ARIC Manuscript Proposal # 3146

PC Reviewed: 04/10/2018 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: DNA methylation in Relation to Diet Quality in the CHARGE Consortium

b. Abbreviated Title (Length 26 characters): DNA methylation and diet quality

2. Writing Group:

Jiantao Ma[1]
Lindsay Reynolds[2]
Kim Braun[3]
Casey M. Rebholz[4]
Rachel Hennein[1]
Emily A. Hu[4]
Rui Xia[5]
Chunyu Liu[1]
Tianxiao Huan[1]
Roby Joehanes[1]
Alexis Frazier-Wood[2]
Trudy Voortman[3]
Yongmei Liu[2]
Myriam Fornage[5]
Daniel Levy[1]

1. Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland and the Framingham Heart Study

2. Department of Epidemiology and Prevention, Wake Forest School of Medicine

3. Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands

4. Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health

5. The Brown Foundation Institute of Molecular Medicine for the Prevention of Human Diseases at The University of Texas Health Science Center at Houston

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

First author: Jiantao Ma, MBBS, PhD
Address: Framingham Heart Study
73 Mt. Wayne Avenue, Suite 2

Framingham, MA 01702-5827

Phone: Fax:
E-mail: jiantao.ma@nih.gov

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: Casey Rebholz
Address: Johns Hopkins Bloomberg School of Public Health
2024 E. Monument Street, Suite 2-600
Baltimore, MD 21287

Phone: 410-502-2049 Fax:
E-mail: crebhol1@jhu.edu

3. Timeline: Phenotypic dataset preparation will be completed by Casey Rebholz and Emily Hu and shared with Myriam Fornage. Data analysis (to be conducted by Myriam Fornage) will be conducted upon approval of this manuscript proposal. Results (not data) will be shared with Jiantao Ma for meta-analysis with other cohorts. Manuscript preparation will begin in summer 2018. We anticipate that a manuscript draft will be prepared by summer 2019 for submission to the ARIC Publications Committee.

4. Rationale:
Prior studies have shown that overall diet quality may have beneficial effects on cardiovascular disease (CVD) (1-4). Prior population-based studies have utilized dietary scores developed to assess diet quality, such as the Mediterranean-style diet score (MDS) and the Alternative Healthy Eating Index (AHEI) score (1, 5-7). In general, adherence to favorable dietary patterns, as assessed by these scores, is associated with less weight gain, lower risk of incident cardiovascular disease, and decreased risk of death (6-9). However, there is a gap in knowledge of the biological underpinnings of the beneficial effects of the high-quality diet. For example, several studies suggest that diet may modify DNA methylation, a major epigenetic phenomenon, and potentially affect metabolic phenotypes (10-13). To date, little is known about the DNA methylation signature associated overall diet quality.

5. Main Hypothesis/Study Questions: To determine the association between a Mediterranean-style dietary pattern (MDS) and DNA methylation profiles and between the Alternative Heath Eating Index (AHEI) score and DNA methylation profiles.

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 sample:

Cohort	Ethnicity	N	Methylation	Dietary assessment

Framingham Heart Study (FHS) [including Offspring and Third Generation cohorts]	Caucasian	~3300	Illumina 450K	Based on FFQ
Rotterdam Study	Caucasian	951	Illumina 450K	Based on FFQ
Atherosclerosis Risk in Communities (ARIC) Study	Caucasian African American	4,000	Illumina 450K	Based on FFQ
Multi-Ethnic Study of Atherosclerosis (MESA)	Caucasian Hispanic African American	1250	Illumina 450K	Based on FFQ

Exclusion criteria (criterion 1 is required, others depend on cohort availability):

1. History of myocardial infarction or stroke
2. Cancer (excluding non-melanoma skin cancer)
3. Bariatric surgery

4*. Heavy alcohol consumers, defined as 21 drinks/week in men or 14 drinks/week in women. A drink is defined as 12 ounces of beer, 4-5 ounces of wine, or 1.5 ounces of liquor, where one drink is equivalent to ~14 g of ethanol.

We will conduct two analyses. The primary analysis will use sample after applying exclusion criteria 1 to 3. In sensitivity analysis, we will also exclude heavy alcohol consumers.

Dependent variable: Methylation beta score (i.e., proportion of CpGs that are methylated, ranging from 0 to 1) normalized using appropriate method. This project will analyze CpGs at autosomal chromosomes. The Framingham Heart Study used DASEN method to normalize methylation beta score. In addition, we calculated surrogate variables using SVA package in R. We then calculated the residuals for methylation beta score using linear regression model with adjustment for sex, age, and surrogate variables that associated with MDS ($P<0.1$).

