ARIC Manuscript Proposal #2644

PC Reviewed: 10/13/15	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Whole Exome Sequence Analysis of Plasma Amyloid-β in African and European Americans; the Atherosclerosis Risk in Communities-Neurocognitive Study

b. Abbreviated Title (Length 26 characters): WES analysis of plasma aβ

2. Writing Group: Jeannette Simino, Zhiying Wang, Jan Bressler, Stephen G. Younkin, Myriam Fornage, Vincent Chouraki, Sudha Seshadri, Eric Boerwinkle, and Thomas Mosley. Others are welcome.

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

First author: Jeannette Simino
Address: 2500 N. State Street, Room G551-04
Jackson, MS 39216

Phone: (601) 815-8128
E-mail: jsimino@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: Thomas Mosley
Address: Office Annex I-Room 216
2500 N. State Street
Jackson, MS 39216

Phone: (601) 984-2763
E-mail: tmosley@umc.edu

3. Timeline:

All data is currently available and the analysis pipeline is complete. We plan to submit the manuscript for publication in less than three months from the date of approval.

4. Rationale:

Alzheimer's disease (AD) is a major public health burden, afflicting 5.1 million Americans aged 65 or older and accounting for nearly twenty percent of Medicare expenditures [1]. Demographic changes will cause the number of affected elderly to triple by 2050, costing the nation $1.1 trillion [1]. Although there are no effective treatments to prevent, slow, or cure AD, researchers have made great strides in dissecting its genetic etiology. The largest (>74,000 participants) and most-recent genome-wide association study (GWAS) of late-onset AD (LOAD) revealed twenty loci [2] containing common variants with small effect sizes (except for the APOE ε4 allele). These variants collectively explained a small proportion of the LOAD heritability [3], inspiring researchers to conduct targeted, whole-exome, and whole-genome sequencing studies to identify a handful of rare and low-frequency variants with large effect sizes [4-13]. These studies implicated variants from the whole allele frequency spectrum [14] and provided insight into the diverse biological processes and pathways involved in LOAD. Unfortunately, the identification of additional variants is constrained by the number of genotyped cases and controls currently available [6].

New approaches, such as the use of AD biomarkers or neuroimaging measures, may facilitate the identification of both common and rare candidate drug targets from a moderate number of sequenced participants. Given a fixed sample size, endophenotypes have a greater statistical power than diagnostic outcomes because they are quantitative, objective, and closer to the level of gene action [15, 16]. Endophenotypes may also be influenced by greater genetic effects and allow analysis of all participants without regard to disease status (i.e. do not restrict the analysis to AD cases and cognitive normals) [15, 16]. A few GWAS have successfully leveraged the two hallmark characteristics of AD, the prevalence of β-amyloid plaques outside brain neurons and the accumulation of tau tangles inside brain neurons[1], as endophenotypes [17-22]. We will employ two less-invasive and less-expensive plasma biomarkers associated with AD [23]: amyloid β42 ($a\beta_{42}$) levels and the ratio of $a\beta_{42}$ to amyloid β40 ($a\beta_{40}$). These plasma amyloid traits are ascertained from blood, exhibit a large heritability (50-80%)[24], and may indicate early amyloid pathogenesis preceding the plaque cascade. Furthermore, measuring amyloids in stored blood from past examinations enables an assessment of age-specific genetic effects and a refined granularity of biological changes occurring over time.

Our aim is to identify common and rare variants that influence plasma amyloids in late middle age, old age, or their changes over time. We will accomplish this aim through a whole exome sequence analysis of plasma amyloids using 1,414 participants (406 African Americans and 1,008 European Americans) from the Atherosclerosis Risk in Communities-Neurocognitive Study (ARIC-NCS). This study is particularly well-suited for this investigation for several reasons. First, participants had two amyloid measurements spaced an average of 18 years apart, with mean ages of 59 and 77 years for

the two blood draws. Second, ARIC-NCS is enriched for atypical cognition, and hence should be enriched for variants contributing to neurocognitive impairment and the amyloid plaque cascade. Third, the inclusion of African Americans allows exploration of rare population-specific variants with large effects which may contribute to health disparities. Overall, this investigation will provide insight into the influence of age on amyloid gene discovery efforts.

