

## ARIC Manuscript Proposal # 3211

PC Reviewed: 8/14/18  
SC Reviewed: \_\_\_\_\_

Status: \_\_\_\_\_  
Status: \_\_\_\_\_

Priority: 2  
Priority: \_\_\_\_\_

**1.a. Full Title:** Socioeconomic Adversity, Epigenetics and Measures of Obesity in the ARIC Study

**b. Abbreviated Title (Length 26 characters):** Socioepigenomics of Obesity

### 2. Writing Group:

Writing group members: Lindsay Fernandez-Rhodes, Anne Justice, Ellen Demerath, Allison E. Aiello, Chantel Martin, LaShaunta Glover, Laura Loehr, Kari E. North and other investigators welcome

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

First Author: Lindsay Fernandez-Rhodes  
Address: 123 West Franklin Street  
Suite 210  
Carolina Population Center  
University of North Carolina at Chapel Hill  
Chapel Hill, NC 27514  
Phone: 919.260.4855 Fax: 919.966.9800  
E-mail: fernandez-rhodes@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: 123 West Franklin Street  
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-August 2018
- Interpretation and meta-analyses: September-October 2018
- Manuscript preparation: November 2018-April 2019
- Manuscript submission: May 2019

#### 4. Rationale:

Epigenetics provides a dynamic mechanistic framework for exploring the interaction between environmental factors, such as socioeconomic adversity, with the genome to ultimately alter disease risk. Epigenetics is also providing new targets for prevention and treatment. Yet, most epigenetic studies of human diseases have been conducted in European ancestry populations, cross-sectional in design, and have been conducted with little replication and insufficient consideration of how socioeconomic or cultural factors may either mediate the underlying disease pathogenesis, or drive diseases which then shape the epigenetic profiles of individuals living with that disease.

Socioeconomic adversity is an established social determinant of health. Yet, it is unclear how socioeconomic adversity becomes translated into greater obesity burden, poor health and may contribute to the current racial disparities in the United States (U.S.). In the U.S., individuals of African American (AA) ancestry have elevated burden of obesity, hypertension, insulin resistance, impaired glucose metabolism, and other ensuing cardiometabolic diseases, compared to other U.S. populations<sup>1-3</sup>. Yet, epigenetics has only recently begun to include more diverse populations, such as AAs, and consider their most pressing health concerns.<sup>4,5</sup>

The epigenetic event of DNA methylation at cytosines proximal to guanine nucleotides (CpG) may represent a mechanism through which stressful exogenous life experiences such as socioeconomic adversity may become embodied. For example, a growing body of literature (see systematic review<sup>6</sup>) supports the role of socioeconomic adversity in early life in patterning adulthood methylation profiles. However, the extent to which this association is mediated by prior obesity is unclear (investigated in Aim A below). It remains an open question if any longitudinal association between socioeconomic adversity and methylation profiles is independent of changes in obesity status (Aim B below). Lastly, methylation profiles at specific loci related to socioeconomic adversity may have a complex pattern of association with obesity across the adult life course,<sup>7</sup> which warrants further study (Aim C below).

Thus, the identification of epigenetic factors that associate with socioeconomic disparities for obesity in AAs, in addition to those already described in small studies of European descent,<sup>8</sup> may allow for additional insights into the biologic pathways influenced by socioeconomic adversity and thus new public health interventions<sup>9</sup> aimed at minimizing racial/ethnic health disparities. Moreover, the discovery of epigenetic variation influencing obesity-related traits has the potential to also identify important pathways for disease prediction and treatment.<sup>10</sup>

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

**Overall Aim:** To conduct an epigenome wide association analysis (EWAS) of socioeconomic adversity with consideration to overall and central obesity status/measures and directly comparing effects across race/ethnicity, using extant phenotypic and Illumina HumanMethylation 450K Beadchip (HM450K) data in up to 2,900 AA and 1,100 European descent (EUD) participants in the Atherosclerosis Risk in Communities (ARIC) study measured at visit 2 or 3. Additionally, we hope to incorporate the visit 5 measurement of methylation in the ARIC cohort into all of our Aims, but specifically using this additional data in our longitudinal Aim B below.

