ARIC Manuscript Proposal # 1774

PC Reviewed: 4/8/11	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: *The effects of interactions between psychosocial factors and genes on blood pressure traits.*

b. Abbreviated Title (Length 26 characters): *Blood pressure interaction.*

2. Writing Group:

Writing group members: *Sharon L.R. Kardia, (University of Michigan), Ana V. Diez-Roux (University of Michigan), Yan V. Sun (University of Michigan), Eric Boerwinkle (University of Texas Health Science Center), Kari North (Carolina Center for Genome Sciences, UNC-CH), Aravinda Chakravarti (Johns Hopkins University School of Medicine), Kathryn Rose, University of North Carolina, Gerardo Heiss, University of North Carolina, Kristin Tomey (University of Michigan)*

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

First author: *Sharon L.R. Kardia, Ph.D.*

Address: *University of Michigan, Department of Epidemiology*
1415 Washington Heights
Ann Arbor, Michigan 48109-2029

Phone: *734-936-0866* Fax: *734-764-1357*
E-mail: skardia@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, PhD*
Address: *Human Genetics Center and Div. of Epidemiology*
1200 Herman Pressler, Suite E-447
Houston, Texas 77030
Phone: *713.500.9816* Fax: *713.500.0900*
E-mail: Eric.Boerwinkle@uth.tmc.edu

3. Timeline:

Proposal submitted: April 7, 2011
Complete manuscript draft developed by: Oct 1, 2011
Submit to journal: Dec 15, 2011

4. Rationale:

A number of recent studies have identified replicable genetic predictors of blood pressure. In a joint meta-analysis with the Global BPgen Consortium, four CHARGE loci attained genome-wide significance for SBP (ATP2B1, CYP17A1, PLEKHA7, SH2B3), six for DBP (ATP2B1, CACNB2, CSK-ULK3, SH2B3, TBX3-TBX5, ULK4) and one for HTN (ATP2B1) (Levy, 2009). Several physiological pathways have been proposed for these genes in the pathogenesis of hypertension many of which are likely to be modulated by environmental factors. For example, the joint meta-analysis identified a genome-wide significant association of ATP2B1 with SBP, DBP and hypertension (17% increase in odds per risk allele and 37% increase for two risk alleles). This gene encodes PMCA1, a plasma membrane calcium/calmodulin-dependent ATPase that is expressed in vascular endothelium and is involved in calcium pumping from the cytosol to the extracellular compartment (Pande, 2006). This signal transduction pathway is influenced by oxidative stress and inflammation which are associated with lifestyle (physical activity, diet, smoking) and their underlying social or psychosocial motivations.

In parallel, a number of investigations have shown that social and psychosocial factors are also related to blood pressure levels. Several studies have linked low socioeconomic position at both the individual (Skodova, 2008) and neighborhood levels (Mujahid, 2010; Chichlowska, 2008) as well as over the life course (Diez-Roux, 2002) to high blood pressure. Possible pathways mediating these effects include effects of low socioeconomic position on health behaviors such as salt intake or lack of physical activity as well as possible links between socioeconomic position and stress, which has been hypothesized to be linked to blood pressure through neuroendocrine mechanisms (Player, 2007).

Psychosocial factors such as anger and anger reactivity have also been linked to high blood pressure, possibly through effects of these factors on sympathetic and HPA axis activity (Player, 2007; Smith, 2004; Manuck, 1986; Skodova, 2008). In ARIC, after adjusting for covariates, high levels of trait anger were associated with progression from prehypertension to hypertension (Player, 2007). Higher levels of vital exhaustion (which involves excessive fatigue and feelings of general malaise, such as hopelessness, listlessness, loss of libido, increased irritability and problems with sleep) have also been positively associated with blood pressure (Kopp, 1998). In addition, social support has been found to be inversely associated with blood pressure (Lett, 2005; Skodova, 2008) and it is thought that greater social support may buffer individuals against stressors that lead to higher blood pressure (Bell, 2010). Nevertheless, important questions remain regarding the relative importance of social and psychosocial factors in explaining variability in blood pressure.

It is often argued that proper understanding and quantification of the etiologic roles of genes and environments in the causation of diseases will require consideration of gene by environment interactions (Thomas, 2010). Specifically the impact of genetic predictors may be enhanced or only be present in the presence of a given environmental context. Likewise, the impact of environmental exposures may only be evident in the presence of genetic predispositions. Yet the investigation of gene by environment interactions remains in its infancy in part because of the lack of large population studies with rich environmental and genetic measurement. In addition,

the vast majority of gene by environment interaction studies have focused on either environmental toxins or health behaviors as the "environmental factors". Very few have examined the social or psychosocial environments in which genes operate. Prior literature suggests that these types of "environments" may be of special importance in the case of hypertension. Moreover, the pathways through which social and psychosocial factors are hypothesized to affect blood pressure (behaviors or stress) could plausibly interact with genetic predispositions.

