ARIC Manuscript Proposal # 3130

PC Reviewed: 3/20/2018 Status: _____ Priority: 2
SC Reviewed: _________ Status: _____ Priority: _____

1.a. Full Title: *KNG1* rs710446, high molecular weight kininogen (HMWK), factor XI (FXI) and activated partial thromboplastin time (aPTT)

b. Abbreviated Title (Length 26 characters): *KNG1* gene, HMWK, FXI, and aPTT

2. Writing Group:
   Writing group members: Weihong Tang, Saonli Basu, James Pankow, Aaron Folsom, Mary Cushman, Nathan Pankratz, others are welcome

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

First author: Weihong Tang
Address: Division of Epidemiology and Community Health
         University of Minnesota

         Phone: 612-626-9140       Fax: 612-624-0315
         E-mail: tang0097@umn.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 R. Folsom
   Address: Division of Epidemiology and Community Health
            University of Minnesota

   Phone: 612-626-8862       Fax: 612-624-0315
   E-mail: folso001@umn.edu

3. Timeline: Fall 2018

4. Rationale:
Coagulation pathways are clearly involved in the etiology of VTE, with higher levels of plasma factor VIII, XI, von Willebrand factor, and D-dimer being predictive of VTE in the Longitudinal Investigation of Thromboembolism (LITE) study, which comprises ARIC and CHS. In addition, many of the common and low-frequency variants consistently associated with VTE risk in
GWAS (ABO, F11, F2, F5, FGG, GP6, KNG1, PROCR, SLC44A2, STXBP5, TSPAN15 and VWF) are involved in coagulation pathways.

HMWK is part of the contact pathway of coagulation, which also includes the proteins factor XI, factor XII, and prekallikrein, the precursor of HMWK. Activation of the contact system leads to procoagulant and proinflammatory reactions. The contact system is essential for surface-initiated coagulation, as exemplified by the aPTT, but whether the contact system has any role in initiating physiologic in vivo coagulation is disputed. Over the last few years, there has been renewed interest, especially because of experimental evidence suggesting that the contact system contributes to thrombosis. Knockout mice deficient in one of the contact proteins were protected against artificially induced thrombosis. Furthermore, inhibiting agents such as monoclonal antibodies, antisense oligonucleotides, and small molecules were found to prevent thrombosis in rodents and primates in both venous and arterial vascular beds. Factor XI is a strong VTE risk factor, and anti-factor XI prevented VTE in orthopedic patients. HMWK helps to optimally position prekallikrein and factor XI next to factor XII; it also inhibits the thrombin- and plasmin-induced aggregation of platelets. Although HMWK and prekallikrein have been associated with arterial CVD in limited studies, to our knowledge, only one retrospective study has examined the association of VTE with plasma HMWK, which was higher in VTE patients than in controls, whereas prekallikrein did not differ. However, those VTE patients were on vitamin K antagonists, which might have influenced results.

Defects in KNG1 (kininogen 1), the HMWK structural gene, can cause HMWK deficiency. Genome-wide association studies (GWAS) in LITE and other studies have shown rs710446, a missense variant in KNG1, to be associated with factor XI, aPTT, and VTE. However, adjustment for FXI did not abolish the association of rs710446 with the aPTT (PMID:28053049), suggesting that KNG1 influences the aPTT by additional mechanisms.

LITE is currently measuring HMWK in ARIC, using a case-cohort design (650 VTE events and a cohort random sample of 4,200). We therefore can test the hypotheses outlined below.

Note: Reference numbers are from the LITE renewal and are available upon request.

5. Main Hypothesis/Study Questions:

1. HMWK explains the associations of KNG1 rs710446 with factor XI (FXI) and aPTT.
2. Exploratory: GWAS will identify additional variants for HMWK and these newly identified variants may be associated with factor XI and aPTT.

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).

