1.a. **Full Title:** Genome-wide scan for meQTLxE Interactions on Central Adiposity: The ARIC Study

b. **Abbreviated Title (Length 26 characters):** meQTLxE of Central Adiposity

2. **Writing Group:**
   Writing group members: Anne Justice, Ellen Demerath, Jan Bressler, Myriam Fornage, Megan Grove, Kari North, Weihua Guan, Eric Boerwinkle and other investigators welcome

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. AEJ

**First Author: Anne Justice**
Address: 137 East Franklin Street, Suite 306  
Department of Epidemiology  
University of North Carolina at Chapel Hill  
Chapel Hill, NC 27514  
Phone: 919.966.1403  
Fax: 919.966.9800  
E-mail: anne.justice@unc.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: Kari North  
Address: 137 East Franklin Street, Suite 306  
Department of Epidemiology  
University of North Carolina at Chapel Hill  
Chapel Hill, NC 27514  
Phone: 919.966.2148  
Fax: 919.966.9800  
E-mail: kari_north@unc.edu

3. **Timeline:** 1 year
   - Study-level statistical analyses: July-September 2016
   - Interpretation and meta-analyses: October 2016-January 2017
   - Manuscript preparation: January-March 2017
   - Manuscript submission: March 2017

4. **Rationale:** Elevated central adiposity is a recognized risk factor for cardiometabolic disease (CMD)\(^4\)-\(^6\); however, rates of obesity, and particularly central obesity, have more
than doubled in the U.S. over the past three decades \cite{7,9}. Further, there are stark differences in obesity risk among minorities \cite{7,10}, and across sexes in body fat distribution and its genetic predisposition \cite{11}. While the genetic influence of central adiposity is well-established \cite{12}, there are multiple critical questions that remain unanswered, which, if answered, could lead to important discoveries about potentially preventable contributors to obesity. One such priority is the importance of epigenetic factors in the pathogenesis of obesity. DNA methylation is an important epigenetic mechanism that links genotypes, the environment, and obesity, but methylation studies have primarily been conducted in small studies of only European descent (EUD) subjects, with limited consideration of the possible role of clinically relevant environmental influences. Also, there is a dearth of epigenetic studies of central adiposity \cite{13-15}. The identification of epigenetic factors that influence the pathogenesis of central obesity may allow for the identification of biomarkers for risk and/or progression, and thus new public health interventions \cite{16}.

An abundance of evidence exists for sex-specific methylation patterns and changes in global methylation \cite{13,17,18}. Several studies have reported sex-specific patterns of methylation in genes related to a wide variety of biological processes \cite{19}. These sex-specific methylation patterns help to explain differences within the organismal environment that exist between men and women exclusive of sex determination. While there is clear evidence of sex-specific global and regional methylation patterns, these patterns have yet to be explicitly examined within the context of adiposity. Given the sex-specific genetic architecture of central adiposity traits, such studies are of great interest.

There is a global decrease in methylation in current smokers as compared to former and never smokers \cite{18,20}. Smoking-specific methylation patterns have been identified using both EWAS and candidate gene approaches across 15 loci, including changes in methylation patterns near CHRN genes \cite{18,21}, a region known to harbor SMK-specific SNP associations with adiposity related traits \cite{22}, supporting the hypothesis that genetic factors interact with smoking to influence adiposity. Further, several studies also indicate that site specific hypomethylation was reversible after smoking cessation \cite{21}. We will explicitly explore the effects of smoking on epigenetic influence to central obesity.

Just as epigenetics promises to identify new pharmacological targets, many existing drugs alter gene expression through epigenetic modifications, including antidepressants, antipsychotics, and antihypertensives \cite{23-25}, and which are also known to influence weight change \cite{3,26,27}. These epigenetic modifications may improve drug response, attenuate drug response, or cause other unexpected downstream effects, such as weight change. For example, valporic acid, lithium, and clozapine used to treat bipolar disorder and schizophrenia, all lead to methylation-induced changes in gene expression \cite{23,25,28}. Interestingly, weight gain is a common side effect of all of these drugs. Additionally, olanzapine has been shown to influence the expression of almost 20 different genes in the brain and liver, including several genes harboring GWAS-implicated SNPs for BMI variation (e.g. $PCDH9$, $CDH13$) \cite{29,30}. So, while the area of pharmaco-epigenetics remains largely unexplored, these studies illustrate the importance of further epigenetic by drug interaction studies. This will be the first study to specifically explore epigenetic by drug interaction on changes in adiposity.