We propose to use unique data available in the ARIC study to investigate interactions between social/psychosocial factors and genetic factors in blood pressure and pulse pressure in this large population sample.

5. Main Hypothesis/Study Questions:

Our general study question is: Do established genetic predictors of continuous blood pressure and pulse pressure interact with social and psychosocial factors? We will use the Global BPGEN (Levy 2009) and ICBP (Ehret, submitted) genes that have genome-wide significance and evidence of replication to test the following hypotheses:

Hypothesis 1: Low socioeconomic position (as assessed by a combined measure over the lifecourse and at the individual and neighborhood levels) will modify the effects of genetic predictors on continuous blood pressure and pulse pressure.

Hypothesis 2: Greater levels of anger/hostility and vital exhaustion and lower levels of social support will modify the effects of genetic predictors on continuous blood pressure and pulse pressure.

In secondary analyses we will investigate interactions with the separate measures of socioeconomic position that are combined into the single indicator evaluated in Hypothesis 1 to better understand the type of interactions being detected.

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).

Study design: The study will be cross-sectional since we will use the average value across the 4 exam periods for the blood pressure and pulse pressure outcomes and independent variables (when possible).

Inclusion/ exclusion criteria: Participants who have genetic data and outcome data during at least one time point will be included in the analysis.

Table 1. Variables of Interest		
Variable	**Source of Data**	**Exam Years Used for Proposal**
Outcome variables		
Systolic BP (mmHg)	ARIC blood pressure data: (*SBPA21, SBPB21 , SBPC22, SBPD19*)	Exams1, 2, 3, 4
Diastolic BP (mmHg)	ARIC blood pressure data: (*SBPA22, SBPB22, SBPC23, SBPD20*)	Exams1, 2, 3, 4
Pulse pressure	ARIC pulse pressure data (*SBPA09, SBPB09, SBPC10 , SBPD10)*	Exams1, 2, 3, 4
Use of antihypertensive medication (yes/ no)	Information on the use of antihypertensive medication from ARIC derived files	Exams1, 2, 3, 4
Medication class and type for antihypertensive medications used	Medication code and medication type	Exams1, 2, 3, 4
Exposure Variable		
Genetic markers as identified in large meta-analyses such as the International Consortium of Blood Pressure Genome-wide Association Studies—ICBP-GWAS	dbGap	Baseline
Psychosocial variables		
Anger/hostility	ARIC variables HPCB1-HPCB10 (from the Health and Life Profile)	Exam 2
Social support	ARIC variables HPAA01-HPAA29 (from the Health and Life Profile) all four visits	Exams 1, 2, 3, 4
VITAL EXHAUSTION	ARIC variables CORA07D, HPBA01- HPBA21 (from the Health and Life Profile)	Exam 3
Individual –level SES variables	ARIC derived datasets and Lifecourse SES Ancillary Study	Individual-level income, education and occupation for all available exams, childhood SES when available.
Neighborhood variables		
Neighborhood disadvantage (composite)	ARIC Lifecourse Ancillary Study	Summary neighborhood SES measures developed by the Lifecourse SES study for lifecourse and each ARIC visit

<u>Genetic data:</u> Our proposal will use existing genome-wide SNP data from the Affymetrix 6.0 platform (~900,000 genome-wide SNP markers; National Institutes of Health (R01 HL086694), A. Chakravarti, PI; Kardia, subcontract PI) and the imputation from the HapMap and 1000Genomes project for the genes identified as genome-wide significant and replicated in the Levy (2009) meta-analysis for blood pressure in Global BPGen and the Ehret (submitted) for ICBP. The genomic data to be used in the proposed study is already available in the Kardia laboratory (University of Michigan, School of Public Health, Ann Arbor, MI). Currently, we have

genotypic information on 9,716 non-Hispanic White, and 3,159 African American ARIC cohort participants.

Summary of Data Analysis:

Preliminary Analyses: Descriptive statistics will be calculated separately for each ethnic and sex specific subgroup, and the distributions will be tested for differences between genders within ethnic groups and between ethnic groups within genders using standard two-sample tests or chi-square tests of independence. Variable transformation will be performed as necessary to achieve adequate distributional properties. Distributions of values will be examined to identify potential outliers or other data issues that may affect modeling relationships. Initial analysis will evaluate bivariate associations between each predictor and outcome variable, associations between predictor variables, and pairwise interactions between predictors.

