Published in final edited form as:

J Thromb Haemost. 2009 October ; 7(10): 1743–1746. doi:10.1111/j.1538-7836.2009.03567.x.

Replication of findings on the association of genetic variation in 24 hemostasis genes and risk of incident venous thrombosis

Nicholas L. Smith, PhD, Kerri L. Wiggins, RD, MS, Alexander P. Reiner, MD, MS, Leslie A. Lange, PhD, Mary Cushman, MD, MSc, Susan R. Heckbert, MD, PhD, Thomas Lumley, PhD, Kenneth M. Rice, PhD, Aaron R. Folsom, MD, and Bruce M. Psaty, MD, PhD

Departments of Epidemiology (NLS, APR, SRH, BMP), Biostatistics (TL, KMR), Health Services (BMP), and Medicine (KLW, BMP), University of Washington; Seattle Epidemiologic Research and Information Center, VA Research and Development (NLS); Group Health Center for Health Studies (SRH, BMP), Seattle WA; Department of Genetics and the Carolina Center for Genome Sciences, (LAL), University of North Carolina, Chapel Hill, NC; Department of Pathology and Medicine (MC), University of Vermont, Burlington, VT; Divisions of Epidemiology and Community Health (ARF), University of Minnesota, Minneapolis, MN

A major obstacle to advancing knowledge of the genetic predictors of disease has been the failure to replicate findings when tested in new populations. In this letter, we present findings that attempted to replicate statistical significant associations from our previously published investigation of variation in 24 hemostasis genes and the risk of incident venous thrombosis (VT) in post-menopausal women.¹ In the original work, we identified 21 single nucleotide polymorphisms (SNPs) and 6 haplotypes across 24 hemostasis genes that were associated with a nominal p-value of less than 0.05. Among these, 5 SNPs remained significant after adjusting for multiple testing using a 0.2 threshold of the false-discovery rate q value.² Three of the 5 were “new” findings not previously reported. In this report, we used data from 2 studies to replicate the 3 findings and to further explore the other 24 associations in addition to the 2 well-characterized variants, factor V Leiden (rs6024) and prothrombin 20210A (rs1799963).

Using meta-analytic methods, we combined data from an on-going, population-based, case-control study in women 18–89 years of age and data from the Cardiovascular Health Study (CHS), a prospective cohort study of men and women aged 65 years and older. None of the subjects in this meta-analysis was included in the original report. The case-control data included only newly recruited subjects to the on-going study that produced the original report. Detailed methods were previously published.¹ Briefly, new cases included 91 incident VT events in premenopausal women that occurred between years 1995 and 2004 and 120 incident VT events in postmenopausal women that occurred in years 2003 and 2004. Also included were 187 and 435 matched controls, respectively. The median age at the time of the VT event was 54 years (interquartile range 41–72 years), 100% were women, and 89% were white. Genotyping was performed on genomic DNA using an Illumina GoldenGate custom panel. The cohort data were incident VT events in CHS identified by the Longitudinal Investigation of Thromboembolism Etiology (LITE). Methods for CHS and LITE were previously published.^{3, 4}

Corresponding authors: Nicholas L. Smith, PhD, Cardiovascular Health Research Unit, University of Washington, 1730 Minor Ave, Suite 1360, Seattle WA 98101, Phone: 206-287-2777; nlsmith@u.washington.edu.

There are no potential conflicts of interest for any author.

Briefly, CHS recruited 5,888 participants in 2 waves: 1989–90 (original cohort) and 1992–93 (minority cohort). Among CHS participants without a history of VT at baseline and for whom we had DNA and consent to use it for research, we identified 150 incident VT events through 2002, the last year of VT event identification in CHS-LITE. The median age at the time of the VT event was 79 years (interquartile range 74–84 years), 55% were women, and 79% were white. Genotyping was performed on genomic DNA using an Illumina GoldenGate custom panel.

