ARIC Manuscript Proposal #2194

PC Reviewed: 8/13/13	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Exome chip and exome sequence analyses for hemostasis traits

b. Abbreviated Title (Length 26 characters): Exome & hemostatic factors

2. Writing Group:

Writing group members: ARIC writing group members include the participants in the CHARGE and ESP hemostasis working groups and individuals involved in generation of the sequence data (alphabetical): Eric Boerwinkle, Aaron Folsom, Alanna Morrison, Nathan Pankratz, Weihong Tang, Peng Wei

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

F**irst author**: Alanna Morrison, PhD
Address: University of Texas Health Science Center at Houston
School of Public Health
1200 Pressler St.; Suite 453E; Houston TX 77030

Phone: 713-500-9913 Fax: 713-500-0900
E-mail: Alanna.C.Morrison@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: Aaron Folsom, MD, MPH
Address: Division of Epidemiology, School of Public Health, University of Minnesota, 1300 South Second Street, Suite 300, Minneapolis, MN 55454

Phone: 612-626-8862 Fax: 612-624-0315
E-mail: folsom@umn.edu

3. Timeline:

Completion of analyses and draft of manuscript(s) by the end of 2013.

4. Rationale:

Fibrinogen, von Willebrand factor (vWF), and factor VIII are hemostatic factors that play a key role in coagulation and thrombosis as well as inflammation. Fibrinogen is a plasma glycoprotein synthesized by the liver. As a part of the coagulation cascade, thrombin converts fibrinogen into fibrin. Also, through activated GPIIbIIIa-mediated binding, fibrinogen plays a key role in platelet cross-linking. Fibrinogen is also a major acute phase reactant. vWF is a multimeric plasma glycoprotein produced in endothelial cells and stored in Weibel-Palade bodies that mediates platelet adhesion to the endothelial surface as well as the sub-endothelial matrix through GP Ib attachment. In the circulation, vWF acts as a carrier for coagulation factor VIII. Factor VIII is synthesized in both the liver and the sinusoidal endothelium, and plays a key role in coagulation as a critical cofactor (lack of factor VIII is hemophilia A) when the activated form (FVIIIa) binds to the enzyme factor IXa and factor X to produce the active enzyme responsible for thrombin production, factor Xa. vWF and factor VIII are modest acute phase reactants.

These hemostatic factors have been consistently associated with the development of cardiovascular disease (CVD).[1,2] Estimates of heritability range from 0.28 to 0.37 for fibrinogen[3], 0.40 to 0.61 for factor VIII, and 0.31 to 0.75 for vWF.[4,5] Characterization of common and low frequency variation influencing inter-individual and inter-population differences in fibrinogen, factor VII, factor VIII, and vWF may lead to improved understand of the role of hemostasis in inflammation and athero-thrombotic risk, as well as potentially reveal novel biologic pathways in which these hemostatic factors are involved.

Recent genome-wide association studies (GWAS) convincingly demonstrate that common polymorphisms in the fibrinogen structural genes on chromosome 4 (FGA, FGB, and FGG) influence fibrinogen levels in individuals of European ancestry.[6] A number of other loci (IL6R, IL1RN, NLRP3, CPS1, PCCB, SCL22A5-IRF1) have also been shown to play a role in fibrinogen in individuals of European and African ancestry.[6,7] Genome-wide studies for vWF have revealed loci (BAI38 and STXBP5, SCARA5, STAB2, STX2, TC2N and CLEC4M9) influencing vWF levels in addition to the well-characterized ABO blood group[9,10] and VWF gene.[9,11] Genome-wide significant results for factor VIII overlapped with the findings for vWF.[9] However, for each of these hemostatic factors, the common polymorphisms identified to-date explain only a small proportion of the heritability.[7,9] For each phenotype, the amount of variation explained by the genome-wide significant SNPs was the difference in R^2 from a model including the SNPs compared to a model containing only non-genetic covariates. It was found that these variants account for 12.8% of the vWF antigen variation, 10.0% of the factor VIII activity variation and the majority of genome-wide significant loci accounted for less than 2.0% of the variance in plasma fibrinogen.[6,9]

5. Main Hypothesis/Study Questions:

The analytical plan outlined here encompasses exome sequence data generated for ARIC Whites and Blacks through the NHLBI Exome Sequencing Project (ESP) and the CHARGE consortium. We will additionally consider exome chip variants genotyped for ARIC Whites and Blacks. Analysis of exome data (either exome sequence or exome

chip) will allow for: (1) identification of novel genes with common/rare variants that contribute to fibrinogen, factor VII, factor VIII, or vWF; (2) identification of rare variation in known candidate genes (e.g., those identified by GWAS) influencing these four hemostatic factors; and (3) assessment of the extent to which rare variants account for the heritability of fibrinogen, factor VII, factor VIII, and vWF.

