1.a. Full Title: Association between Mitochondrial DNA Copy Number and atrial fibrillation: Findings from the Atherosclerosis Risk in Communities Study (ARIC)

b. Abbreviated Title (Length 26 characters): mtDNA copy number and AF

2. Writing Group:
   Eliseo Guallar
   Ryan J. Longchamps
   Christina A. Castellani
   Yunsoo Hong
   Alvaro Alonso
   Dan E. Arking

   additional interested ARIC investigators

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

First author: Di Zhao
Address: 2024 E Monument St. 2-635
Baltimore, MD 21205

E-mail: dizhao@jhu.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: We anticipate a manuscript will be ready by the end of summer 2017.

4. Rationale:
Atrial fibrillation (AF) is the most common form of clinical cardiac arrhythmia, with rising prevalence and incidence worldwide due to aging population. It is estimated that the number of adults with AF in the US will double by year 2050, affecting more than 8 million people. AF is associated with increased risk of stroke, heart failure, mortality, and imposes tremendous healthcare costs. Consequently, it is critical to identify effective prevention or novel treatment of this disease and its complications.

Mitochondria are the organelles that generate energy for the cell through converting nutrients and oxygen into adenosine triphosphate (ATP). Unlike other organelles, mitochondria has its own circular DNA (mtDNA), which is essential in encoding the oxidative phosphorylation process. Each cell contains on average $10^3$ to $10^4$ copies of mtDNA, with variations of cell type and development phase. Mitochondrial DNA copy number (mtDNA-CN) is proportional to transcription of mitochondrial genes and a determinant of mitochondrial function. The decrease of mtDNA-CN is a marker of mtDNA dysfunction, which is related to adverse cardiovascular events including heart failure, left ventricular function, arrhythmias, sudden cardiac death, as well as risk factors for cardiovascular disease including hypertension, diabetes, atherosclerosis and chronic kidney disease.

Emerging evidence suggests that mtDNA dysfunction is associated with increased risk of AF through reduced ATP production and elevated reactive oxidative species (ROS). However, these results were mostly from molecular studies. The association between mtDNA-CN and incident AF in the general population remains unknown. In the present study, we will examine the prospective association between baseline mtDNA-CN and the risk of AF among participants from the Atherosclerosis Risk in Communities (ARIC) study.

5. Main Hypothesis/Study Questions:

We hypothesize that mtDNA copy number will be a significant predictor of atrial fibrillation.

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: prospective study.

Inclusion/exclusion: We will include Whites and Blacks for whom DNA is available for mtDNA-CN estimation. We will exclude non-whites and non-blacks, Blacks from Minnesota or Maryland center, as well as participants with prevalent AF at the time of mtDNA-CN measurements, and participants missing AF data or other covariates.

Measurement of AF events
The method for ascertainment of AF has been described previously. Briefly, AF cases were identified three ways: electrocardiograms (ECGs) performed during study visits, review of hospital discharge codes, and death certificates. At each study exam, a supine 12-lead resting ECG was performed and transmitted to the ARIC ECG Reading Center (Epidemiological Cardiology Research Center, Wake Forest School of Medicine, Winston Salem, NC) for automatic coding with E Marquette 12-SL program (GE Marquette, Milwaukee, WI). AF or atrial flutter was detected automatically by computer and confirmed by a cardiologist.

Hospitalization information during follow-up was obtained through annual follow-up phone calls and surveillance of local hospitals. Trained abstractors collected hospital discharge codes. The presence of AF was identified if the ICD-9-CM codes 427.31 (AF) or 427.32 (atrial flutter) were listed. AF cases detected in the same hospitalization with open cardiac surgery were excluded. Identification of AF from hospital discharge codes has proven to be valid method in epidemiological studies.

Finally, AF was identified from death certificates if ICD-9 427.3 or ICD-10 I48 were listed as a cause of death. The AF date was determined as the date of the first ECG with AF (4%), the time of first hospital discharge with AF codes (96%), or when AF was listed as a cause of death (0.1%), whichever occurred first.

Measurement of mtDNA Copy Number

DNA samples were isolated from buffy coat and gynotyped on the Affymetrix Genome-Wide Human SNP Array 6.0 (www.genesis.org). Mitochondrial SNPs were collected across all samples and were signaled with 25 high-quality mitochondrial probes. Raw mtDNA CN was determined by the median probe intensity difference across all mitochondrial SNPs. To correct for technical artifacts, batch effects, DNA quality, and starting DNA quantity, surrogate variable analysis was applied to probe intensities of 43,316 autosomal SNPs. We calculated residuals using a linear regression model with raw mtDNA CN as the dependent variable and the surrogate variable, age, sex, and enrollment center as independent variable. The calculated residuals were then used as measurement for mtDNA-CN for all subsequent analyses.

Statistical Analyses

DNA for mtDNA-CN analysis was collected in visit 1 (1987-1989) for 484 participants (4.2%), visit 2 (1990-1992) for 9,112 participants (79.6%), visit 3 (1993-1995) for 1,791 participants (15.6%), and visit 4 (1996-1998) for 66 participants (0.6%). The visit of DNA collection in each participant will be considered the baseline visit and all covariates will be obtained from that visit. Follow-up for events started from the baseline visit and continued until incident AF or through December 31, 2014, whichever came first.

Baseline characteristics of the study population will be compared across quintiles of mtDNA-CN. We will use Cox proportional hazards model to estimate hazard ratios (HR) and 95% confidence intervals (CI) for the association between mtDNA-CN and AF. In the primary analysis, we will categorize mtDNA-CN into quintiles based on the
sample distribution. Tests for linear trend across quintiles will be conducted. In secondary analysis, we will model mtDNA as a continuous variable and estimated the HR for AF comparing the 10th to the 90th percentile of mtDNA-CN. Additionally, we will model mtDNA-CN as restricted cubic splines with knots at the 5th, 35th, 65th and 95th percentiles of its distribution to provide a smooth yet flexible description of the dose-response relationship between mtDNA-CN and AF.

To adjust for potential confounders, we will use 3 multivariate models with progressive degrees of adjustment as described below:

Model 1: age, sex and race/enrollment center groups.
Model 2: model 1 + body mass index, height, smoking, alcohol intake, physical activity
Model 3: model 2 + total and HDL cholesterol, cholesterol medication, systolic and diastolic blood pressure, hypertensive medication, diabetes, prevalent CHD, and prevalent heart failure, and prevalent CKD at baseline.

Additionally, we will perform pre-specified subgroup analyses by race, sex, other comorbidities, and tested for potential interactions. We will also perform sensitivity analyses by excluding participants with prevalent CVD at baseline. Furthermore, we will adjust for log-transformed white blood cell count since it has been found to be associated with mtDNA-CN from peripheral blood, as well as adjust for established biomarkers of AF risk such as NT-proBNP (available at visit 2 and visit 4).

All statistical analyses will be performed using STATA version 14 (StataCorp LP, College Station, Texas).

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.csc.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)?

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* _af_inc14.dta_____)

___ 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.csc.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.csc.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: