ARIC Manuscript Proposal # 1717

PC Reviewed: 11/9/10	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Contribution of variants in a priori defined structural, miRNAs and epigenetic genes to predicting clinically low plasma HDL-C

b. Abbreviated Title (Length 26 characters):

2. Writing Group:

Writing group members: C. M. Lusk, T. J. Rea, G. Dyson, K. Volcik, A. Brautbar, E. Boerwinkle, C.M. Ballantyne, C. F. Sing

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

First author: Christine Lusk
Address: University of Michigan
Dept. of Human Genetics
1241 E. Catherine St./5928 Buhl Bldg.
Ann Arbor, MI 48109

Phone: 734-647-3151 Fax: 734-763-5277
E-mail: cmbz@umich.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: Eric Boerwinkle
Address: University of Texas Health Science Center at Houston
RAS E447
1200 Herman Pressler Dr.
Houston, Texas 77030

Phone: 713-500-9914 Fax: 713-500-0900
E-mail: Eric.Boerwinkle@uth.tmc.edu

3. Timeline: Manuscript to be completed by December 2010

4. Rationale: Despite progress made by GWAS and candidate gene studies in identifying associations between HDL-C and individual SNPs, the genetic architecture (i.e., the number of genes involved, the number of variants and the relative frequencies of these variants) of HDL metabolism remains to be resolved. By considering the gene region, rather than individual SNPs, as the unit of inference, we evaluate the added value of structural candidate genes, miRNAs and genes involved in epigenetic regulation to explaining variability in low HDL-C beyond the contribution of lipid related covariates.

5. Main Hypothesis/Study Questions:
What is the contribution of variation in structural candidate gene regions, miRNA regions and epigenetic gene regions to variability in low HDL-C risk beyond the contribution of lipid related covariates (e.g., age, BMI, smoking, etc.)? Which multilocus genotypes predict low HDL-C in the gene regions that make a significant contribution to explaining variation in low HDL-C beyond the lipid related covariates?

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).
We evaluate high/low HDL-C (dichotomized at 45 mg/dl) in 8,929 female (N=4,755) and male (N=4,174) European-American and XXXX female (N=XXXX) and male (N=XXXX) African-American participants in the ARIC study with clinical examination data collected at the baseline (1987-1989) exam and DNA samples available. We consider 1,739 SNPs genotyped on either the Affy 6.0 platform or the IBC 50k SNP Array in 22 gene regions (six structural candidate regions, six miRNA regions and 10 epigenetic regions) for our analysis. Selection of structural candidate regions was based on evidence of significant, replicable effects on HDL-C from genome-wide association studies, combined with evidence of both epigenetic and miRNA regulation. The selected structural candidate regions are: *ABCA1*, *APOA1/C3/A4/A5*, *CETP*, *GALNT2*, *GRIN3A* and *LPL*. We selected miRNA genes if they were predicted to target at least one of the structural candidates and were not located in an mRNA transcript. Epigenetic regions included the DNA methyltransferases (N=5) and methyl-binding proteins (N=5). In each of the 22 regions we first evaluate whether any individual SNPs predict low HDL-C beyond genotype plates and a set of lipid related covariates by comparing two logistic regression models (reduced vs. complete), separately for females and males. The reduced model includes genotype plate, age, BMI, blood pressure, diabetes, geographic location, alcohol consumption and smoking status. The complete model includes all aforementioned covariates plus the SNP. Using backward stepwise logistic regression we next find a more parsimonious set of SNPs in each region and then use pseudo-R^2 measures to estimate the contribution of genotype plates, traditional risk factors and SNPs to variability in risk of low HDL-C. We finally determine whether there are multilocus combinations of SNPs in particular gene regions that contribute to trait variability beyond genotype plates and traditional risk factors in both females and males.

7.a. Will the data be used for non-CVD analysis in this manuscript? __X__ Yes ____ 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 ICTDER03 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"?
____ 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

_____ 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)?

The only manuscript close to this work is the GWAS work already published as part of the global lipids consortium.

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* _____)
_x__ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* __2006.05_____ _____ _____) The results are a joint effort of regular ARIC and this ancillary study.

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12. 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. I agree.