ARIC Manuscript Proposal #2927

PC Reviewed: 02/03/17	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title:
Meta-analysis of genome-wide DNA methylation and serum CRP levels

b. Abbreviated Title (Length 26 characters):
Epigenetics and CRP levels

2. Writing Group:
ARIC: Pascal Schlosser, Liz Selvin, Weihua Guan, Eric Boerwinkle, Anna Köttgen, additional members of the ARIC epigenetics working group as authorship guidelines in the international meta-analysis, to which the data will be contributed, allow

The international consortium that ARIC will join with this data is coordinated by John C. Chambers, Marjo-Riita Jarvelin, Benjamin C. Lehne, and Matthias Wielscher from Imperial College, London, UK. The following cohorts have currently agreed to contribute data to the discovery step in this consortium: LOLIPOP, NFBC studies, DFKZ, KORA. The initiators are currently considering contacting additional CHARGE studies.

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

First author: **Pascal Schlosser, M.Sc.**
University Medical Center Freiburg
Institute for Medical Biometry and Statistics
Division of Genetic Epidemiology
Stefan-Meier-Str. 26
79104 Freiburg, Germany
Phone: +49 761 203 7704
E-mail: schlosser@imbi.uni-freiburg.de

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: **Anna Köttgen, MD MPH**
Address: Dept. of Epidemiology
Johns Hopkins School of Public Health
Baltimore, MD 21205
akottge1@jhu.edu
Postal Address: University Medical Center Freiburg
Division of Genetic Epidemiology
Hugstetter Str. 49, 79106 Freiburg, Germany

phone: +49 761 270-78050

3. Timeline:
Data analysis will start immediately upon ARIC manuscript approval. A manuscript is expected to be prepared within 12 months.

4. Rationale:

Chronic inflammation plays a significant role in the development of several disorders such as type 2 diabetes, hypertension, cardiovascular disease as well as cancer and chronic obstructive pulmonary disease.[1-4] Here we propose to link measurements of serum hsCRP levels as a proxy for chronic inflammation to blood DNA methylation levels to gain insights into the underlying mechanisms, because the actual mechanisms how inflammation may impact the development of the above-mentioned diseases is poorly understood. DNA methylation as a regulator of gene expression is closely intertwined with DNA sequence and environmental exposures. Our study aims to improve insights into this connection between DNA methylation, chronic inflammation and related diseases by conducting a large-scale forward epigenome-wide screen to illuminate underlying pathways. A recently published paper from the CHARGE Consortium illustrates the potential of the analytical approach.[5] In this paper, Ligthart and colleagues identified and replicated 58 differentially methylated CpG sites in association with serum CRP concentrations, which together explained 6% of the CRP variation and were associated with cardiovascular traits and diseases. The discovery step of this study was carried out among individuals of European ancestry, with replication among African Americans. Trans-ethnic discovery EWAS provide an added opportunity to identify additional epigenetic factors associated with low-grade inflammation.

5. Main Hypothesis/Study Questions:

There will be associations between serum CRP levels and some CpG specific DNA methylation patterns in whole blood and/or peripheral blood mononuclear cells.

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).

Inclusion criteria:
All African American ARIC participants with HM450K data passing quality control with available data on high-sensitivity serum CRP levels (hsCRP). Outliers (>4 SD from the mean in ln(hsCRP) will be excluded, for a final sample size of N≤2802.

Outcomes:
Primary outcome: hsCRP at ARIC visit 2, when DNA methylation was quantified. hsCRP was measured as part of Dr. Selvin's ancillary study (AS # 2009.16) in serum using a latex-particle enhanced immunoturbidimetric assay kit (Roche Diagnostics, Indianapolis, IN 46250) and read on the Roche Modular P Chemistry analyzer (Roche Diagnostics). The reference range is 0 – 5 mg/L. The laboratory inter-assay CV is 4.5%.

Predictor:
Methylation levels at 473,788 CpG sites passing quality control. All DNA methylation data distributed as released by the ARIC epigenetics working group. Data is already transferred to the compute cluster at JHU, where it is being used for other epigenome-wide association studies.