Independent variable: MDS and AHEI score. Detailed interpretation regarding the two scores are shown in the separate documents (slices for developing the two dietary scores). Scores are standardized by cohort-specific standard deviation.

Covariates:

1. Sex (men or women, dichotomous)
2. Age (years, continuous)
3. Energy intake (kcal/day, continuous), assessed using cohort-specific dietary assessment tool such as food frequency questionnaire, 24-hour dietary recall, or weighted food record.
4. Smoking status. Current smoker yes or no, but may vary by cohort (e.g., current smoker, former smokers, & never).

a. Current smoker: Study volunteer who has smoked at least one cigarette a day in the past year.
b. Former smoker: Study volunteer who has previously smoked at least one cigarette a day but has stopped for at least one year.
c. Never smoker: Study volunteer who has never smoked.

5. Physical activity level. Cohort-specific definition. Preliminary analysis in the Framingham Heart Study used questionnaire based physical activity score.
6. BMI (kg/m^2)
7. White cell counts. Cohort-specific data. The Framingham Heart Study used estimated cell counts for CD8T, CD4T, NK, Bcell, and Mono based on approach proposed by Houseman et al. Omit this if your cohort has methylation from single cell type.
8. Lab for methylation assessment. As needed by cohort.
9. Technical variables. Cohort-specific. Preliminary analysis in the Framingham Heart Study calculated residuals for methylation beta score after adjustment for surrogate variables and used the residuals as dependent variables.

Statistical model (Please perform analysis in overall sample as well as in each ethnicity, i.e., in Caucasian, Hispanic, and American African separately):

Models:*

Model 1: Residuals of methylation beta score = MDS/AHEI + sex + age + energy intake + ethnicity + cell counts + labs

Model 2: Residuals of methylation beta score = MDS/AHEI + sex + age + energy intake + ethnicity + smoking status + physical activity level + cell counts + labs

Model 3 **: Residuals of methylation beta score = Model 2 covariates + BMI

* Use multiple linear regression model or linear mixed model for family data.

** Model 3 are secondary model.

Cohort-specific analysis and meta-analysis

1. Discovery-Replication: The primary analysis is using model 1 results in Caucasian samples.
   a. Discovery EWAS in the FHS Caucasian samples
      i. False discovery rate (FDR) < 0.05
   b. Replication in Caucasian samples of ARIC, MESA, and RS
      i. Cohort-specific analysis in ARIC, MESA, and RS Caucasian samples.
      ii. meta-analysis using fixed effect models or random effect models (if I-squared > 25%).
      iii. Bonferroni corrected p-value (0.05/number of discovered CpGs)

2. We will run parallel discovery-replication analysis using model 2 and 3 in Caucasian samples. To increase sample size, we will meta-analyze all Caucasian samples. Same fixed or random effect model will be used. We will compare the three models in this analysis.

3. Separate analysis will be conducted in MESA Hispanic samples. We will test if there are CpGs significant at FDR < 0.05. We will test if the significant CpGs are differentially methylated in Caucasian samples (Bonferroni correction will be applied, i.e., 0.05/the number of significant CpGs). We will also test if replicated CpGs in Caucasian samples are significant in Hispanic samples (at Bonferroni corrected p-value threshold).

4. In African American samples. EWAS will be conducted in African American samples in each cohort. We will conduct fixed effect models or random effect models (if I-squared > 25%) to meta-analyze cohort-specific results. Similar in Hispanic samples, we will select CpGs (at FDR < 0.05) and test if they are differentially methylated in Caucasian samples (Bonferroni-corrected threshold). we will also test if replicated CpGs in Caucasian samples are significant in African American samples.
5. We will meta-analyze all samples to test if there are common differentially methylated CpGs in all ethnicities.

Exploration of potential function of identified CpGs: We will test the association between diet scores, identified CpGs and gene expression levels in cohorts with available data. In addition, we will conduct gene set enrichment analysis, pathway analysis, enrichment analysis for localization to different genomic features.

Testing association of identified CpGs and change in CVD risk factors: We will test the prospective association of CpGs with longitudinal changes in CVD risk factors including waist circumference, systolic blood pressure, diastolic blood pressure, fasting glucose, fasting insulin, fasting triglycerides, fasting HDL, and fasting LDL (measured or calculated using the Friedewald equation). We will also test if CpGs predict the development of incident metabolic syndrome, hypertension, hypertriglyceridemia, low HDL, and prediabetes and type 2 diabetes. In this analysis, we test will if and to what extent the observed association between CpGs and CVD risk factors is mediated by MDS. This analysis will be conducted in all cohorts with available data. A meta-analysis will be implemented to analyze cohort-specific results. In addition to analyze single CpG, we will create a weighted methylation score using all significant CpGs and test its association with CVD risk factors.

