muscles, and the liver.[9] The plasma levels of these Aß peptides can be influenced by the function of other organs. The association between these Aß peptides and circulating proteins may inform the pathways that regulates these Aß peptides and their role as plasma biomarkers of cerebral amyloidosis. In addition, recently the potential pathogenic role of these Aß peptides on peripheral organs, particularly in blood vessels and the heart, has been gaining attention.[10] Some hypothesize that these Aß peptides may link to arterial stiffness and the risk of cardiovascular disease in the context of aging.[10] Recent genetic studies suggest the levels of complex traits may be determined by gene networks involving thousands of variants in many genes across the genome.[11,12] A cross-sectional association study of the levels of these Aß peptides and the circulating proteome may identify genes and pathways related to the regulation of these Aß peptides and help to gain insight on the potential pathogenic role of these Aß peptides or their coregulation on diseases.

Visits 3 and 5 of the ARIC study has measures of these Aß peptides quantified using a sandwich immunoassay[13] in approximately 2,500 participants and a large number of plasma proteins (~5000) quantified using the SomaScan assay in the overall cohort.[14] A discovery study on the association between these SomaScan proteins and the plasma levels of the Aß peptides may inform the mechanisms that determine the levels of these Aß peptides, potentially identifying novel targets for interventions.

5. Main Hypothesis/Study Questions:

Study question: To identify plasma proteins that are associated with Aß$_{1-42}$, Aß$_{1-40}$, and the ratio of Aß$_{1-42}$ and Aß$_{1-40}$ peptides in midlife and late life. The associated proteins may be related to co-regulation of Aß clearance from the brain or the production of Aß from peripheral organs in these two periods of life. The associations discovered in midlife may suggest mechanisms leading to dementia, while the associations discovered in late life may suggest biomarkers of dementia versus normal cognition.

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 study at ARIC visits 3 and 5.
Plasma Aß$_{1-40}$ and Aß$_{1-42}$ peptides were measured in 2,588 participants, who were a subset of those selected using a case-cohort sampling that included (1) all participants who failed at least 1

cognitive domain at visit 5 and (2) an age-stratified random sample of cognitively normal participants.

The SomaScan assay are available for all participants who provided full consent and did not refuse industry research. So we anticipate the vast majority of the participants with measures of the Aß peptides also have SomaScan protein measures.

Inclusion criteria: All participants with measures in plasma Aß peptides, SomaScan proteins and covariates.

Dependent variable: SomaScan proteins. We will include those with flag2 = 0, i.e. human proteins that do not bind to related proteins with similar affinity, or have not been tested for binding specificity.

Independent variable: $Aß_{1-42}$, $Aß_{1-40}$, $Aß_{1-42}$ /$Aß_{1-40}$ ratio in separate analyses.

With 3 independent variables and SomaScan proteins at each visit, three sets of association analyses will be conducted at one visit for a total of six sets of association analyses across the two visits.

Other variables:
Demographics and genetics: age, sex, race-center, APOE genotype
Other variables at visit 5: smoking, BMI, CRP, triglycerides, total cholesterol, prevalent diabetes and hypertension, fasting glucose levels, eGFR, prevalent cardiovascular disease (coronary heart disease, stroke, heart failure),

Data analysis:

Controlling for unmeasured broad factors that affect plasma protein levels. We aim to identify proteins that are more specifically related to the regulation of Aß. Unmeasured factors, e.g. relative abundance, could give rise to spurious associations unrelated to Aß regulation. In addition, low molecular weight proteins tend to have inverse correlation with kidney function due reduced filtration, secretion, or catabolism by the kidney as kidney function decline.[15-17] Although estimated glomerular filtration rate (eGFR) is available in the ARIC study, due to measurement errors of eGFR in the intermediate range,[18] controlling for eGFR may not adequately adjust out the influence of kidney function on protein levels. We will use the SomaScan protein measures to generate probabilistic estimation of expression residuals (PEER) factors to represent board shared latent factors influencing protein levels.[19] The PEER method is based on factor analysis and has shown to be more sensitive to the underlying complexities of the data than other similar methods.[20] We will use factor relevance to select the appropriate number of PEER factors as recommended by the developers of this methods.[19]

Modeling of the SomaScan protein and Aß peptide measures. We will investigate appropriate transformation of the protein measures for linear regression, including log2, inverse normal transformation, and winsorization.

Covariates: age, sex, race-center, eGFR, and an appropriate number of PEER factors.

The limited number of covariates is motivated by:

1) Many commonly used covariates in dementia research, such as APOE4 genotype, BMI, and diabetes, etc. may be upstream effectors of the co-regulation pathways between some plasma proteins and Aß peptides. Including these variables would reduce the chance of identifying proteins in the co-regulation pathways.
2) PEER factors, which represent broad laden factors influencing protein levels, will be included.

Secondary analysis. For significant proteins, we will conducted secondary analysis that include additional covariates: APOE4 carrier status, cognition status, smoking, systolic blood pressure, BMI, CRP, triglycerides, total cholesterol, prevalent diabetes and hypertension, fasting glucose levels, prevalent cardiovascular disease (coronary heart disease, stroke, heart failure). The change in significance after controlling for additional covariates may help to identify upstream effectors of the association between a plasma protein and an Aß peptide.

Significant threshold. We will use the Bonferroni method to determine significance threshold for each set of analysis (~ 0.05/5000 = 1e-5).

Associated proteins will be prioritized by binding specificity. Tier 1 association will be those that have been tested and did not bind to related proteins. The rest will be tier 2.

