ARIC Manuscript Proposal #1985

PC Reviewed: 8/14/12  Status: A  Priority: 2
SC Reviewed: _________  Status: _____  Priority: ____

1.a. Full Title:
Relationship loci (rQTL) among lipid traits, blood pressure, incident CHD and loci that interact with them

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

2. Writing Group:
Writing group members:
Taylor J. Maxwell, Christie M. Ballantyne, James M. Cheverud, Cameron S. Guild, Chiadi E. Ndumele, 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 build on the work of our analyses that showed APOE as an example of an rQTL and demonstrate that there are rQTL (relationship loci) that significantly effect the correlations between lipids and between lipids and incident CHD. A relationship locus (rQTL) is a locus that affects 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. These rQTL loci may be important loci that can modify the relationship between know risk factors (lipids and blood pressure) and CHD. They may also give us clues to genes and biological pathways that connect lipids to each other and to CHD.

5. Main Hypothesis/Study Questions:

Hypothesis 1: There are loci (rQTL) in the genome where the relationship between pairs of lipid traits with each other and between systolic (SBP) and diastolic blood (DBP) pressure (not with lipids) that significantly differ by genotype (i.e. APOE for Total Cholesterol and Triglycerides). We intend to test this separately in both European-American and African-Americans.

Hypothesis 2: Because lipids and blood pressure are risk factors for incident CHD, significant rQTL from hypothesis 1 are likely to modulate the relationship between the lipids in the model and CHD and represent a priori tests within the population it was found.

Hypothesis 3: There are loci (rQTL) in the genome where the relationship between CHD and individual lipid and blood pressure traits differs by genotype.

Hypothesis 4: Significant rQTL from hypotheses 1-3 interact with other loci to effect one or both of the traits in the original rQTL model.

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 GWAS 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 1 we will perform a genome-wide screen for each pair of lipids (total cholesterol, LDL, HLD, Triglycerides) and among systolic and diastolic blood pressure to test the significance of the interaction term in the following linear model separately in African-American and European-American populations.

\[ Trait_{ijk} = u + age + sex + bmi + SNP_{Genotypic} + Trait2 + SNP_{Genotypic} \cdot Trait2 + e_{ijk} \]
Where trait1 and trait2 represent two lipid traits in that particular scan. Age, sex, bmi, and cholesterol medication will be used as covariates in the lipid models and hypertension medication will replace cholesterol medication in the systolic and diastolic blood pressure models. We will not pair total cholesterol with LDL or HDL. The SNP is treated as a factor by genotypes where only SNPs with at least 2 genotypes having 10 or more observations, genotypes with fewer than 10 observations will be excluded from the test. 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. Significance will be based on typical genome-wide significance thresholds for each trait pair scan. The natural log of triglycerides will be used in all analyses.

For Hypothesis 2, for each significant rQTL within a population from hypothesis 1, we will fit two models, one for each lipid trait in the original rQTL model with respect to incident CHD (ARIC variables ln_07sp & futimea) in an analogous Cox Proportional Hazards model framework. Below is the model including where Trait represent one of the two lipid or blood pressure traits from the original significant rQTL model.

\[
CHD_{ij} = \text{age} + \text{sex} + \text{bmi} + \text{SNP}\text{genotype} + \text{Trait} + \text{SNP}\text{genotype} \times \text{Trait} + e_{ijk}
\]

Significance of the interaction term rejects the null hypothesis that the lipid trait related risk for incident CHD is equivalent across the rQTL genotypes. The significance of the test will be based on a likelihood ratio test. Because these tests are motivated by significant results from hypothesis 1, nominal (alpha = 0.05) significance threshold will suffice.

For Hypothesis 3 we will perform a genome-wide screen for each CHD/Lipid pair to test the significance of the interaction term in the using the same Cox Proportional Hazards Model as in hypothesis 2 separately in African-American and European-American populations. Typical genome-wide significance thresholds will be use to assess significance.

A significant rQTL (from the results of Hypotheses 1-3) creates a prior hypothesis (Hypothesis 4) 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 from a given significant rQTL model, 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. Because our original rQTL model was by genotypes we continue to treat them as genotypic 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 at most 2 degrees of freedom. Below are the general models (linear model & cox proportional hazards model) where we are interested in the rQTL*SNP_{add} interaction term.

\[
\text{Trait}_{ijk} = \text{u} + \text{age} + \text{sex} + \text{bmi} + r\text{QTL}_{\text{genotype}} + \text{SNP}_{\text{add}} + r\text{QTL}_{\text{genotype}} \times \text{SNP}_{\text{add}} + e_{ijk}
\]
\[ CHD_{ij} = age + sex + bmi + rQTL_{genotype} + SNP_{add} + rQTL_{genotype} \cdot SNP_{add} + e_{ijk} \]

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 rQTL genotype (and vice versa); in other words if significant, the beta value for that SNP changes among analogous models within each rQTL genotype. Scans will only be done for phenotypes from significant rQTL models and genome-wide significance thresholds will be employed.

7.a. Will the data be used for non-CVD analysis in this manuscript?  ____ 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”?  ____ 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.csc.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 pleiotropy and gene-by-gene interactions other than my previous ARIC manuscript proposal #1912, “APOE modulates the relationship
among triglycerides, cholesterol, and CHD through pleiotropy and gene-gene interactions”. This previous proposal is a specific case of an rQTL (APOE) and the current proposal proposes to identify other rQTL among lipids and CHD and subsequently loci that interact with them.

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.