We will begin our preliminary genetic analysis by estimating a) allele frequencies, b) haplotype frequencies, and c) linkage disequilibrium using both D' and r-squared measures using a standard population genetic routine that we've implemented in R and S-Plus. Since population substructure is a potential source of confounding in population-based genetic association studies we will estimate principal components within ethnic group and incorporate them into the linear regression modeling as covariates.

We will take two approaches to handling the use of blood pressure medications. The first approach will be to simply add 10mm Hg to systolic blood pressure and 5 mm Hg to diastolic blood pressure to anyone on blood pressure medications to create a shift in their distributions. The second approach involves classifying medications into two broad classes - monotherapies versus combination therapies - and then classifying them by drug type (beta blocker, RAAS inhibitor, diuretics, calcium channel blockers, and research drugs.

For SES we will build on combined SES indices previously developed in ARIC by the Lifecourse SES study. These indices combine information over the lifecourse for both individual and contextual SES. WE WILL NOT NEED ACCESS TO GEOGRAPHIC IDENTIFIERS FOR THIS ANALYSIS. Primary analyses will focus on the summary index, because the summary index is likely to best capture the multiple ways in which SES may affect blood pressure and modify genetic effects. However in secondary analyses we will also investigate selected components separately.

For the psychosocial measures (anger/hostility, vital exhaustion, and social support) we will use previously developed scales. Primary analyses will focus on each scale separately because they each capture different constructs. However in secondary analyses, we will also explore a summary psychosocial score.

Covariates will include age, sex, race, height, weight, waist circumference, and body mass index.

***Assessing SNP-Risk Factor Interactions**: We will then estimate and compare the effects of social and psychosocial risk factors and SNPs on blood pressure and pulse pressure outcomes using ordinary least squares linear regression. Multilevel models (Diez-Roux, 2000) will also be employed as appropriate. SNP genotypes will be dummy coded to represent the additive and dominance deviations of each variation (i.e. for genotypes AA, Aa, and aa we will create two dummy variables X_1 and X_2 where X_1 = 1,0,-1 and X_2= 0,-1, 0). We will first estimate the effects of each individual level risk factor and SNP on average blood pressure or pulse pressure levels.*

Given the sometimes large number of SNPs implicated in the gene regions (named above) we will use two approaches to aggregate information. First, we will estimate the principal component the variation in a particular gene region to represent variation in the entire gene locus. Second, we will use a moving window approach (taking 2-4 SNPs) and performing haplotype-based analyses.

The lack of replication of genetic effects or gene-environment interactions on complex traits such as blood pressure is a major issue in the field of human genetic association studies. In order to guard against type I errors, we will use the False Discovery Rate to adjust for multiple testing (Benjamini, 1995; Storey, 2002). We will also use a four-fold cross validation strategy to estimate the percent variability predicted by the linear regression models estimated above (Stone, 1974; Browne, 2000) and repeat this procedure 10 times to estimate an average predictive accuracy for the models.

Modeling: *Linear regression models will be estimated using a cross-sectional approach, with dependent variables (systolic blood pressure, diastolic blood pressure or pulse pressure) constructed as average values from exams 1-4. Independent variables that are available at more than one time point will also be averaged. The primary analyses will have identified key interactions to provide a shortened list of variables to use in a forward selection modeling procedure to estimate the combined contributions of multiple genetic and psychosocial/social interactions on blood pressure. Briefly, the set of validated main effects will be allowed to enter the model first. Each time a variable is added we will test whether it provides significantly more explanatory power (using a partial F-test). After the best set of main effects has been selected, we will fix them into the model, and then allow interactions to be forwardly selected. Once all the interactions that are significant contributors to explaining the associations with average blood pressure or pulse pressure levels have been entered into the model, we will then make sure that the main effects for each term in the interaction are present in the model. If not, we will force the main effect into the multivariable model and re-estimate the joint parameter space. We will use the four-fold cross validation method to estimate the predictive ability of these models on independent (test) cases to provide a more accurate measure of their utility outside of the ARIC cohort.*

Anticipated Methodologic Limitations or Challenges: The genetic variations selected to be on the Affymetrix chip were not selected to be functional polymorphisms so we are likely to miss key associations or interactions because we do not have the functional variants to assess. Also, in a different vein, there are challenges in the interpretation of interactions. We will be cautious in our write-up of this study both because of the potentially sensitive nature of the topic but also because statistical interactions are only the very first step toward understanding why some individuals are at greater risk of high blood pressure than others.

