PC Reviewed: 08/09/16   **Status: _____**   **Priority: ____**
SC Reviewed: _____   **Status: _____**   **Priority: ____**

1.a. Full Title: B-vitamin Dietary Intake & DNA methylation: The CHARGE Consortium

  b. Abbreviated Title (Length 26 characters): EWAS of B-vitamin

2.   Writing Group:
        Writing group members: CHARGE Epigenetics and Nutrition Working Groups, ARIC Members include Anne Justice, Misa Graff, Kari North, and other investigators welcome

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

Contact 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-August 2016
   • Interpretation and meta-analyses: September-December 2016
   • Manuscript preparation: January- February 2017
   • Manuscript submission: March 2017

4. Rationale: One carbon metabolism is an important biological network which is linked to DNA methylation via several metabolites, enzyme activities and dietary micronutrients. These factors serve as methyl donors to homocysteine, which in turn incorporates the methyl groups to the DNA by s-adenosylmethionine. One of the widely studied groups of methyl-donor micronutrients are B-vitamins (Figure 1) [1,2]. Their role in regards to clinical outcomes are fairly known; however, little is studied about their influence in changing epigenetic patterns [1,3-5]. A few studies have investigated their associations with global (LINE-1) methylation [6-8], but their association with methylation at specific locations on a genome-wide scale remains unknown.

5.   Main Hypothesis/Study Questions:
The objective of this study is to investigate the associations of intake of methyl-donating B-vitamins (Vitamin B2, B6, B12, folate & choline) as well as intake of methionine with leukocytic genome-wide DNA methylation, measured using Infinium Illumina HM450k arrays in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium.

Figure 1. Diagram illustrating the role of methyl-donors in the one-carbon pathway that are of interest in the current study.
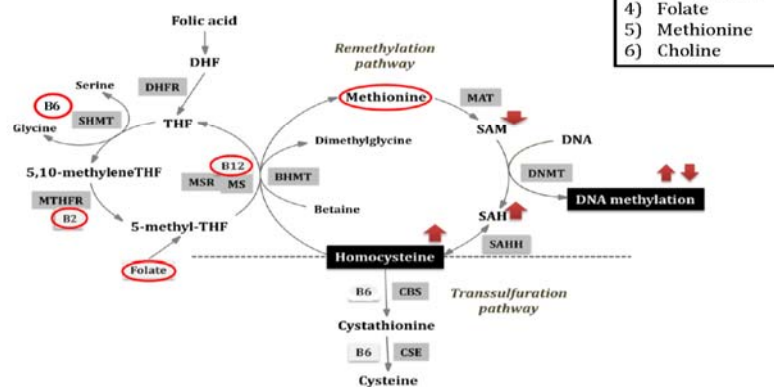

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

<u>Study design</u>: This study will be a meta-analysis of study-specific association results available through the CHARGE Consortium (Table 1). For the ARIC study results, we will use cross-sectional data taken from **visit 2 or 3** (dependent on availability of nutritional data for those with methylation typing, N≈562) in the ARIC African American participants.

<u>Exclusion</u>:
- Individuals < 18 years of age
- Pregnant women
- Missing epigenetic data
- Missing or unrealistic values (e.g. Participants with almost no (<500) or extreme energy (>5000) intake must be excluded)
- Missing covariates
- Subjects with cancer

<u>Exposures</u>:
- Vitamin B2
- Vitamin B6
- Vitamin B12
- Folate
- Methionine
- Choline

<u>Outcome data</u>:
- Genome-wide DNA methylation Beta values assessed by 450k arrays

<u>Biological covariates:</u>

- Total energy Intake
- Total protein Intake only for the "methionine intake data", since methionine intake is associated with Total protein intake.
- Age
- Sex
- BMI
- Technical covariates (e.g. Differential % white blood cell counts, batch effects, row effects).
- Smoking status (Current/Former/Never)

Table 1. Replication cohorts with nutrient data available through CHARGE Consortium.