Exploration of the KNG1 SNP with HMWK and HMWK GWAS. ARIC measured genome-wide SNPs imputed to the 1000 Genomes reference panel. We will test the association of HMWK with dosage of genotyped or imputed SNPs, including rs710446, in linear regression
with adjustment for potential confounders including age, sex, and principal components for population stratification effects, after excluding outliers and anticoagulant users. To avoid bias caused by enrichment of disease alleles in VTE cases, the genetic analysis will focus on the random cohort sample with HMWK and be stratified by ethnicity (approximately 75% European Americans (EA)). We will use SNPTEST\textsuperscript{188} to perform the rs710446-HMWK analysis and the GWAS for HMWK. We will use METAL\textsuperscript{189} to perform a fixed effect inverse-variance weighting meta-analysis to pool results from EAs and African Americans (AA). Significance will be at p<0.05 for testing rs710446-HMWK and p = 2.5\times10^{-8} for the exploratory GWAS.\textsuperscript{190} To evaluate how adjustment for HMWK changes the association between rs710446, FXI, and aPTT, we will use mediation packages “Mediation”\textsuperscript{191} and “RMediation” implemented in R\textsuperscript{197} to measure the percent reduction in the effect size of rs710446 on FXI and aPTT when adjusted for HMWK.\textsuperscript{198,199} We will perform sensitivity analyses recommended for the “Mediation” package. Using the R package “powerMediation”, we estimate 80% power to detect a 4.5% change in the effect size of a continuous mediator with SD=1 and a correlation of 0.30 with the SNP at $\alpha=0.05$.

We are aware that aPTT was measured in Visit 1 samples and HMWK and FXI in Visit 3 samples. It is unknown how different sample visits might affect the mediation analysis for aPTT. In the literature, there are no data on correlation/variability of aPTT over time. One study conducted in ARIC showed that the correlation of fibrinogen and factor VIIc between measurements of 6 years apart was 0.51 for fibrinogen and 0.52 for factor VIIc (PMID: 10669661). Therefore, we might be able to assume that aPTT measured in samples of Visits 1 and 3 was also correlated. However, given this time issue, we will treat the mediation analysis that involves aPTT as an exploratory analysis.

For the exploratory HMWK GWAS, to be conservative, we estimated power assuming a cohort random sample of approximately 3,000 EAs. For the rs710446-HMWK analysis, at minor allele frequency (MAF) of 40% and $\alpha=0.05$, we have 86% power to detect genotypic differences of 0.08 SD. At minimum MAF of 9% and $\alpha=2.5\times10^{-8}$, we have 86% power to detect 0.3 SD genotypic differences. There are no genetic association study reports for HMWK, so we do not know the effect size of genetic variants with HMWK. In our GWAS for aPTT and FXI, the effect sizes of rs710446 (MAF=0.4-0.5) on aPTT and FXI were 0.35 SD\textsuperscript{54,86} and 0.38 SD (PMID:28053049), respectively. It is reasonable to assume that the effect size of \textit{KNG1} SNPs on HMWK is not less than that on FXI and aPTT because the \textit{KNG1} SNPs likely act through HMWK to influence FXI and aPTT.\textsuperscript{192} The RATIO study,\textsuperscript{92} comprising 638 healthy women aged 18-50, has agreed to replicate HMWK findings, and we will actively seek other replications. In RATIO, assuming five independent loci come forward to replication in a final sample size of 600, at the MAF of 40% and 9%, we will have at least 80% power to detect a minimal effect size of 0.19 SD and 0.32 SD, respectively. We also will compare our results for EAs and AAs (N about 1,000), as the rs710446-aPTT association replicated between racial groups.\textsuperscript{86}

\textbf{Phenotype:} HMWK, FXI, and aPTT  
\textbf{Covariates:} age, sex, field center, and principal components of ancestry for the genetic analysis.

7.a. Will the data be used for non-CVD analysis in this manuscript? \textbf{____ 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 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? __xx__ 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”? __xx__ 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.cscu.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)?

None

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

11.b. If yes, is the proposal
___xx__ A. primarily the result of an ancillary study (list number* ___2001.15 LITE____)
___ 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.cscu.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. Per Data Use Agreement Addendum for the Use of Linked ARIC CMS Data, approved manuscripts using linked ARIC CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes __X__ No.