1.a. Full Title:
Role of the Human Metabolome in Incident Heart Failure Etiology among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): Metabolomics of heart failure

2. Writing Group:
Writing group members:
Yan Zheng, Bing Yu, Danny Alexander, Teri Manolio, David Aguilar, Gerardo Heiss, Eric Boerwinkle, Jennifer Nettleton

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

First author: Yan Zheng, M.D., M.PH
Address: Division of Epidemiology, Human Genetics & Environmental Sciences
University of Texas Health Science Center
1200 Herman Pressler, suite E-405
Houston, TX  77030
Phone: (713) 500-9823
E-mail: Yan.Zheng@uth.tmc.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: Jennifer A. Nettleton, Ph.D., Assistant Professor
Division of Epidemiology, Human Genetics & Environmental Sciences
University of Texas Health Science Center
1200 Herman Pressler, suite E-641
Houston, TX  77030
Phone: (713) 500-9367   Fax: 713-500-9264
Email: Jennifer.A.Nettleton@uth.tmc.edu

3. Timeline:
Data analyses and preparation will begin upon approval and will take around 2 weeks; manuscript drafting will commence once suitable analytical models are finalized, and will also take around 2 weeks; another 2 weeks will be needed for manuscript review by the writing group.

4. Rationale:
Studies of the relationship between metabolomics and incident heart failure (HF) in African Americans are few. The human metabolome may improve our understanding of HF etiology. Currently, there are no longitudinal data to assess the relations between aspects of the human metabolome and HF. Therefore, we plan to explore the metabolomic antecedents of HF in a well-characterized population-based sample of African Americans in the Atherosclerosis Risk in Communities (ARIC) study.

5. **Main Hypothesis/Study Questions:**

We propose explore the role of individual metabolites and composites of related metabolites (reflecting specific pathways) in determining risk of HF. We will also evaluate the role of these metabolites or collection of metabolites (reflecting pathways) above and beyond that of baseline traditional HF risk factors in ~2,000 ARIC African Americans with metabolomic data. These metabolites were measured by GC-MS/LC-MS (n = 204, filtering on metabolites with high reliability coefficients and low prevalence of ‘missing’ values) via ARIC ancillary study 2008.16 “Metabolomics & Heart Failure: A Novel Approach to Biomarker Discovery.”

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 and Sample**

The present study sample will consist of African American ARIC participants at the baseline (visit 1) examination free of prevalent heart failure.

**Exclusions:**
- Non-African American race
- Prevalent heart failure at visit 1
- Missing heart failure outcome or follow-up time information
- Missing covariates information

**Outcome:**
- Incident hospitalized heart failure (by ICD-9 or ICD-10 in the first position)

**Follow-up time:**
- From visit 1 to event or to December 31, 2008.

**Primary models will adjust for basic risk factors, including the following:**
- Age
- Gender
- (field center is omitted since only Jackson African Americans have serum metabolomic data)

We will further evaluate how metabolites (or composite of metabolites) relate to HF after adjustment for several baseline HF risk factors in extended models—which
may be either confounders or mediators of the relations between the human metabolome and HF risk. These potential confounders or mediator will be introduced one at a time and then in combination in a forward step-wise manner.

- Body mass index (kg/m²)
- Current smoking status
- Physical activity
- Education level
- Systolic blood pressure
- Antihypertensive medication
- Prevalent diabetes status
- Serum glucose (by Clinical Chemistry Laboratory with calibration)
- Prevalent coronary heart disease status
- Left ventricular hypertrophy (by electrocardiograph using Cornell criteria)
- Serum total cholesterol
- HDL cholesterol (or triglycerides)
- LDL cholesterol
- Lipid-lowering medication use
- Serum creatinine

Metabolomics data:
Based on both practical and theoretical considerations, we have placed each measured metabolite into three groups by reliability coefficient (RC; from either the medium-term reliability study or the blind duplicates) and missing percentage.

- Group 1 contains metabolites (n=187) that are reliably measured (RC ≥ 0.60) and have missing values in fewer than 50% of the sample. The metabolites are to be treated as continuous variables during data analysis with the missingness of metabolites are replaced by the lowest measured value.

- Group 2 contains metabolites (n=17) that are reliably measured (RC ≥ 0.60) but have a moderate amount of missing data (values missing in 50-80% of the sample). For this group, we consider missing values as category 1. For the measured (non-missing) values, we consider values below the median as category 2 and values above the median as category 3. Ordinal regression is to be used during data analysis.

- Group 3 contains metabolites (n=398) that have >80% missing data or RC < 0.6; this group is not included in data analysis.

Statistical Methods
Multivariable Models and Estimation of Heart Failure Risk Function
For each metabolite, hazard ratios (HRs) with multivariate adjustment will be estimated using Cox proportional hazard models, described above - a primary model and an extended model adjusting for different potential combination of baseline risk factors respectively. These potential combinations of baseline risk factors will be researched by inspecting how beta-coefficient for the metabolite (or collection of metabolites) changes as we add individual risk factor in the extended model for each metabolite. The proportionality assumption will be evaluated by including a time×risk interaction term.
Statistical significance for the metabolomic data will be pre-specified with an experiment-wise $\alpha = 0.05$ (2 tailed). We will use modified Bonferroni procedures considering the correlations among metabolites.\textsuperscript{1,2} All statistical analyses are to be performed in SAS version 9.2 (SAS Institute, Cary, NC).

References:


7.a. Will the data be used for non-CVD analysis in this manuscript? \textbf{No}
7.b. \textbf{NA}
8.a. Will the DNA data be used in this manuscript? \textbf{No}
8.b. \textbf{NA}

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](http://www.cscc.unc.edu/ARIC/search.php)

\textbf{Yes. There is no overlap between this proposal and current proposals/published manuscripts.}

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? \textbf{Yes}
11.b. If yes, is the proposal
_____ A. primarily the result of an ancillary study (list number* \textbf{2008.16} )
_____ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* \textbf{2008.16 “Metabolomics & Heart Failure: A Novel Approach to Biomarker Discovery”})
*ancillary studies are listed by number at [http://www.cscc.unc.edu/aric/forms/](http://www.cscc.unc.edu/aric/forms/)

12. 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.
Yes, the lead author is aware that manuscript preparation is expected to be completed in 1-3 years, and if this expectation is not met, the manuscript proposal will expire.