- A) To identify methylation loci associated with adulthood socioeconomic adversity as measured using highest education of the ARIC AA and EUD participants, and then determine if these socioepigenetic associations are independent of earlier measures of overall and central obesity by additionally adjusting for BMI and/or WHR.
- B) To examine the extent to which changes in household income associate with changes in methylation profiles, independent of changes in obesity status, or its related quantitative measures: BMI and WHR.
- C) Among loci identified in Aims A-B or previously associated with socioeconomic adversity in other populations
  - a. To conduct a structural equation model to examine the extent to which the effect of socio-economic adversity measured in adulthood on methylation profiles is 1) mediated by changes in obesity status/measures or 2) predictive of future obesity risk.
  - b. To replicate the cross-sectional or longitudinal (if available) findings in the studies of the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (CHARGE, see Table below), or similar consortium with data on socioeconomic adversity and methylation.

**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 longitudinal analysis using methylation data taken from visit 2 or 3 and visit 5 in the ARIC population-based cohort.

**Inclusion:**

- Adults  $\geq$  45 years of age
- AA and EUD participants

**Exclusion:**

- Individuals < 45 years of age
- Pregnant women
- Missing epigenetic data or socioeconomic adversity data
- Missing or unrealistic values ( $\pm$  4SD from the mean) for obesity-related measures (body mass index, waist-to-hip ratio)
- Missing covariates (age, PCs, study center, smoking status, sex, and BMI and/or WHR, childhood SEA, parental highest education from SES13, to be used in Aim C only)

**Exposures:** **Adulthood Socioeconomic Adversity:** Highest Educational Attainment (HOM54 at visit 1), Household Income (HOM62 at visit 1, SES6 at visit 4)

**Outcomes:** Methylation  $\beta$ -values

**Key Covariates:** Body mass index (BMI), waist hip ratio (WHR) in adulthood (as captured during one of the clinical examinations; BMI at 25 years of age (Aim A and C only)

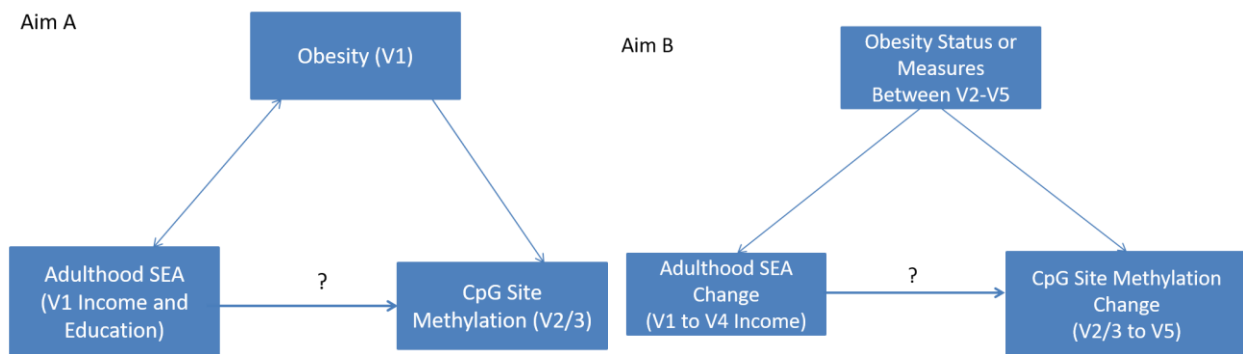
**Genotype data:**

- HM450 or EPIC array (as available)

**Summary data analysis:**

**Discovery Analyses.** To determine if socioeconomic adversity in adulthood associates with site specific  $\beta$ -values of CpG methylation probes and if this is mediated by obesity (tested in Aims A-B; see Figure 1A-B below) or another driver of obesity risk (tested in Aim C; see Figure 2 below).

