ARIC Manuscript Proposal #2529

PC Reviewed: 4/14/15	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: The role of mitochondrial heteroplasmy and genetic variation in successful aging

b. Abbreviated Title (Length 26 characters): mtDNA heteroplasmy and survival

2. Writing Group:
Jeremy Walston
Foram Ashar
B Gwen Windham
Eliseo Guallar
Josef Coresh
Megan Grove
Eric Boerwinkle
Additional interested ARIC investigators

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

F**irst author**: **Foram N. Ashar**
Address: 733 N Broadway, BRB 515
Baltimore, MD 21205

Phone: (443) 287-0251 Fax: (410) 614-8600
E-mail: fashar1@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: **Dan E. Arking**
Address: 733 N Broadway, BRB 447
Baltimore, MD 21205

Phone: (410) 502-4867 Fax: (410) 614-8600
E-mail: arking@jhmi.edu

3. Timeline: The initial pipeline to determine heteroplasmy and full mtDNA sequence is currently being implemented. We expect that initial data will be ready for analysis by Fall 2015. Analysis will take 3-6 months, and a first manuscript should be ready for Spring/Summer 2016.

4. Rationale:
Energy metabolism has long been proposed to play a critical role in human disease and aging, and mitochondria are the key organelle involved in energy production (oxidative phosphorylation, OXPHOS). Studying the role of mitochondrial function in human disease is complicated by two major features leading to increased variation beyond inherited genetic variation: mitochondrial DNA copy number (mtDNA-CN) and heteroplasmy. mtDNA-CN reflects the fact that each cell contains 100s of mitochondria, with a high degree of variation in the number of mitochondrial genomes within each individual mitochondrial cell, and in the number of mitochondria per cell [1]. Heteroplasmy, which is universally observed [2], is largely a function of the higher mutation rate that mitochondria are subject to relative to the nuclear genome (~70% somatic, ~30% inherited), leading to an accumulation of heteroplasmy with age [3].

We have previously developed methods to determine mtDNA-CN from existing ARIC genotyping arrays, and demonstrated that mtDNA-CN measured in peripheral blood cells declines longitudinally with age and is associated with general health among the elderly, and ultimately, mortality (age- and sex-adjusted relative risk comparing the lowest to the highest quintiles mtDNA-CN of 1.47 (95% CI 1.33-1.63, $P=4.24x10^{-14}$) [4]. The entire ARIC cohort is currently undergoing whole-exome (WES) and/or whole-genome sequencing (WGS) (WES/WGS is already available on >6,000 samples, and the remaining ~9,000 will be released over the next 6 months). Recent tools have been developed to extract mtDNA sequence from WES/WGS, moreover, with the deep sequencing coverage (>1000x), we are also able to detect heteroplasmy (often due to somatic mutation) down to ~0.5% [5].

5. Main Hypothesis/Study Questions:
Mitochondrial variation, in the form of genetic variation and heteroplasmy, influences the risk for healthy aging and overall mortality.

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).
Full mtDNA sequence and heteroplasmy will be determined using the MToolBox automated software pipeline [5]. Briefly, this pipeline assembles the mtDNA sequence and calculates the heteroplasmic fraction (HF) for each variant site using mitochondrial reads extracted from WES/WGS data. A key step is the removal of reads that map to nuclear copies of the mitochondria (NumtS), along with incorporation of stringent quality filters, resulting in 60% sensitivity and 100% specificity for HF down to 0.5% with 1000x coverage, and 100% sensitivity and specificity when reads reach 2000x.

Specific Aim 1: Determine the association of mtDNA genetic variants with frailty, physical function performance, CVD, and overall mortality. With the full mitochondrial genome sequence available, we will analyze both common (standard single SNP tests) and rare (SKAT) inherited genetic variation with two measures of successful aging, frailty and HAI, as well as