5. Main Hypothesis/Study Questions:

Our aim is to identify common and rare exomic variants that influence plasma amyloids ($a\beta_{42}$ levels and the $a\beta_{42}$: $a\beta_{40}$ ratio) in late middle age, old age, or their changes over time. We will compare and contrast the findings in the different life phases to explore the impact of age on plasma amyloid gene discovery efforts in African and European Americans.

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 gene-discovery study at both the third and fifth visits. In addition, the fold-change in amyloids between visits will be used as a trait in the whole exome sequence analysis.

Exclusion Criteria:

- Participants missing amyloids or whole exome sequence
- Race other than African or European American
- Participants missing covariates (listed below)
- Individuals who did not consent to DNA use

Outcomes (all in plasma):

1) $a\beta_{42}$ levels at visit 3
2) $a\beta_{42}$ levels at visit 5
3) fold-change in $a\beta_{42}$ levels from visit 3 to visit 5
4) $a\beta_{42}$: $a\beta_{40}$ ratio at visit 3
5) $a\beta_{42}$: $a\beta_{40}$ ratio at visit 5
6) fold-change in $a\beta_{42}$: $a\beta_{40}$ ratio from visit 3 to 5

Covariates:

- visit-specific age (in years)
- gender
- apolipoprotein E ε4 carrier status (dichotomous yes/no; ascertained from TaqMan assays)

- time elapsed between visits (in years; included with the third visit age when modeling the fold-changes in plasma amyloids only)
- Eigenstrat-derived principal components to adjust for population substructure

Exome Sequence:

Of the participants with amyloids, 1,414 (406 AAs, 1,008 EAs) had whole exome sequence and all covariates available. This sequence data was quality-controlled as part of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. A full annotation file and description of the methods is available.

Statistical Methods:

Single-visit amyloid measures will be linearly regressed onto age, gender, and *APOE* ε4 carrier status separately by race. The fold-change amyloid traits will be regressed onto age at the third visit, the time elapsed between visits, gender, and APOE ε4 carrier status separately by race. Following rank-based inverse normal transformation of the residuals, we will use seqMeta version 1.5 to conduct ethnic-specific single-variant and gene-based association tests while adjusting for population structure (the first ten PCs). Linear regression models will be fit on all autosomal SNPs with minor allele frequency (MAF) ≥ 0.01 while Sequence Kernel Association Tests (SKAT) and T5 gene-based tests will be conducted on nonsynonymous, splicing, stopgain, stoploss, or frameshift autosomal variants with MAF ≤ 0.05. For both the single-variant and gene-based tests, we will conduct a meta-analysis of the African and European American results and apply a Bonferonni correction for the number of variants or genes tested.

Replication:

Sudha Seshadri from the Framingham Heart Study (FHS) has agreed to replicate our single-visit findings. FHS has multiple generations and can tailor replication of each significant variant or gene to the age group that mirrors our ARIC visit. Given the paucity of African American amyloid data, we may attempt replication of ethnic-specific findings in ARIC participants that are missing exome sequence data but have exome chip or 1000G GWAS data; however, there is no guarantee that the variants of interest will be genotyped and/or of high imputation quality.

Limitations:

- No readily available replication samples for variants or genes (if any) contributing to amyloid fold-changes over time.

- We do not have a representative sample. Plasma amyloids were retroactively measured in a sample of participants that survived to visit 5. Thus those with amyloids measured at the third visit have a survivor bias. As this is a gene-discovery project, no weighting will be used in the primary analysis but replication in the FHS will be of the upmost importance.

- Amyloids were measured on 69 plates; we will perform a sensitivity analysis to ascertain whether adjustment for these batch effects alters any significant associations. We will not correct for plate effects in the primary analysis because the small number of participants per plate may introduce too much statistical noise. The largest published GWAS of plasma amyloid-β to date ignored such plate effects [25].

