1.a. **Full Title**: MyomiRs, cardiac structure and function, and incident heart failure risk

b. **Abbreviated Title (Length 26 characters)**: microRNA 208a, echo, and heart failure

2. **Writing Group**:
   Writing group members: Amil M Shah, Eric Boerwinkle Christie Ballantyne, Scott Solomon, Calum Macrae, Bing Yu; Others welcome.

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

**First author**: Amil M Shah, MD MPH  
**Address**: 75 Francis Street  
Boston, MA 02115

Phone: 617-525-6733  
Fax: 617-582-6027  
E-mail: ashah11@rics.bwh.harvard.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**: Amil M Shah, MD MPH  
**Address**: same as above

Phone:  
Fax:  
E-mail:

3. **Timeline**:
   Analysis will begin once this manuscript proposal is approved. Anticipate manuscript completion in approximately the following 3 months.

4. **Rationale**:

   Heart failure (HF) is a major public health problem and predominantly affects the elderly, with over 80% of HF hospitalizations occurring in persons over 65 years of age.\(^1\) Heart failure with preserved ejection fraction (HFpEF), in particular, is increasing in prevalence, and causes substantial morbidity, mortality, and resource utilization.\(^2,3\) While the cardiac phenotype in HFpEF is diverse,\(^4\) it has classically been characterized by LV hypertrophy and diastolic dysfunction.\(^5,6\) Abnormalities of LV systolic function are also
being increasingly appreciated. MicroRNAs (miRs) are small non-coding RNAs typically of 19- to 25-nucleotides, which have been implicated in the regulation of multiple physiologic and pathophysiologic processes including HF.

miR208a is one of the only cardiac myocyte specific miRs and has been extensively studied in rodent models of HF. miR208a is encoded in the intronic sequence of the α-myosin heavy chain (α-MHC) gene (MYH6) and is involved in regulating the cardiac response to hemodynamic stress and hyperthyroidism. Knockout mice for 208a do not demonstrate overt abnormalities of cardiac structure, but demonstrate higher levels of natriuretic peptide transcripts and develop evidence of mild systolic dysfunction in an age-dependent fashion. Intriguingly, when exposed to thoracic aortic banding, these knockout mice demonstrate an absence of β-MHC upregulation and a blunted hypertrophic and fibrotic response. Conversely, overexpression of miR208a results in increased ventricular size and wall thickness, myocyte hypertrophy, and upregulation of β-MHC expression — particularly in areas of fibrosis. miR208a overexpression was also associated with cardiac conduction system abnormalities, including first and second degree AV block. miR 208a appears to regulate the cardiac response to stress via two related miRs: mir208b, which is encoded in an intronic sequence of the β-MHC (MYH7) gene, and miR499, which is encoded in an intronic sequence of the MYH7B gene. Collectively, this family of myocyte-specific miRs has been referred to as MyomiRs. Both miR208b and miR499 are expressed in both cardiac and skeletal myocytes, where they promote β-MHC and slow-twitch (type 1) myofibril expression respectively. Additional regulatory targets of this family of miRs include Thrap1, Sox6, Purβ, and Sp3. In a rat model of hypertension-induced HF, the administration of an antagonir to neutralize miR208a resulted in improved LV diastolic function, decreased myocyte hypertrophy, and decreased periarteriolar fibrosis.

While the role of miR208a has been well described in rodent models of hypertension and HF, little data on miR208a in human HF exist. While MyomiRs are thought to act largely intracellularly or in a paracrine fashion, MyomiRs can be detected in circulating plasma. We have recently found that circulating levels of miR208a are higher among patients with HFpEF compared to controls using a small, highly phenotyped cohort based at the Brigham and Women’s hospital who underwent cardiopulmonary exercise testing with concomitant invasive hemodynamic monitoring (unpublished). 19 HFpEF patients (VO2max <80% predicted; exercise pulmonary capillary wedge pressure >20 mmHg) demonstrated higher resting mixed venous miR208a levels than 10 controls with normal functional capacity and hemodynamic response to exercise. Resting miR208a level correlated significantly with percent predicted VO2max and cardiac output. These findings are concordant with data from animal models, implicating miR208a in cardiac adaptations to hemodynamic stress that are believed to underlie HFpEF. However, whether circulation levels of miR208a relate to LV structure, diastolic function, and systolic function in humans in unknown. In addition, whether genetic variants resulting in alteration in 208a levels influence risk of HF is unknown.

In this study, we propose to (1) define the association between circulating WBC levels of miR208a, miR208b, and miR499 and cardiac structure and function, biomarkers of myocardial stress and injury, and measures of cardiac conduction disease (PR interval, AV block, prevalent atrial fibrillation); (2) use whole genome sequence data to identify genomic mutations in the coding sequences for these miR that relate to
miR levels; and (3) determine the prognostic relevance of these ‘loss of function’ or ‘gain of function’ mutations for incident heart failure and incident atrial fibrillation.