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).

An analytic pipeline has been established to evaluate exome sequence or exome chip genotypes. A full description of the components of this pipeline is available on the CHARGE wiki (http://depts.washington.edu/chargeco/wiki/SKATmeta). Analyses are conducted separately in ARIC Whites and Blacks. The skatCohort portion of the analytic pipeline produces Rdata objects for the ARIC study that may then be meta-analyzed with other cohorts, or across projects (e.g., meta-analysis of CHARGE and ESP exomes).

Analysis of hemostasis traits will follow two main approaches:

1) Single variant association analyses

Single markers of a suitable frequency (e.g., >1% minor allele frequency, depending on the total sample size) will be analyzed for their association with the phenotype in regression models. The phenotypes and covariates are described below.

2) Gene-based analyses

For variants of low frequency (e.g., <1% minor allele frequency), we will evaluate the rare variants in aggregate within a gene using SKAT and T1 tests. Rare variants may be further subset to only include those of possible functional consequence (e.g., nonsynonymous, splicing, stopgain, or stoploss).

Meta-analysis

Meta-analysis will be conducted across the participating studies for single variants using methodologies developed through the CHARGE consortium (i.e., skatSingleSNP). Meta-analysis of gene-based tests will also be conducted using methodologies developed through the CHARGE consortium (i.e., skatMeta).

Phenotypes: fibrinogen, factor VII, factor VIII, and vWF at the first visit. Fibrinogen will be transformed using the natural log.
Covariates: age at visit 1, sex, and study center or principal components if suitable

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 ICTDER 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)?

There are no manuscript proposals for the evaluation of exome sequence data and hemostatic factors. This proposal is similar to other manuscript proposals to evaluate exome sequence data and other traits (e.g., EKG phenotypes, atrial fibrillation, pulmonary measures, blood pressure).

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data?

___ Yes _X_ No

11.b. If yes, is the proposal

___ A. primarily the result of an ancillary study (list number* _____)

___ 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.

13. References

1. Kaptoge S, et al. Associations of plasma fibrinogen levels with established cardiovascular disease risk factors, inflammatory markers, and other characteristics: individual participant meta-analysis of 154,211 adults in 31 prospective studies: The Fibrinogen Studies Collaboration. *American Journal of Epidemiology*. 2007;166:867-879.
2. Danesh J, et al. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *Journal of the American Medical Association*. 2005;294:1799-1809.
3. Freeman M, et al. Genetic contribution to circulating levels of hemostatic factors in healthy families with effects of known genetic polymorphisms on heritability. *Arteriosclerosis, Thrombosis and Vascular Biology*. 2002;22:506-510.
4. Souto J, et al. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation*. 2000;101.
5. deLange M, et al. The genetics of haemostasis: a twin study. *Lancet*. 2001;357:101-105.
6. Dehghan A, et al. Association of novel genetic loci with circulating fibrinogen levels: a genome-wide association study in 6 population-based cohorts. *Circulation: Cardiovascular Genetics*. 2009;2:125-133.
7. Wassel C, et al. Association of genomic loci from a cardiovascular gene SNP array with fibrinogen levels in European Americans and African-Americans from six cohort studies: the Candidate gene Association Resource (CARe). *Blood*. 2010;.
8. Antoni G, et al. A multi-stage multi-design strategy provides strong evidence that the BAI3 locus is associated with early-onset venous thromboembolism. *Journal of Thrombosis and Haemostasis*. 2010.
9. Smith N, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: the CHARGE Consortium. *Circulation*. 2010;121:1382-1392.
10. DeVisser M, et al. Linkage analysis of factor VIII and von Willebrand factor loci as quantitative trait loci. *Journal of Thrombosis and Haemostasis*. 2003;1:1771-1776.
11. Keightley A, et al. Variation at the von Willebrand (vWF) gene locus is associated with plasma vWF:Ag levels: identification of three novel single nucleotide polymorphisms in the vWF gene promoter. *Blood*. 1999;93:4277-4283.