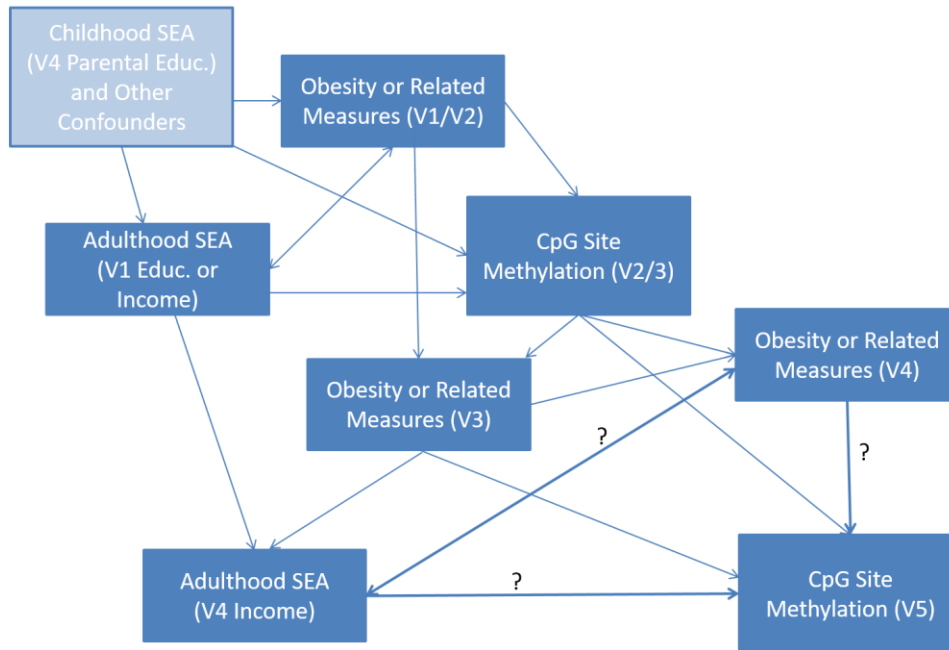
We will analyze both obesity using established thresholds (e.g.  $\geq 30$  kg/m<sup>2</sup>) as well as using related quantitative measures taken during the clinical examinations (i.e. BMI and WHR), to determine if the mediation by obesity is primarily acting through overall or central adiposity. Additionally, in Aims A and C, we may also explore replacing BMI from the clinical examinations with a measure of early adulthood overall obesity, as captured as BMI at 25 years of age (=self-reported weight in kg at 25 years/visit 1 height in m<sup>2</sup>).



**Figure 1. Directed Acyclic Graphs of the relationship between socioeconomic adversity, obesity, and CpG Methylation in the ARIC Study, as modeled in Aims A and B.**

In Aims A-B (see Figure 1A-B above), we will employ linear model in R with adulthood socioeconomic adversity (or their change) as the independent variable, and methylation  $\beta$  values (or their change) as the dependent variable. If missingness of any socioeconomic adversity variable(s) exceeds 10% we will consider multiple imputation of the socioeconomic adversity variable, as has been previously implemented by Dr. Whitsel within an EWAS context. In order to account for batch effects, we will use COMBAT to generate standardized methylation  $\beta$  values, accounting for 10 principal components scores (PCs) from the HM450 array and WBC count. Additionally, the following variables will be tested for inclusion as potential fixed effects: study center, sex, age, current smoking status, current alcohol consumption, current physical activity, 10 PCs of previously collected genetic data to account for potential confounding by genetic ancestry. Additionally, in Aim A we will adjust for BMI and/or WHI measured at visit 1 and in Aim B changes in obesity status or its related measures between visit 2-5. In Aims A-B, we may explore the impact of additional adjustments for childhood socioeconomic adversity (SES13 capturing their parent highest level of education), to estimate the independent impact of adulthood socioeconomic adversity on obesity. 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 detecting socioepigenomic loci. BMA will be implemented using the R package BMS v0.3.0.

Aim C



**Figure 2. Proposed structural equation model in Aim C to capture the direct and indirect (shown in wide lined arrows) pathways connecting socioeconomic adversity, obesity, and CpG Methylation in the longitudinal data of the ARIC Study.**

In the first phase of Aim C, we will use Mplus to generate a structural equation model (see Figure 2 above) to simultaneously assess the direct and indirect effects between adulthood socioeconomic adversity, methylation and obesity each measured at up to two time points, independent of the key confounders identified in Aims A-B (light blue box). The advantage of this particular approach is that it can account for missingness using full information maximum likelihood as well as investigate the direction of effect of any variables that are measured concurrently.