Power analysis. Based on a type-I error rate of 1e-5 and a sample size of 2,000, we have 80% power to detect ≥ 0.13 SD difference in the plasma protein levels per SD difference in Aß measures.

Follow-up analysis. We will investigate the following follow-up analysis:

1) Enrichment analysis. We will perform follow-up analysis to identify the gene sets or pathways that are enriched for the associated proteins. Tools for performing these analyses include Database for Annotation, Visualization and Integrated Discovery (DAVID)[21,22] and Ingenuity Pathway Analysis (IPA), etc. These analyses can help to interpret the associations and generate hypotheses on potential mechanisms.
2) Association analysis of the Aß-associated proteins with Aß burden detected in PET scan in ARIC.

Limitations:

1) External validation**.** We will seek other cohorts for external validation.
2) If some proteins are related to Aß regulation and have non-linear relation with the Aß variables, this will be considered in other future manuscript proposal.

3) The potential relation between changes in the levels of the Aß peptides from visit 3 to visit 5 and the SomaScan proteins will be considered in other future manuscript proposal.

4) Given that the measures of the SomaScan proteins are from relative quantification, studies on the change in RFU of the SomaScan proteins from visit 3 to visit 5 need to consider a) the potential effects of differences in the assay technique details between the two visits on RFU levels, and b) the interpretation of the difference between RFU values between the two visits. Studies of the change in RFU of the SomaScan proteins from visit 3 to visit 5 will be considered in other manuscript proposals.

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/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)?

3489 Plasma Amyloid-Beta and Risk of MCI/Dementia in ARIC-NCS

3354 Plasma beta-amyloid and late-onset epilepsy: The ARIC Neurocognitive Study

2644 Whole Exome Sequence Analysis of Plasma Amyloid-ß in African and European Americans; the Atherosclerosis Risk in Communities-Neurocognitive Study

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

ARIC NCS

11.b. If yes, is the proposal

_X_ A. primarily the result of an ancillary study (list number* _ARIC Hemostatic factor and NCS ancillary studies**____)**

___ 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.

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 ____ No.

References

1. Nakamura A, Kaneko N, Villemagne VL, et al. High performance plasma amyloid-beta biomarkers for Alzheimer's disease. Nature 2018;554:249-54.
2. Vergallo A, Megret L, Lista S, et al. Plasma amyloid beta 40/42 ratio predicts cerebral amyloidosis in cognitively normal individuals at risk for Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association 2019;15:764-75.
3. Ritchie C, Smailagic N, Noel-Storr AH, et al. Plasma and cerebrospinal fluid amyloid beta for the diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev 2014:CD008782.
4. Mori H, Takio K, Ogawara M, Selkoe DJ. Mass spectrometry of purified amyloid beta protein in Alzheimer's disease. J Biol Chem 1992;267:17082-6.
5. Koyama A, Okereke OI, Yang T, Blacker D, Selkoe DJ, Grodstein F. Plasma amyloid-beta as a predictor of dementia and cognitive decline: a systematic review and meta-analysis. Arch Neurol 2012;69:824-31.
6. Chouraki V, Beiser A, Younkin L, et al. Plasma amyloid-beta and risk of Alzheimer's disease in the Framingham Heart Study. Alzheimer's & dementia : the journal of the Alzheimer's Association 2015;11:249-57 e1.
7. Muller UC, Deller T, Korte M. Not just amyloid: physiological functions of the amyloid precursor protein family. Nat Rev Neurosci 2017;18:281-98.
8. Olsson F, Schmidt S, Althoff V, et al. Characterization of intermediate steps in amyloid beta (Abeta) production under near-native conditions. J Biol Chem 2014;289:1540-50.

9. Roher AE, Esh CL, Kokjohn TA, et al. Amyloid beta peptides in human plasma and tissues and their significance for Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association 2009;5:18-29.
10. Stakos DA, Stamatelopoulos K, Bampatsias D, et al. The Alzheimer's Disease Amyloid-Beta Hypothesis in Cardiovascular Aging and Disease: JACC Focus Seminar. Journal of the American College of Cardiology 2020;75:952-67.
11. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell 2017;169:1177-86.
12. Liu X, Li YI, Pritchard JK. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. Cell 2019;177:1022-34.e6.
13. Blennow K, De Meyer G, Hansson O, et al. Evolution of Abeta42 and Abeta40 levels and Abeta42/Abeta40 ratio in plasma during progression of Alzheimer's disease: a multicenter assessment. J Nutr Health Aging 2009;13:205-8.
14. Gold L, Ayers D, Bertino J, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS One 2010;5:e15004.
15. Yang J, Brody Edward N, Murthy Ashwin C, et al. Impact of Kidney Function on the Blood Proteome and on Protein Cardiovascular Risk Biomarkers in Patients With Stable Coronary Heart Disease. Journal of the American Heart Association 2020;9:e016463.
16. Maack T, Johnson V, Kau ST, Figueiredo J, Sigulem D. Renal filtration, transport, and metabolism of low-molecular-weight proteins: a review. Kidney Int 1979;16:251-70.
17. Mogielnicki RP, Waldmann TA, Strober W. Renal handling of low molecular weight proteins. I. L-Chain metabolism in experimental renal disease. The Journal of clinical investigation 1971;50:901-9.
18. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150:604-12.
19. Stegle O, Parts L, Piipari M, Winn J, Durbin R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat Protoc 2012;7:500-7.
20. Stegle O, Parts L, Durbin R, Winn J. A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. PLoS Comput Biol 2010;6:e1000770.
21. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009;37:1-13.
22. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44-57.