overall mortality. Frailty is considered a clinical syndrome of decreased function and resistance to stressors, due to cumulative declines across multiple physiologic systems that culminates in increased vulnerability to adverse outcomes. It is operationalized in ARIC based upon the Fried et al criteria [6], with slight modification to energy expenditure component based on available ARIC data, and incorporates measures of physical function, weight loss, strength, activity, and energy levels. It is considered a syndrome related to, but distinct from, disability. Consistent with existing literature, the ARIC frailty definition is associated with older age, female sex, and black race. The Short Physical Performance Battery (SPPB) is a validated measure of lower extremity performance associated with risk of incident disability, institutionalization, falls, mortality and other adverse outcomes in older adults.[7,8] These measures are available from ARIC visit 5 (~6,500 individuals We will also assess whether genetic variation is a predictor of enrollment in visit 5 (i.e. those who survive and are well-enough to participate). For frailty and SPPB, we will use standard exclusion criteria (not self-identified as White or Black, genetic outliers as determined from GWAS data, first-degree relatives) and include covariates for adjustment (age, sex, principal components from GWAS, study center) in race-stratified analysis.

Survival analyses methods will be employed to evaluate potential differences in the incidence of CVD and mortality. After model construction steps, to accommodate the survival outcome and potential for sex specific differences, we anticipate that final models will incorporate age, sex, and be stratified on self-reported race to address population substructure. The magnitude, direction, and statistical significance of coefficients will be assessed. We will also examine the relationship between mtDNA genetic variation and CVD and mortality, incorporating known CVD risk factors, as part of a mediation analysis to see whether mtDNA variants are acting through specific pathways.

All analyses will be performed assuming a homoplasmic state (sites with heteroplasmy greater than 25% will be coded as missing), as heteroplasmy will be specifically tested in Aim 2.

Specific Aim 2: Determine the association of the heteroplasmy with frailty, SPPB, CVD, and overall mortality. Using the methods described above, for each heteroplasmic site, we will test the association of HF with the phenotype of interest. We will also assess the overall level of heteroplasmy across the mitochondrial genome (i.e. percentage of mitochondrial genomes carrying a mutation, allowing for different mutations on different mitochondrial chromosomes within an individual), as an explicit test of whether heteroplasmic burden is a risk factor for frailty, SPPB, and/or overall mortality.

Specific Aim 3: Determine whether there is an epistatic interaction between mtDNA-CN and heteroplasmy on the association with frailty, the SPPB, and/or overall mortality. One potential consequence of somatic mutation, and the resultant decrease in mitochondrial function, is an upregulation in the production of mitochondria. Thus, while we have demonstrated, that on average, increased mtDNA-CN is protective for overall survival, we may be significantly underestimating the impact of mtDNA-CN by not taking into account the quality of the mitochondria. Specifically, one could imagine that high mtDNA-CN in the presence of high heteroplasmy may actually be harmful rather than protective for survival. Thus, we propose to combine these measures in association analyses, including a specific interaction term to test this hypothesis.

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)?
Our group has several manuscript proposals related to mtDNA and CVD/frailty/mortality.

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* 2009.14)
___ 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.

References

1. Chen, X. J. & Butow, R. A. The organization and inheritance of the mitochondrial genome. *Nat. Rev. Genet.* **6,** 815–825 (2005).
2. Payne, B. A. I. *et al.* Universal heteroplasmy of human mitochondrial DNA. *Hum. Mol. Genet.* **22,** 384–390 (2013).
3. Sondheimer, N. *et al.* Neutral mitochondrial heteroplasmy and the influence of aging. *Hum. Mol. Genet.* **20,** 1653–1659 (2011).
4. Ashar, F. N. *et al.* Association of mitochondrial DNA levels with frailty and all-cause mortality. *J. Mol. Med. Berl. Ger.* **93,** 177–186 (2015).
5. Calabrese, C. *et al.* MToolBox: a highly automated pipeline for heteroplasmy annotation and prioritization analysis of human mitochondrial variants in high-throughput sequencing. *Bioinforma. Oxf. Engl.* **30,** 3115–3117 (2014).
6. Fried, L. P. *et al.* Frailty in older adults: evidence for a phenotype. *J. Gerontol. A. Biol. Sci. Med. Sci.* **56,** M146–156 (2001).
7. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-Extremity Function in Persons over the Age of 70 Years as a Predictor of Subsequent Disability. *The New England journal of medicine.* 1995;332(9):556-562.
8. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol.* 1994;49(2):M85-94.