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.cscc.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 #2322: DNA methylation-related SNPs interact with fatty acids and triglycerides

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, approved manuscripts using 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

References

1. Fung TT, Rexrode KM, Mantzoros CS, Manson JE, Willett WC, Hu FB. Mediterranean diet and incidence of and mortality from coronary heart disease and stroke in women. Circulation 2009;119(8):1093-100. doi: 10.1161/CIRCULATIONAHA.108.816736.
2. Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. The New England journal of medicine 2003;348(26):2599-608. doi: 10.1056/NEJMoa025039.
3. Estruch R, Ros E, Salas-Salvado J, Covas MI, Corella D, Aros F, Gomez-Gracia E, Ruiz-Gutierrez V, Fiol M, Lapetra J, et al. Primary prevention of cardiovascular disease with a

Mediterranean diet. The New England journal of medicine 2013;368(14):1279-90. doi: 10.1056/NEJMoa1200303.
4. Stewart RA, Wallentin L, Benatar J, Danchin N, Hagstrom E, Held C, Husted S, Lonn E, Stebbins A, Chiswell K, et al. Dietary patterns and the risk of major adverse cardiovascular events in a global study of high-risk patients with stable coronary heart disease. European heart journal 2016;37(25):1993-2001. doi: 10.1093/eurheartj/ehw125.
5. Antonia Trichopoulou TC, Christina Bamia, Dimitrios Trichopoulos. Adherence to a Mediterranean Diet and Survival in a Greek Population. N Engl J Med 2003(348):2599.
6. Sotos-Prieto M, Bhupathiraju SN, Mattei J, Fung TT, Li Y, Pan A, Willett WC, Rimm EB, Hu FB. Association of Changes in Diet Quality with Total and Cause-Specific Mortality. The New England journal of medicine 2017;377(2):143-53. doi: 10.1056/NEJMoa1613502.
7. Chiuve SE, Fung TT, Rimm EB, Hu FB, McCullough ML, Wang M, Stampfer MJ, Willett WC. Alternative dietary indices both strongly predict risk of chronic disease. The Journal of nutrition 2012;142(6):1009-18. doi: 10.3945/jn.111.157222.
8. Fung TT PA, Hou T, et al. Long-Term Change in Diet Quality Is Associated with Body Weight Change in Men and Women. J Nutr 2015;145(8):1850-6.
9. Sotos-Prieto M, Bhupathiraju SN, Mattei J, Fung TT, Li Y, Pan A, Willett WC, Rimm EB, Hu FB. Changes in Diet Quality Scores and Risk of Cardiovascular Disease Among US Men and Women. Circulation 2015;132(23):2212-9. doi: 10.1161/CIRCULATIONAHA.115.017158.
10. Ceccarelli V, Racanicchi S, Martelli MP, Nocentini G, Fettucciari K, Riccardi C, Marconi P, Di Nardo P, Grignani F, Binaglia L, et al. Eicosapentaenoic acid demethylates a single CpG that mediates expression of tumor suppressor CCAAT/enhancer-binding protein delta in U937 leukemia cells. The Journal of biological chemistry 2011;286(31):27092-102. doi: 10.1074/jbc.M111.253609.
11. Burdge GC, Lillycrop KA, Phillips ES, Slater-Jefferies JL, Jackson AA, Hanson MA. Folic acid supplementation during the juvenile-pubertal period in rats modifies the phenotype and epigenotype induced by prenatal nutrition. The Journal of nutrition 2009;139(6):1054-60. doi: 10.3945/jn.109.104653.
12. Li Q, Chen H. Epigenetic modifications of metastasis suppressor genes in colon cancer metastasis. Epigenetics 2011;6(7):849-52. doi: 10.4161/epi.6.7.16314.
13. Arpon A, Riezu-Boj JI, Milagro FI, Razquin C, Martinez-Gonzalez MA, Corella D, Estruch R, Casas R, Fito M, Ros E, et al. Adherence to Mediterranean diet is associated with methylation changes in inflammation-related genes in peripheral blood cells. Journal of physiology and biochemistry 2017. doi: 10.1007/s13105-017-0552-6.