ARIC Manuscript Proposal #2332

PC Reviewed: 3/11/14   Status: A   Priority: 2
SC Reviewed: _________   Status: _____   Priority: ____

1.a. Full Title:
Relationship loci (rQTL) and variance heterogeneity loci (vQTL) among blood pressure measures, incident CHD and loci that interact with them.

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

2. Writing Group:
Writing group members:
Taylor J. Maxwell, Peng Wei, Ying Cao, 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 (Maxwell et al., 2013) and demonstrate that there are rQTL (relationship loci) that significantly effect the correlations between systolic and diastolic blood pressure and BP 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 known risk factors (blood pressure) and CHD. They may also give us clues to genes and biological pathways that blood pressure measures to each other and to CHD. rQTL are sometimes related to variance heterogeneity loci (vQTL) where a quantitative trait has variance heterogeneity among the genotypes of the locus. These can also be the result of gene-by-gene interactions. Variance heterogeneity loci (vQTL) are also closely related to rQTL and context dependent interactions. Methodological work advanced by Dr. Maxwell (Cao et al., 2014) has produced an efficient and statistically powerful method to detect vQTL. vQTL can be the result of gene-by-gene interactions. They can also produce rQTL at the same locus between that trait and other (particularly disease). An rQTL for a pair of traits may be a vQTL for any linear combination of those traits, in particular, any vQTL for a composite trait (i.e. total cholesterol) may be an rQTL for the subcomponents of that trait (i.e. HDL & LDL).


5. Main Hypothesis/Study Questions:

Hypothesis 1: There are loci (rQTL) in the genome where the relationship between systolic (SBP) and diastolic blood (DBP) pressure 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 blood pressure is a risk factor for incident CHD, significant rQTL from hypothesis 1 are likely to modulate the relationship between the blood pressure traits 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 blood pressure traits differ by genotype.
Hypothesis 4: Some significant rQTL are also variance heterogeneity loci (vQTL) for one or both of quantitative traits in the model.

Hypothesis 5: There are variance heterogeneity loci (vQTL) for each blood pressure trait.

Hypothesis 6: Significant vQTL may be rQTL for that blood pressure trait and incident CHD.

Hypothesis 7: Significant rQTL and vQTL from hypotheses 1-3 interact with other coding loci to affect 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 with genotype data is available from the exome SNP chip. We will exclude individuals for which permissions were not granted for DNA use. For all analyses, typical covariates related to hypertension and CHD will be used such as age, sex, bmi, and hypertension medication.

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 (Ilkram et al, 2009).

For Hypothesis 1 we will perform a genome-wide screen between 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.

\[ \text{Trait}1_{ijk} = u + \text{age} + \text{sex} + \text{bmi} + \text{SNP}_{\text{Genotypic}} + \text{Trait}2 + \text{SNP}_{\text{Genotypic}} \times \text{Trait}2 + e_{ijk} \]

Where trait1 and trait2 represent two blood pressure measures in that particular scan. Age, sex, bmi, and hypertension medication will be used as covariates in the systolic and diastolic blood pressure model. 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 blood pressure measure 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 blood pressure measures from the original significant rQTL model.

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

Significance of the interaction term rejects the null hypothesis that the blood pressure measure 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/blood pressure measure 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 used to assess significance.

For Hypothesis 4 each quantitative trait from a significant rQTL will be tested for variance heterogeneity by genotype at the locus. We will use the omnibus method developed under Dr. Maxwell’s directions (Cao et al., 2014).

For Hypothesis 5 we will perform a genome-wide screen for each blood pressure trait using the omnibus test from Cao et al. 2014. Typical genome-wide significance thresholds will be used to assess significance of the omnibus tests. Subsequent variance heterogeneity tests will be performed for each globally significant omnibus test and a Bonferroni adjustment will be applied for that subset of tests.

For Hypothesis 6 any significant vQTL for a lipid trait may act as an rQTL for that blood pressure trait and incident CHD (i.e. it may act to modify the risk relationship between that lipid trait and incident CHD).

A significant rQTL (from the results of Hypotheses 1-6) creates a prior hypothesis (Hypothesis 7) that it interacts with other loci that affect one or both traits involved in the rQTL model (Pavlicev et al., 2011) or the individual trait for the vQTL. 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). A recent paper (Pavlicev et al. 2013) suggests that rQTL are often found in noncoding regions and that loci that interact with them are found in coding regions. For that reason, we would like to use the Exome SNP Chip genotype data for the gene-by-gene interaction screens. It would reduce the number of tests that need to be corrected for, it would increase the density of likely functional variants, and there is some evidence that these this specific type of variant interacts with rQTL. For each phenotype from a given significant rQTL model, we will perform a genome-wide scan for loci that interact with the rQTL 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.

\[
\text{Trait}_{ijk} = u + \text{age} + \text{sex} + \text{bmi} + rQTL_{\text{genotype}} + \text{SNP}_{\text{add}} + rQTL_{\text{genotype}} \cdot \text{SNP}_{\text{add}} + e_{ijk}
\]

\[
\text{CHD}_{ijk} = \text{age} + \text{sex} + \text{bmi} + rQTL_{\text{genotype}} + \text{SNP}_{\text{add}} + rQTL_{\text{genotype}} \cdot \text{SNP}_{\text{add}} + e_{ijk}
\]

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  ____ 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?  ____ Yes  ____ X  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.cscce.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 rQTL among BP traits and CHD and subsequently loci that interact with them. This paper has now been accepted by Genetics (2013 195: 1397-1405).  

I previously submitted and had approved #1985, “Relationship loci (rQTL) among lipid traits, blood pressure, incident CHD and loci that interact with them”. This proposal is the same but I would like this this proposal to supersede the other because I would prefer to use the exome SNP chip data instead of the GWAS data and I would like to split the rQTL analyses into 2 manuscripts by separating the blood pressure analyses from the lipid analyses.  

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.cscu.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.