**Replication Analyses in the second phase Aim C.** Socioepigenomic association p values  $<1.03 \times 10^{-7}$  (chip-wide significance [CWS] corrected for number of CpG variants tested) for each socioeconomic adversity trait will be carried forward for replication in participating/interested CHARGE studies (see Table below for estimates of minimum sample sizes of possible collaborating CHARGE studies), or other consortium with similar socioeconomic measures and methylation data. Perhaps the largest and most diverse replication cohort for this study would be the Women’s Health Initiative (WHI). We estimate that we would be able to replicate our ARIC findings among at least 4,633 women (1,183 AAs, 2,744 EAs and 706 Hispanic/Latinas), as well as to replicate our longitudinal findings from Aim B in up to 243 women.

Given the different analytic strategies and cell types used among the replication studies, we will conduct a sample size-weighted meta-analysis implemented in METAL. Replication will be defined as consistent direction of the beta coefficient, and a CWS meta-analysis p value.

**Limitations/challenges:** The methods discussed herein may be supplemented or altered as newly established methods develop. I will discuss my preferred analytical method with respect to its specific limitations, and propose an alternative should such limitations arise.

**Table.** Replication cohorts available through  
The CHARGE Consortium to be used in Aim C.

<b>Cohort Name</b>	<b>Ethnicity(ies)</b>	<b>Minimum Estimated Sample Size</b>
Amish	European American	192
BLSA	African, European American	542
CHS	African, European American	385
FHS/FHSgen2	European American	2,846
GENOA	African American	239
GOLDN	European American	993
GS	European	227
HBCS	European	122
HyperGEN	African American	466
InChianti	European	472
KORA	European	1,488
LBC1921/1936	European	1,348
MESA	African, European, Hispanic/Latino American	1,183
NAS	European American	674
NTR	European	2753
RS	European	1,331
SYS adults	French Canadian	111
TwinsUK	European	724
WHI (BAA23, EMPC, and AS311 if willing to participate)	African, European, Hispanic/Latino American	4,633

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.csc.unc.edu/ARIC/search.php>

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
- "Epigenetic Factors Influencing Central Adiposity: The ARIC Study" #2802

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

11.b. If yes, is the proposal

A. primarily the result of an ancillary study (list number\* multiple studies related to the collection of the ARIC methylation data)

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](http://publicaccess.nih.gov/submit_process_journals.htm) shows you which journals automatically upload articles to PubMed central.

## References

1. Stevens GA, Singh GM, Lu Y, et al. National, regional, and global trends in adult overweight and obesity prevalences. *Population health metrics* 2012; **10**(1): 22.
2. Finucane MM, Stevens GA, Cowan MJ, et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet* 2011; **377**(9765): 557-67.
3. Klein S, Allison DB, Heymsfield SB, et al. Waist Circumference and Cardiometabolic Risk: a Consensus Statement from Shaping America's Health: Association for Weight Management and Obesity Prevention; NAASO,

the Obesity Society; the American Society for Nutrition; and the American Diabetes Association. *Obesity (Silver Spring)* 2007; **15**(5): 1061-7.

4. Demerath EW, Guan W, Grove ML, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Human molecular genetics* 2015; **24**(15): 4464-79.
5. Bose M, Wu C, Pankow JS, et al. Evaluation of microarray-based DNA methylation measurement using technical replicates: the Atherosclerosis Risk In Communities (ARIC) Study. *BMC bioinformatics* 2014; **15**: 312.
6. Demetriou CA, van Veldhoven K, Relton C, Stringhini S, Kyriacou K, Vineis P. Biological embedding of early-life exposures and disease risk in humans: a role for DNA methylation. *European journal of clinical investigation* 2015; **45**(3): 303-32.
7. Xu K, Zhang X, Wang Z, Hu Y, Sinha R. Epigenome-wide association analysis revealed that SOCS3 methylation influences the effect of cumulative stress on obesity. *Biological psychology* 2018; **131**: 63-71.
8. Loucks EB, Huang YT, Agha G, et al. Epigenetic Mediators Between Childhood Socioeconomic Disadvantage and Mid-Life Body Mass Index: The New England Family Study. *Psychosomatic medicine* 2016; **78**(9): 1053-65.
9. Heyn H, Esteller M. DNA methylation profiling in the clinic: applications and challenges. *Nature reviews Genetics* 2012; **13**(10): 679-92.
10. Green ED, Guyer MS, National Human Genome Research I. Charting a course for genomic medicine from base pairs to bedside. *Nature* 2011; **470**(7333): 204-13.