Cohort Name	Ethnicity(ies)	Sample Size
FHS/FHSgen2	European Descent	2,498
GOLDN	European Descent	991
RS-III	European	457
RS-BIOS	European	642
CODAM	European	186
LBC36	European	900
CHS	European	227
MESA	European, African, and Hispanic Americans	1,264
TwinsUK	European	708
YFS	European	167
ARIC	African American	562

- Physical Activity
  - Note: Following the harmonization methods in Kilpeläinen et al. [9] categorizing it into a dichotomous variable (physically inactive versus active).
- Supplement Intake for Folate/Vitamin B12/Multivitamins/B-vitamins (Users/Non-users)

Summary data analysis:
Exposures
We will examine both tertiles of nutrient intake and nutrient intake in continuous values.

Tertiles: For the Residual variables of each nutrient, we will calculate tertiles, coded as 0,1, and 2. Tertiles will be used for the analysis with linear mixed models. No outliers are removed since we are using tertiles.

Continuous: In addition to tertiles, separate analysis will also be done on continuous variables. Outliers in this case will be removed for each exposure that are $< Q1 - (M \times IQR)$ or $> Q3 + (M \times IQR)$, where Q1 is the Quartile 1 value, Q3 is Quartile 3 value, IQR is the observed interquartile range (IQR), and M, a multiplying factor, is equal to 3.

Outcome Variables
Probe filtering: Probe filters may be study specific. Probes which fail the general QC criteria should be removed (e.g. missing rate ≥5%)

Normalization: Depends on the cohort. ARIC used BMIQ.

Association Analyses
Linear Mixed Models will be used, wherein the technical covariates are treated as random factors and biological covariates are treated as fixed factors. Each analysis will be performed stratified by dietary supplement status and ancestry. Therefore, each nutrient will have a total of four analyses per ancestry (e.g. 1) Tertiles: Dietary Supplement-Yes; 2) Tertiles: Dietary Supplement-No; 3) Continuous: Dietary Supplement-Yes; 4) Continuous: Dietary Supplement-No), for a maximum of 24 models across six dietary exposures. Note: ARIC does not have data on choline, and therefore will have only 20 models (4 x B2, 4 x B6, 4 x B12, 4 x methionine, 4 x folate).

Meta-Analysis
Given the different analytic strategies and cell types used among the participating studies, we will conduct a sample size-weighted meta-analysis implemented in METAL. Significance will be assessed based on P<0.05/# of methylation sites examined (~480,000 sites or P<1E-7).

Limitations/challenges: The methods discussed herein may be supplemented or altered as newly established methods develop or should harmonization issues arise among participating studies.

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)?**
N/A

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

1. Plazar N, Jurdana M. Hyperhomocysteinemia and the role of B vitamins in cancer. *Radiology and oncology.* Jun 2010;44(2):79-85.
2. Selhub J. Folate, vitamin B12 and vitamin B6 and one carbon metabolism. *The journal of nutrition, health & aging.* 2002;6(1):39-42.
3. Lippi G, Plebani M. Hyperhomocysteinemia in health and disease: where we are now, and where do we go from here ? *Clinical chemistry and laboratory medicine.* Dec 2012;50(12):2075-2080.
4. Ramakrishnan S, Sulochana KN, Lakshmi S, Selvi R, Angayarkanni N. Biochemistry of homocysteine in health and diseases. *Indian journal of biochemistry & biophysics.* Oct 2006;43(5):275-283.
5. Strain JJ, Dowey L, Ward M, Pentieva K, McNulty H. B-vitamins, homocysteine metabolism and CVD. *The Proceedings of the Nutrition Society.* Nov 2004;63(4):597-603.
6. Perng W, Villamor E, Shroff MR, et al. Dietary intake, plasma homocysteine, and repetitive element DNA methylation in the Multi-Ethnic Study of Atherosclerosis (MESA). *Nutrition, metabolism, and cardiovascular diseases : NMCD.* Jun 2014;24(6):614-622.

7. Boeke CE, Baccarelli A, Kleinman KP, et al. Gestational intake of methyl donors and global LINE-1 DNA methylation in maternal and cord blood: prospective results from a folate-replete population. *Epigenetics.* Mar 2012;7(3):253-260.

8. Takumi S, Okamura K, Yanagisawa H, Sano T, Kobayashi Y, Nohara K. The effect of a methyl-deficient diet on the global DNA methylation and the DNA methylation regulatory pathways. *Journal of applied toxicology : JAT.* Dec 2015;35(12):1550-1556.

9. Kilpelainen TO, Qi L, Brage S, et al. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS medicine.* Nov 2011;8(11):e1001116.