References:

Bell CN, Thorpe RJ Jr, Laveist TA. Race/Ethnicity and hypertension: the role of social support. Am J Hypertens. 2010;23:534-40.

*Benjamini Y and Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J Royal Statistical Society **57**, 289-300 (1995).*

*Browne, M. W. Cross-Validation Methods. J. Math. Psychol. **44**, 108-132 (2000).*

Chichlowska KL, Rose KM, Diez-Roux AV, Golden SH, McNeill AM, Heiss G. Individual and neighborhood socioeconomic status characteristics and prevalence of metabolic syndrome: the Atherosclerosis Risk in Communities (ARIC) Study.

Psychosom Med. 2008 Nov;70(9):986-92. Epub 2008 Sep 16.
*Diez-Roux, A. V. Multilevel analysis in public health research. Annu. Rev. Public Health **21**, 171-192 (2000).*
Diez-Roux AV, Chambless LE, Merkin SS, Arnett D, Eigenbrodt M, Nieto FJ, Szklo M, Sorlie P. Socioeconomic disadvantage and change in blood pressure associated with aging. Circulation 2002;106:703-10.
Ehret, G.B. et al. Common Polymorphisms Impacting Blood Pressure and Cardiovascular Disease in Diverse Populations Highlight Novel Biological Pathways. Submitted.
Kopp MS, Falger PR, Appels A, Szedmak S. Depressive Symptomatology and Vital Exhaustion Are Differentially Related to Behavioral Risk Factors for Coronary Artery Disease. Psychosomatic Medicine 1998; 60:752-758.
Lett HS, Blumenthal JA, Babyak MA, Strauman TJ, Robins C, Sherwood A. Social support and coronary heart disease: epidemiologic evidence and implications for treatment. Psychosom Med. 2005;67:869-78.
Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, et al. Genome-wide association study of blood pressure and hypertension. Nat Genet. 2009 Jun;41(6):677-87. Epub 2009 May 10.
Manuck SB, Kaplan JR, Matthews KA. Behavioral antecedents of coronary heart disease and atherosclerosis. Arteriosclerosis. 1986;6:2-14.
Mujahid MS, Diez Roux AV, Cooper RC, Shea S, Williams DR. Neighborhood Stressors and Race/Ethnic Differences in Hypertension Prevalence (The Multi-Ethnic Study of Atherosclerosis). Am J Hypertens. 2010 Sep 16.
Pande J, Mallhi KK, Sawh A, Szewczyk MM, Simpson F, Grover AK. Aortic smooth muscle and endothelial plasma membrane CA2+ pump isoforms are inhibited differently by the extracellular inhibitor caloxin 1b1. Am. J. Physiol. 2006;290: 1341–1349.
Player MS, King DE, Mainous AG 3rd, Geesey ME. Psychosocial factors and progression from prehypertension to hypertension or coronary heart disease. Ann Fam Med. 2007;5:403-11.
Skodova Z, Nagyova I, van Dijk JP, Sudzinova A, Vargova H, Studencan M, Reijneveld SA. Socioeconomic differences in psychosocial factors contributing to coronary heart disease: a review. J Clin Psychol Med Settings. 2008 Sep;15:204-13.
Smith TW, Glazer K, Ruiz JM, Gallo LC. Hostility, anger, aggressiveness, and coronary heart disease: an interpersonal perspective on personality, emotion, and health. J Pers. 2004;72:1217-70.
*Stone, M. Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society **Series B, 36**, 111-147 (1974).*
*Storey, J. D. A direct approach to false discovery rates. Journal of the Royal Statistical Society **Series B, 64**, 479-498 (2002).*
Thomas D. Gene-environment-wide association studies: emerging approaches. Nature. 2010:11:259-272.

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? ***N/A***
____ 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)?

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

Yes, , *(1) GWA for loci influencing incident CHD and other HLB phenotypes (NHLBI RFA for large scale genotyping) (STAMPEED) (GEI) (this is for ARIC whites); (2) The National Heart Lung and Blood Institute's Candidate Gene Association Resource (CARE): Phase I (CARE) (this is for ARIC African Americans); and (3) Life course SES, social context, and CVD (SESCVD)*

11.b. If yes, is the proposal

x___ A. primarily the result of an ancillary study (list number* *(1) 2006.03 for ARIC Whites, (2) 2007.02 for ARIC AAs and (3) 1998.02 (1998.02)***)**

___ 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/

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.