5. Main Hypothesis/Study Questions:
Aims: Conduct an epigenome wide association analysis (EWAS) of central adiposity to identify meQTLs associated with central adiposity while accounting for the influence of environmental exposures (sex, smoking behavior, and obesogenic and anorexigenic medication use) using extant phenotypic and Illumina HumanMethylation 450K Beadchip (HM450K) data in 2,861 African American (AA) and 939 EUD participants in the Atherosclerosis Risk in Communities (ARIC) study. A) Identify methylation quantitative trait loci (meQTL) associated with central adiposity (waist circumference [WC], WC to hip ratio [WHR], and WC to height ratio [WHtR]) while accounting for environmental exposures. B) Replicate findings in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium.

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 analysis using data taken from visit 2 or 3 (dependent on concordantly measured methylation data) in the ARIC population-based cohort.

Inclusion:
- Adults ≥ 18 years of age
- European and African American Ancestry

Exclusion:
- Individuals < 18 years of age
- Pregnant women
- Missing epigenetic data
- Missing or unrealistic values (+/- 4SD from the mean) for outcome (waist circumference, hip circumference, waist-to-hip ratio, height) data
- Missing covariate (BMI, age, PCs, study center, smoking status, and sex) data

Outcomes: WC, WHR ratio, and WHtR ratio, all adjusted and unadjusted for BMI.

Exposures: Sex, Smoking, Obesogenic medication use

Genotype data:
- HM450 array

Summary data analysis:

Discovery Analyses. To determine if site specific β values of methylation probes are associated with central adiposity after adjusting for environmental exposures, measured as WC and WHR, and WtHR, we will employ linear mixed models (LMM) in R with methylation β values as the independent variable, central adiposity as the dependent variable, and with chip array specified as a random effect. In addition to smoking, sex, and obesogenic medication use, the following variables will be tested for inclusion as potential fixed effects: 10 principal components scores (PCs) from the HM450 array to account for potential confounding by

<table>
<thead>
<tr>
<th>Indication</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antidepressants</td>
<td>TCA</td>
</tr>
<tr>
<td>Antidepressants</td>
<td>SSRI</td>
</tr>
<tr>
<td>Allergies</td>
<td>Histamine-1 receptor antagonist</td>
</tr>
<tr>
<td>Mood Disorder</td>
<td>Atypical antipsychotics</td>
</tr>
<tr>
<td>Allergies</td>
<td>Antihistamine</td>
</tr>
<tr>
<td>Antihypertensives</td>
<td>Oral beta blocker</td>
</tr>
<tr>
<td>Antiinflammatory/</td>
<td>Corticosteroids</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Thiazolidinediones</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Sulfonyureas</td>
</tr>
<tr>
<td>Seizures/Mood</td>
<td>Mood stabilizer</td>
</tr>
<tr>
<td>Seizures/Migraines</td>
<td>Anagase</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Ammoniumcitrate</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Biguanides</td>
</tr>
<tr>
<td>Diabetes Medications</td>
<td>Incretin mimetics</td>
</tr>
<tr>
<td>Diabetes Medications</td>
<td>DPP-4 inhibitor</td>
</tr>
<tr>
<td>Epilepsy/Seizure</td>
<td>Anticonvulsants</td>
</tr>
<tr>
<td>Parkinsons/Restless Leg</td>
<td>Dopamine precursor</td>
</tr>
<tr>
<td>ADHD</td>
<td>CNS stimulant</td>
</tr>
<tr>
<td>ADHD</td>
<td>Methylxanthine</td>
</tr>
</tbody>
</table>