5. Main Hypothesis/Study Questions:

We hypothesize that (1) higher levels of miR208a will be associated with greater LV mass, worse LV diastolic function, and worse LV systolic function, and (2) mutations associated with lower/loss of miR208a expression will be associated with a lower risk of incident HF and of incident 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:
The study design to address hypothesis 1 will be cross-sectional at Visit 5, using echocardiographic data and miR levels obtained through Dr Boerwinkle’s ancillary study ‘RNA signatures of cardiovascular and metabolic disease risk factors’. Hypothesis 2 will be addressed using a prospective cohort study design. Mutations in the MYH6, MYH7, and MYH7B gene associated with alterations in the levels of miR208a, 208b, and 499 levels will be associated with the risk of incident cardiovascular events (HF or atrial fibrillation).

Inclusion/exclusion criteria:
For analysis 1, participants without data for miR levels or with missing data for key echocardiographic criteria (E wave, A wave, E wave deceleration time, TDI E’, and LAVi) or biomarkers (hsTnT or NT-proBNP) will be excluded.

Key variables of interest:
1. Genetic/RNA sequence data: levels of miR208a, 208b, and 499 at Visit 5; whole genome sequence variants for MYH6 intronic region coding miR208a, MYH7 intronic region coding miR208b, MYH7b intronic region coding miR499, and for the 3’ UTR of Thrap1, Sox6, Purβ, and Sp3
2. Echocardiographic variables (visit 5 echo) of LV diastolic function (E wave, A wave, E wave deceleration time, TDI E’, and LAVi), LV structure (wall thickness, relative wall thickness, systolic and diastolic diameters and volumes), RV function (FAC, TA S’), and pulmonary artery systolic pressure
3. Electrocardiographic variables (visit 5): heart rate, rhythm, PR interval, QRS interval, QT interval, LVH criteria
4. Laboratory values (visit 5): high sensitivity troponin T, NT-proBNP, serum albumin and creatinine, urine albumin and creatinine, hemoglobin and hematocrit, glucose, hemoglobin A1C, total cholesterol, triglycerides, HDL, LDL
5. Clinical covariates (visit 5): age, gender, race/ethnicity, height, weight, blood pressure, heart rate, history of hypertension, diabetes, dyslipidemia, coronary artery disease, prior MI or revascularization procedure, prior stroke or TIA,
Peripheral arterial disease, heart failure, prior hospitalization for heart failure, atrial fibrillation

Data analysis:
For hypothesis 1, miR208a levels will be modeled as both an ordered categorical and a continuous variable. For the categorical analysis, the study population will be divided into groups based on miR208a level: participants with undetectable level will be placed in one group and the other 4 groups will be generated by splitting the observed miR208a levels into approximate fourths. Clinical covariates, laboratory variables, echocardiographic parameters of structure and function, and echocardiographic parameters of diastolic function (E’, E wave/E’, LAVi, E/A ratio, E wave deceleration time) will be described by hsTnT category. For the continuous analysis, miR208a levels will be used directly or will be log transformed depending on the normality of its distribution, with undetectable values assigned a value just below the lower detection limit of the assay. Correlation of miR208a level with echocardiographic parameters of LV structure, systolic function, and diastolic function will be assessed by univariate linear regression and by multivariable linear regression after adjusting for age, gender, race, blood pressure at time of echo, history of hypertension, diabetes, coronary artery disease, and eGFR. Additional analysis will be performed, stratifying by the presence of thyroid disease defined by the use of thyroid medications around the time of Visit 5. Similar analyses will be performed for miR208b and miR499.

For hypothesis 2, genomic variants in the intronic sequences of MYH6, MYH7, and MYH7B coding for miR208a, 208b, and 499 pri-microRNA respectively will be identified using whole genome sequencing data. Variants associated with significantly high or low levels of the miRs of interest will be identified. These variants will then be associated with (1) incident HF using Cox proportional hazards modeling, (2) incident atrial fibrillation, (3) cardiac structure and function at Visit 5, and (5) NT-proBNP and hsTnT levels at Visits 4 and 5. Unadjusted associations will be determined. Subsequently, as we expect these variants to influence the cardiac response to hemodynamic stress, multivariable models will be employed, adjusting for hypertension, diabetes, coronary disease, eGFR, and thyroid disease.

Anticipated methodologic limitations:
Circulating miR levels in ARIC are measured from peripheral leukocytes, while existing data on miR208a derive from myocyte expression (animal studies) or free circulating plasma (humans). Analysis 1 is cross-sectional and therefore causality cannot be inferred. The regulatory actions of the miR208a family of microRNAs have been described in animal models, where αMHC is the predominant isoform expressed in adults, while in human adults, βMHC is the predominant isoform expressed. Therefore, the described regulatory biology may vary in humans from what has been described in animal models. However, the conservation of these miRs across species makes this less likely and our preliminary findings are concordant with expectations based on the animal model data.
7.a. Will the data be used for non-CVD analysis in this manuscript?  ____ Yes ___x__ 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?  ____ Yes ___x__ No

(This file ICTDER03 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.cscce.unc.edu/ARIC/search.php](http://www.cscce.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)?

   MS#2188- (Graff et al) Identification of miRNA-Mediated Functional Genetic Variants for Risk Factors of Cardiovascular Disease: Functional Genome-Wide Association and Interaction Studies in the CHARGE Consortium

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* __2014.3__)  ____

   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.cscce.unc.edu/aric/forms/](http://www.cscce.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.
References

10. van Rooij E, Quiat D, Johnson BA, Sutherland LB, Qi X, Richardson JA, Kelm RJ, Olson EN. A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. Dev Cell 2009;17:662-73.