Covariates:
Visit 2 age, gender, race, BMI, smoking, diabetes and glycemic measures, estimated cell type (Houseman method), white blood cell count if available, 10 principal components generated using Affymetric 6.0 autosomal genotype data to control for population substructure and for integration with genetic information. Further covariates may be needed for sensitivity analyses depending on the findings.

Data analysis:

We will collaborate with Dr. Lehne from Imperial College London, who is coordinating a large international meta-analysis combining results from EWAS of hs-CRP across studies, to which the ARIC study was invited to contribute data. The meta-analysis will include more than 7 multi-ethnic studies with a total of >15,000 participants. To allow for meta-analysis of data from different studies, we will use the cleaned and normalized data according to the ARIC epigenetics working group protocol and then follow a central analysis plan.

Specifically, quality-filtering criteria of the methylation data in the ARIC Study will be the same as those reported in Demerath et al.[6] We will work with the beta values based on the BMIQ-normalized (PMID: [7] and cleaned data distributed by the ARIC Epigenetics working group. Beta values will be evaluated for their distribution and inverse normal transformed if necessary. To adjust for white blood cell subtype composition and batch effects, linear mixed effect models will be used with chip number as the random effect and plate, chip row, and imputed cell counts of neutrophils, lymphocytes, monocytes and eosinophils using the Houseman method (distributed with the data[6]), as fixed effects to generate residuals of the beta values for association analysis with CRP concentrations. Technical covariates for adjustment are 10 principle components generated from available genome-wide genetic SNP data on these individuals. Additionally, we will conduct sensitivity analyses using estimated cell type composition using known cell-type specific sites as the reference for imputation (Houseman approach, rather than using measured differentials as the reference).

The association between methylation levels at each individual CpG site i and hsCRP concentrations will be evaluated using linear regression modeling, using two models:

1) *Model_1:* lnCRP ~ res_{beta}[i,] + age + sex + technical covariates + HousemanEstimates

2) Model_2: lnCRP ~ res_{beta}[i,] + age + sex + technical covariates + HousemanEstimates+ BMI

Summary statistics will then be shared with our collaborators at Imperial College and be used for a meta-analysis across 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? _X_ 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)?

#1928: Genome-wide methylation analyses of cardiovascular disease (CVD) and its risk factors
#2576: Epigenome-wide association study of incident chronic kidney disease
2176 r: Identification and Characterization of Genetic Risk Associated with Measures of Thyroid Function and Disease

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

11.b. If yes, is the proposal
X A. primarily the result of an ancillary study (list number* __ 2009.16__)
___ 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 Donath, M. Y. & Shoelson, S. E. Type 2 diabetes as an inflammatory disease. *Nature reviews. Immunology* **11**, 98-107, doi:10.1038/nri2925 (2011).

2 Hansson, G. K. & Hermansson, A. The immune system in atherosclerosis. *Nature immunology* **12**, 204-212, doi:10.1038/ni.2001 (2011).

3 Fernandes, J. V. *et al.* The role of the mediators of inflammation in cancer development. *Pathology oncology research : POR* **21**, 527-534, doi:10.1007/s12253-015-9913-z (2015).

4 Provinciali, M., Cardelli, M. & Marchegiani, F. Inflammation, chronic obstructive pulmonary disease and aging. *Current opinion in pulmonary medicine* **17 Suppl 1**, S3-10, doi:10.1097/01.mcp.0000410742.90463.1f (2011).

5 Ligthart, S. *et al.* DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome biology* **17**, 255, doi:10.1186/s13059-016-1119-5 (2016).

6 Demerath, E. W. *et al.* Epigenome-wide Association Study (EWAS) of BMI, BMI Change, and Waist Circumference in African American Adults Identifies Multiple Replicated Loci. *Hum Mol Genet*, doi:10.1093/hmg/ddv161 (2015).

7 Teschendorff, A. E. *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189-196, doi:10.1093/bioinformatics/bts680 (2013).