- Some amyloids were below the assay detection limit. For the single-visit analyses of $a\beta_{42}$, samples with intensities below the minimal detectable level will be assigned the threshold concentrations (12 pg/ml). These individuals will be omitted from the visit-specific analysis of the $a\beta_{42}$:$a\beta_{40}$ ratio and the fold-changes in both $\alpha\beta_{42}$ and the ratio since the ranks of those with missing data relative to those with nonmissing values are unknown.

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)?

Genetics of Cognitive Traits:

MSP 1704: Genetic variants identified in genome-wide association studies of dementia and cognitive change in middle age: The Atherosclerosis Risk in Communities (ARIC) study (Lead author: Jan Bressler)

MSP 2079: Genome-wide studies of human verbal declarative memory: The CHARGE consortium (Lead author: Jan Bressler)

MSP 2080: Genome-wide association studies for executive function and processing speed indicate a role for genes in neurotransmission (Lead author: Jan Bressler)

Investigations Using Amyloid-β:

MSP 2511: Vascular risk factors and brain amyloid deposition: The ARIC-PET Study (Lead author: Rebecca Gottesman)

MSP 2466: The ARIC-PET Amyloid Imaging Study: Differences in brain amyloid deposition by age, race, sex, and apoE genotype (Lead author: Rebecca Gottesman)

MSP 2544: Arterial Stiffness and β-amyloid deposition in the ARIC-PET Study (Lead author Timothy M Hughes)

Exome Sequence Reference:

MSP 2344: Evaluation of whole exome sequence for loss of function variation influencing quantitative traits (lead author: Alexander Li)

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* 2008.06**)**

___ 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. Accepted.

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. Accepted.

13. Per Data Use Agreement Addendum for the Use of Linked ARIC CMS Data, approved manuscripts using linked ARIC 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. Alzheimer's, A., *2015 Alzheimer's disease facts and figures.* Alzheimers Dement, 2015. **11**(3): p. 332-84.
2. Lambert, J.C., C.A. Ibrahim-Verbaas, D. Harold, A.C. Naj, R. Sims, et al., *Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease.* Nat Genet, 2013. **45**(12): p. 1452-8.
3. Calero, M., A. Gomez-Ramos, O. Calero, E. Soriano, J. Avila, et al., *Additional mechanisms conferring genetic susceptibility to Alzheimer's disease.* Front Cell Neurosci, 2015. **9**: p. 138.
4. Vardarajan, B.N., M. Ghani, A. Kahn, S. Sheikh, C. Sato, et al., *Rare coding mutations identified by sequencing of Alzheimer disease genome-wide association studies loci.* Ann Neurol, 2015.
5. Steinberg, S., H. Stefansson, T. Jonsson, H. Johannsdottir, A. Ingason, et al., *Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease.* Nat Genet, 2015. **47**(5): p. 445-7.
6. Cruchaga, C., C.M. Karch, S.C. Jin, B.A. Benitez, Y. Cai, et al., *Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease.* Nature, 2014. **505**(7484): p. 550-4.
7. Jonsson, T., H. Stefansson, S. Steinberg, I. Jonsdottir, P.V. Jonsson, et al., *Variant of TREM2 associated with the risk of Alzheimer's disease.* N Engl J Med, 2013. **368**(2): p. 107-16.
8. Guerreiro, R., A. Wojtas, J. Bras, M. Carrasquillo, E. Rogaeva, et al., *TREM2 variants in Alzheimer's disease.* N Engl J Med, 2013. **368**(2): p. 117-27.
9. Jonsson, T., J.K. Atwal, S. Steinberg, J. Snaedal, P.V. Jonsson, et al., *A mutation in APP protects against Alzheimer's disease and age-related cognitive decline.* Nature, 2012. **488**(7409): p. 96-9.
10. Vardarajan, B.N., Y. Zhang, J.H. Lee, R. Cheng, C. Bohm, et al., *Coding mutations in SORL1 and Alzheimer disease.* Ann Neurol, 2015. **77**(2): p. 215-27.
11. Kim, M., J. Suh, D. Romano, M.H. Truong, K. Mullin, et al., *Potential late-onset Alzheimer's disease-associated mutations in the ADAM10 gene attenuate {alpha}-secretase activity.* Hum Mol Genet, 2009. **18**(20): p. 3987-96.