Table 3. Known classes of obesogenic and anorexigenic medications.
genetic ancestry, study center, WBC count, age, study center, education, household income, current alcohol consumption, and physical activity. The final choice of covariates will be based on the Bayesian model averaging (BMA) algorithm for linear regression models to choose the best fit model for central adiposity without meQTL45. BMA will be implemented using the R package BMS v0.3.0. To identify site-specific methylation β value-by-environment interactions (meQTLxE) with central adiposity measures, we will employ a LMM in R, but with the addition of an environment by methylation interaction term (smoking - never, ever, or current smoker; sex: men, women; pharmaceutical obesogen/anorexigen: yes/no). We will employ a 1 degree of freedom (1df) test for meQTLxE, and as a secondary approach, we will perform a 2df test considering the covariance of the methylation variant with the exposure for the interaction meta-analysis, thus testing the joint p-value of main effect of the meQTL and the meQTLxE effects. The 2df test is used in addition to the 1df test due to its robustness in detecting effects (e.g. when they are directionally consistent but weaker in one stratum). So, for the joint 2df test, much smaller sample sizes are needed to detect modest effects. Additionally, all analyses will be conducted stratified and meta-analyzed across self-identified race/ethnicity.

**Exposures.** Questionnaires were used to assess current smoking status (coded as current, former, and never smoker). Participants brought prescription medications to each visit, and drug names were recorded by ARIC study staff in a standardized medication survey. This medication survey will be used to identify drugs known to cause changes in weight (see Table 1 for list of medication classes) [https://www2.cscc.unc.edu/aric/cohort-manuals](https://www2.cscc.unc.edu/aric/cohort-manuals). As studies in the CHARGE consortium were collected at variable times and age groups, and medication use changes though time, this comprehensive list, which may include drugs not present in the ARIC study, will be used to classify drugs in replication cohorts.

**Replication Analyses.** Like as in Aim 1 replication, meQTL sites with smoking or sex interaction association p values <1.03x10^{-7} (CWS corrected for number of CpG variants tested) for WC, WHR, or WtHR will be carried forward for replication in participating CHARGE cohorts with phenotype, covariate and methylation data available (Table 2). Every attempt to harmonize the smoking variable (current, former, never smoker) will be made. In cases in which former smoker cannot be determined, current smokers and never smokers will be coded in line with coding in the ARIC study. Again, the meta-analysis will be conducted on p values using a sample size-weighted method implemented in METAL, which also has an option for the 2df model. Replication will be defined as consistent direction of the effect with discovery, and a meta-analysis p value

<table>
<thead>
<tr>
<th>Cohort Name</th>
<th>Ethnicity(ies)</th>
<th>Smoking Available Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGCUT</td>
<td>European Descent</td>
<td>Y</td>
</tr>
<tr>
<td>FHS/FHSgen2</td>
<td>European Descent</td>
<td>Y</td>
</tr>
<tr>
<td>GOLDN</td>
<td>European Descent</td>
<td>Y</td>
</tr>
<tr>
<td>InChianti</td>
<td>European Descent</td>
<td>Y</td>
</tr>
<tr>
<td>KORA</td>
<td>European Descent</td>
<td>Y</td>
</tr>
<tr>
<td>RS</td>
<td>European Descent</td>
<td>Y</td>
</tr>
<tr>
<td>TwinsUK</td>
<td>European Descent</td>
<td>Y</td>
</tr>
<tr>
<td>WHI</td>
<td>African- and European- American</td>
<td>Y</td>
</tr>
</tbody>
</table>
As in the discovery analysis, all analyses will be conducted stratified and meta-analyzed across self-identified race/ethnicity.

**Limitations/challenges:** The methods discussed herein may be supplemented or altered as newly established methods develop. I discuss my preferred analytical method with an understanding of the limitations and propose an alternative should such limitations arise.

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)?

-“Genome-wide methylation analyses of cardiovascular disease (CVD) and its risk factors” #1928
-“Epigenome-wide association study of obesity traits in African American adults: The Atherosclerosis Risk in Communities (ARIC) Study” #2106

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.

13. Per Data Use Agreement Addendum for the Use of Linked ARIC CMS Data, approved manuscripts using linked ARIC CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes __X__ No.

Bibliography


