1.a. Full Title: Adiposity-related DNA methylation variants as predictors of Type 2 Diabetes and Coronary Heart Disease in African Americans: The ARIC Study.

b. Abbreviated Title (Length 26 characters):

2. Writing Group:
   Writing group members:
   Ellen W. Demerath
   Steven Nguyen
   Weihua Guan
   Jim Pankow
   Kari North
   Jan Bressler
   Megan Grove
   Myriam Fornage
   Eric Boerwinkle
   Others Welcome

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

First author: Steven Nguyen
   Address: 1405 Jessamine Avenue West, Saint Paul, MN 55108
   Phone: 408-807-9552   Fax: 408-807-9552
   E-mail: nguy2295@umn.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: Ellen W. Demerath, PhD
   Address: University of Minnesota School of Public Health
   Division of Epidemiology and Community Health
   1300 S. Second Street, Suite 300
   Minneapolis, MN 55454
   Phone: 612-624-8231   Fax: 612-624-8231
   E-mail: ewd@umn.edu

3. Timeline:
August, 2015: Create dataset and conduct analyses
September, 2015: Complete analyses and draft abstract for ARIC P and P approval
October, 2015: Submit results to 2016 AHA EPINPAM meeting
November-January, 2015: Write manuscript and submit to ARIC P and P for approval
February, 2016: Submit manuscript for review to journal

4. Rationale:

Epigenetics is the study of heritable DNA modifications not involving changes in the underlying DNA sequence which can affect transcription and phenotype. One form of epigenetic modification is DNA methylation, involving the addition or removal of methyl groups in CpG dinucleotides which can affect gene expression particularly when they occur in promoters (Jones, 2012). Previous research has identified environmental factors observed to be associated with epigenetic modification such as smoking (Shenker et al, 2012) and BMI (Demerath et al, 2014). Epigenetic modification may be relevant to chronic disease pathophysiology where changes in the epigenome affecting transcription and phenotype may lead to dysregulation of biochemical processes such as the role of chronic inflammation in coronary heart disease and insulin resistance in type 2 diabetes (Paneni et al, 2014).

A recent EWAS carried out using data from the ARIC study identified multiple BMI-related CpG sites that replicated in the Framingham Heart Study and GOLDN study with some replication in the muTHER cohort (Demerath et al, 2014). Obesity has been observed to be a precursor to type 2 diabetes and coronary heart disease (NHLBI, 2013) and while not all molecular mechanisms involved in the pathogenesis of these obesity-related diseases are known as yet, research has suggested overlapping pathways and processes such as chronic inflammation likely underlie this epidemiologic association (Esser et al, 2014). Our EWAS of adiposity traits in ARIC identified 38 replicated CpG sites associated with BMI or WC, many of which are involved in immunological activation and lipid metabolism pathways. Therefore, we propose to use these already identified CpG sites as a more targeted and statistically efficient means of detecting methylation marks predictive of cardiovascular disease than would be EWAS of those disease outcomes.

Our approach follows the current conceptual thinking (see Figure below) on how upstream environmental or behavioral factors like excess body weight change the epigenome over the lifecourse and those changes (including DNA methylation alterations) then cause disease. We cannot infer causation in this analysis, but if there are significant findings, we will examine them more deeply using Mendelian Randomization approaches to assess whether obesity-related methylation variation indeed mediates the relationship of obesity to diabetes and coronary heart disease.

Adiposity $\rightarrow$ Methylation $\rightarrow$ Disease

5. Main Hypothesis/Study Questions

Previously assessed Proposed
It is hypothesized that there are associations between BMI-related DNA methylation variation and the risk of 1) prevalent type 2 diabetes mellitus (T2DM), 2) incident T2DM, and 3) incident coronary heart disease (CHD) in African Americans, presumably due to the effect of adiposity and its downstream changes (dyslipidemia; blood pressure) on methylation and gene expression in genomic regions involved in these related diseases. It is hypothesized that a targeted association study such as this will yield additional hits beyond those identified via EWAS because of the much lower stringency of the multiple comparisons adjustment for statistical significance, and because, theoretically, the obesity-related sites chosen are more likely to be involved in chronic disease.

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 and Prospective Cohort

Inclusion/Exclusion: African-American subjects in ARIC with DNA methylation data at visit 2/3, and concurrent phenotype information (including sex, age, BMI, smoking status, physical activity level).