12. Logue, M.W., M. Schu, B.N. Vardarajan, J. Farrell, D.A. Bennett, et al., *Two rare AKAP9 variants are associated with Alzheimer's disease in African Americans.* Alzheimers Dement, 2014. **10**(6): p. 609-618 e11.
13. Wetzel-Smith, M.K., J. Hunkapiller, T.R. Bhangale, K. Srinivasan, J.A. Maloney, et al., *A rare mutation in UNC5C predisposes to late-onset Alzheimer's disease and increases neuronal cell death.* Nat Med, 2014. **20**(12): p. 1452-7.
14. Lord, J., A.J. Lu, and C. Cruchaga, *Identification of rare variants in Alzheimer's disease.* Front Genet, 2014. **5**: p. 369.
15. Glahn, D.C., E.E. Knowles, D.R. McKay, E. Sprooten, H. Raventos, et al., *Arguments for the sake of endophenotypes: examining common misconceptions about the use of endophenotypes in psychiatric genetics.* Am J Med Genet B Neuropsychiatr Genet, 2014. **165B**(2): p. 122-30.
16. Walters, J.T. and M.J. Owen, *Endophenotypes in psychiatric genetics.* Mol Psychiatry, 2007. **12**(10): p. 886-90.
17. Ramanan, V.K., S.L. Risacher, K. Nho, S. Kim, S. Swaminathan, et al., *APOE and BCHE as modulators of cerebral amyloid deposition: a florbetapir PET genome-wide association study.* Mol Psychiatry, 2014. **19**(3): p. 351-7.
18. Han, M.R., G.D. Schellenberg, L.S. Wang, and I. Alzheimer's Disease Neuroimaging, *Genome-wide association reveals genetic effects on human Abeta42 and tau protein levels in cerebrospinal fluids: a case control study.* BMC Neurol, 2010. **10**: p. 90.
19. Kim, S., S. Swaminathan, L. Shen, S.L. Risacher, K. Nho, et al., *Genome-wide association study of CSF biomarkers Abeta1-42, t-tau, and p-tau181p in the ADNI cohort.* Neurology, 2011. **76**(1): p. 69-79.
20. Cruchaga, C., J.S. Kauwe, O. Harari, S.C. Jin, Y. Cai, et al., *GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease.* Neuron, 2013. **78**(2): p. 256-68.
21. Ramirez, A., W.M. van der Flier, C. Herold, D. Ramonet, S. Heilmann, et al., *SUCLG2 identified as both a determinator of CSF Abeta1-42 levels and an attenuator of cognitive decline in Alzheimer's disease.* Hum Mol Genet, 2014. **23**(24): p. 6644-58.
22. Ramanan, V.K., S.L. Risacher, K. Nho, S. Kim, L. Shen, et al., *GWAS of longitudinal amyloid accumulation on 18F-florbetapir PET in Alzheimer's disease implicates microglial activation gene IL1RAP.* Brain, 2015. **138**(Pt 10): p. 3076-88.
23. Chouraki, V., A. Beiser, L. Younkin, S.R. Preis, G. Weinstein, et al., *Plasma amyloid-beta and risk of Alzheimer's disease in the Framingham Heart Study.* Alzheimers Dement, 2015. **11**(3): p. 249-57 e1.
24. Ertekin-Taner, N., N. Graff-Radford, L.H. Younkin, C. Eckman, M. Baker, et al., *Linkage of plasma Abeta42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees.* Science, 2000. **290**(5500): p. 2303-4.
25. Chouraki, V., R.F. De Bruijn, J. Chapuis, J.C. Bis, C. Reitz, et al., *A genome-wide association meta-analysis of plasma Abeta peptides concentrations in the elderly.* Mol Psychiatry, 2014. **19**(12): p. 1326-35.