ARIC Manuscript Proposal # 1912

PC Reviewed: 3/20/12 **Status: A** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title:
APOE modulates the relationship among triglycerides, cholesterol, and CHD through pleiotropy and gene-gene interactions.

b. Abbreviated Title (Length 26 characters):
APOE, Lipids and CHD risk

2. Writing Group:
Writing group members:
Taylor J. Maxwell, Christie M. Ballantyne, James M. Cheverud, and Eric Boerwinkle (Other authors are invited to join if desired.)

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

First author: **Taylor J. Maxwell**
Address: 1200 Hermann Pressler, RAS E-447, Houston, Texas 77030

Phone: 713-500-9896 Fax: 713-500-0900
E-mail: Taylor.J.Maxwell@uth.tmc.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: 1200 Hermann Pressler, RAS E-531, Houston, Texas 77030

Phone: 713-500-9816 Fax: 713-500-0900
E-mail: Eric.Boerwinkle@uth.tmc.edu

3. Timeline:
The data is currently available and analyses will begin as soon as approval is granted.

4. Rationale:

To introduce the concept and potential of relationship loci (rQTL) and how it relates to pleiotropy and epistasis (gene-by-gene interaction) using Apolipoprotein E (APOE) as an example in humans for two lipid traits and incident coronary heart disease (CHD). This is motivated by a desire to extend the results of Boerwinke et al. (1987) who reported that the correlation between total cholesterol (TC) and triglycerides (TG) differed by APOE genotype. A rQTL is a locus that effects the relationship between two biological phenotypes and may or may not have any direct association with either phenotype. Based on theoretical and empirical work of James Cheverud and his lab using mouse strains (Pavlicev *et al.*, 2007; Pavlicev *et al.*, 2011), rQTL are likely to result in variation in pleiotropy and are typically exist due to gene-by-gene or gene-by-environment interactions

5. Main Hypothesis/Study Questions:

Hypothesis 1: The relationship between total cholesterol and triglycerides varies by APOE. This will establish that APOE is a relationship locus. We intend to test this separately in both European-American and African-Americans.

Hypothesis 2: If APOE is an rQTL for total cholesterol and triglycerides (hypothesis 1), we predict that APOE genotypes modulate the relationship between total cholesterol or triglycerides with incident CHD.

Hypothesis 3: If APOE is an rQTL for total cholesterol and triglycerides, we predict that other loci interact with APOE to effect one or both phenotypes. If APOE effects the correlation between total cholesterol or triglycerides and incident CHD (hypothesis 2), we predict that other loci interact with APOE to effect incident CHD.

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).

For all hypotheses we will only include individuals for which APOE genotype data is available. We will exclude individuals for which permissions were not granted for DNA use. Furthermore, we will use the same quality control criteria established in previous GWAS analyses based on sex mismatch, first degree relatives, outliers based on average identity by descent (Ikram et al, 2009). For Hypothesis 3, we will include only individuals for which both GWAS and APOE genotype data are available.

For **Hypothesis 1** we will test the significance of the interaction term in the following linear model separately in African-American and European-American populations.

$$TC_{ijk} = u + age + sex + bmi + APOE + TG + APOE * TG + e_{ijk}$$

The significance of the interaction term will be assessed using a full versus reduced model. Significance of this test rejects the null hypothesis that the relationship between

the two traits is equal across genotype classes and that the beta coefficient from each bivariate regression within genotypes does not differ.

For **Hypothesis 2**, we will use a Cox Proportional Hazards model framework in each population, to separately test an APOE by TC and an APOE by TG interaction term with incident CHD (ARIC variables ln_07sp & futimea) on the dependent side of the equation using a likelihood ratio test of full versus reduced models. Below is the the model including TG, the other model replaces TG with TC.

$$CHD_{ijk} = age + sex + bmi + APOE + TG + APOE * TG + e_{ijk}$$

Significance of either of these interaction terms rejects the null hypothesis that TC or TG related risk for incident CHD is equivalent across APOE genotypes. Most studies that use triglyceride levels use a natural log transformation (logTG) for analyses; however, the original analysis by Boerwinkle et al. (1987) did not. We will do both for comparison while only using logTG in the subsequent interaction analyses.

A significant rQTL (from the results of Hypothesis 1 or Hypothesis 2) creates a prior hypothesis (**Hypothesis 3**) that it interacts with other loci that affect one or both traits involved in the rQTL model (Pavlicev et al., 2011). In fact, it suggests that a locus that interacts with the rQTL either effects one trait and not the other, or it effects both traits but in opposing patterns of relationships with each trait (Pavlicev et al., 2008). For each phenotype (TC, logTG, incident CHD), we will perform a genome-wide scan for loci that interact with APOE at a genome-wide significance level. Because there is a prior hypothesis for each phenotype (from the significant rQTL model), we need only correct for genome-wide significance within each scan, not across all scans. Traditional bi-allelic two-locus models partition the 4 degrees of freedom for interaction into the additive-by-additive, additive-by-dominance, dominance-by-additive, dominance-by-dominance components (Cheverud 1995). Instead this type of partitioning, because we have three alleles for APOE, we have decided to treat the APOE genotypes as factors and to collapse the second locus into a simple continuous additive (-1, 0, 1) variable taking up only one degree of freedom which limits the interaction test to five degrees of freedom. Below is the general model where we are interested in the APOE*SNPadd interaction term.

$$TC_{ijk} = u + age + sex + bmi + APOE + SNP_{add} + APOE * SNP_{add} + e_{ijk}$$

For TC and logTG we will use a standard linear model and for incident CHD we will use an analogous Cox Proportional Hazards model. Significance of either of the interaction terms rejects the null hypothesis that the relationship of the additive parameterization of the SNP to the phenotype does not change depending on the APOE genotype (an vice versa); in other words if significant, the beta value for that SNP changes if you did analogous models among the APOE genotypes**.**

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 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"?
__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)?
I found a proposal by Franceschini (2006.03-GWACHD, 2007.02-CARE 1448) entitled "Genome-wide genotype-by-sex interaction of subclinical atherosclerosis phenotypes: the ARIC Study" that is related to interactions at a genome-wide level. However, the Franceschini manuscript is analyzing carotid artery wall thickness.

I could not find any manuscripts related to APOE, pleiotropy, or gene-by-gene interactions.

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* __2006.03 & 2007.2)
___ 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.