Outcome/variables of interest: Prevalent type 2 diabetes, incident type 2 diabetes, prevalent coronary heart disease, incident coronary heart disease

Primary exposure/ independent variable: CpG methylation level (beta values) from Illumina 450K Human Methylation BeadChip from visit 2 for 41 CpG sites. The selection of CpG sites to examine in this study are informed by a recently published study which identified multiple BMI-related loci that replicated in the Framingham Heart Study, GOLDN study, and with partial replication in the muTHER cohort. 37 of these CpG sites were observed to be BMI-related, one site was associated with waist-circumference, and three other sites (near HIF3A) were observed to be BMI-related in ARIC only and did not replicate however have been reported in literature to be BMI related (Dick et al, 2014)

<table>
<thead>
<tr>
<th>CpG Site</th>
<th>CHR</th>
<th>Nearest Gene</th>
<th>Reason for inclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg04869770</td>
<td>1</td>
<td>PBX1</td>
<td>Replicated association with BMI</td>
</tr>
<tr>
<td>cg17901584</td>
<td>1</td>
<td>DHCR24</td>
<td>Replicated association with WC</td>
</tr>
<tr>
<td>cg23998749</td>
<td>1</td>
<td>Intergenic</td>
<td>Replicated association with BMI</td>
</tr>
<tr>
<td>cg20954977</td>
<td>2</td>
<td>B3GNT7</td>
<td>Replicated association with BMI</td>
</tr>
<tr>
<td>cg06876354</td>
<td>2</td>
<td>RALB</td>
<td>Replicated association with BMI</td>
</tr>
<tr>
<td>cg14017402</td>
<td>2</td>
<td>Intergenic</td>
<td>Replicated association with BMI</td>
</tr>
<tr>
<td>cg12992827</td>
<td>3</td>
<td>NA</td>
<td>Replicated association with BMI</td>
</tr>
<tr>
<td>cg18307303</td>
<td>5</td>
<td>IL12B</td>
<td>Replicated association with BMI</td>
</tr>
<tr>
<td>cg26403843</td>
<td>5</td>
<td>RNF145</td>
<td>Replicated association with BMI</td>
</tr>
</tbody>
</table>
Secondary exposures/independent variables: Participant age, gender, BMI, smoking history, alcohol use, education, physical activity, income, white blood cell count and imputed cell count differentials (lymphocytes, monocytes, neutrophils, eosinophils), and follow up time to report of disease.

Data Analysis:

Statistical analyses will be conducted using SAS 9.4. The dataset includes methylation data on approximately 2,800 African American adults from Visit 2. Mean methylation and covariate information will be statistically described in participants by disease status after participants with missing information relevant to this study are excluded. To individually test the association between BMI-related DNA methylation for
the 41 CpG sites, separate crude, adjusted, and full models will be fitted for each CpG site.

First, logistic regression will be used to examine the relationship between BMI-related DNA methylation and prevalent diabetes by calculating odds ratios and 95% confidence intervals to describe the odds of prevalent disease based on methylation beta value for a particular CpG site. Crude models will only contain CpG information while adjusted models contain all other covariates besides BMI and full models additionally contain BMI.

Next, Cox proportional hazards regression is used to examine the association between BMI-related DNA methylation and incident diabetes by calculating hazard ratios and 95% confidence intervals. The proportional hazards assumption will be tested by specifying a natural log-time to follow up to report of disease or last report of non-disease status for participants who did not develop diabetes during follow up and DNA methylation interaction term in models. Where the proportional hazards assumption does not hold at the p<0.05 level, follow up time to event or last report of non-disease will be coded as time periods in between visits (e.g., visit 2 to 3, visit 3 to 4, etc), and re-run to calculate hazard ratios and 95% confidence intervals for each period of time. Crude models contain only CpG information while adjusted models contain other covariates except for BMI with full models additionally containing BMI.

Cox proportional hazards regression will similarly be used to examine the association between BMI-related DNA methylation and incident coronary heart disease. However, where violations of the proportional hazards assumptions are observed, follow up time will be recoded into tertiles of time to event for those with events, and hazard ratios and 95% confidence intervals for each third of follow up time will be calculated. Crude models will contain only CpG information while adjusted models contain potential confounders, and a fully adjusted model additionally includes BMI itself and its downstream disease factors related to the outcomes as a way of examining potential mediation (BMI, hypertension, TC, and HDL).

To account for multiple comparisons regarding 41 CpG sites, the threshold for significance will be calculated as 0.05/41 or 0.00122 when determining significance for the association between a particular CpG site and each disease outcome.

References:


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 ____ 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.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)?
11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ___ Yes ___x___ No

11.b. If yes, is the proposal
   ___ A. primarily the result of an ancillary study (list number* __________)
   ___ 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.