Each study independently conducted association analyses of the 23 SNPs (including factor V Leiden and prothrombin 20210A variants) and 6 haplotypes with the risk of incident VT. Logistic regression was used to estimate relative risk (RR) of VT in the case-control study and Cox proportional hazards regression was used to estimate risk of VT in the cohort study. For both analyses, age- and race (white versus non-white) adjusted additive genetic models were used for SNP and haplotype associations. The case-control study also adjusted for case-control matching variables, which included index year, menopausal status, and hypertension status. The cohort study also adjusted for sex and clinic site as well as sampling variables, which included hypertension status and duration of follow-up. We used a fixed-effects, inverse-variance-weighted meta-analysis that combined risk-estimate coefficients and standard errors across the studies to produce a summary RR with alpha-based confidence intervals and p-value for each test. For the 3 SNPs that were “new” findings in the original publication and declared significant after adjustment for multiple testing, the 1-sided p-value threshold for replication was 0.0167 (0.05/3 with Bonferroni correction). For the 15 SNPs and haplotypes that were “new” associations in the original publication but not declared significant after adjusting for multiple testing, the 2-sided p-value threshold for significance was 0.0033 (0.05/15 with Bonferroni correction). For the remaining 11 SNPs for which there were previous reports in the literature of an association at the time of the original study, we attempted confirmation in the meta-analyzed data and used a 1-sided p-value threshold of 0.05. Furthermore, we tested all comparisons across the original and meta-analysis data to see if risk estimates were not statistically different. Statistical power was calculated based on the minor allele frequency (MAF) and relative risk (RR) from the original study.

The results of the meta-analysis that included 361 incident VT events are listed in the Table along with the original findings. For each SNP and haplotype, the MAF and RR with 95% confidence interval and p-value are provided. Also listed are the status of the finding at the time of the original publication (Original Status) and the current status (Updated Status) based on the results of the meta-analysis.

Replication of Significant “New” Findings from Original Report

Of the 3 “new” associations that were significant in the original study, only the factor V 45888 variant (rs4524), a non-synonymous substitution of lysine (K) for arginine (R) at amino acid position 858, was significantly associated with risk in the meta-analysis: 1-sided p-value 0.002. Contributing risk estimates in the case-control and cohort studies were similar: odds ratio = 0.66 and hazard ratio 0.84, respectively (see Supplemental Table). Linkage disequilibrium (LD) between factor V Leiden and 45888 was negligible ($r^2=0.02$) and the risk estimate for 45888 changed minimally when analyses were restricted to those without the Leiden variant.

The other 2 “new” associations that were significant in the original study did not reach statistical significance in the meta-analysis: all 1-sided p-values were greater than 0.27. The factor XI 22771 (rs2289252) variant was not associated with an increased risk of VT (RR = 1.02; 1-sided p-value 0.43). The p-value for the test of a difference in risk estimates was

0.05. Statistical power to detect an association of the magnitude originally reported for this SNP was 66%. Although this meta-analysis failed to replicate the association for the SNP, it was replicated recently by others in a report that used data from 2 case-control studies in the Netherlands.⁵

Bezemer and colleagues reported an odds ratio of 1.27 (95% CI: 1.16 – 1.38) for the rs3756008 SNP in factor XI, which is in high LD ($r^2=0.88$) with the factor XI 22771 (rs2289252) SNP. The other significant “new” finding in the original report was in the protein C gene (11310; rs5937) and was not significantly associated with VTE in the meta-analysis: RR = 1.06; 1-sided p-value 0.27). The p-value testing no difference in risk estimates was 0.05. Statistical power to detect an association of the magnitude originally reported for this SNP was 75%.

Exploration of Other “New” Associations from Original Report

Of the 15 SNPs and haplotypes that were “new” associations in the original publication but not declared significant after adjusting for multiple testing, none reached statistical significance in the meta-analysis: 2-sided p-values were 0.14 or greater or the risk estimate was in the opposite direction. These non-significant findings are consistent with our original findings.

Confirmation of Previously Replicated Associations from the Original Report

There were 11 associations in the original report that were previously reported in the literature. The meta-analysis confirmed 6 of the 11. Confirmations included the prothrombin 20210A variant, the factor V Leiden variant, and the 4 SNPs in the 3 fibrinogen genes.⁶ The meta-analysis did not confirm associations between the 3 SNPs in the protein C genes, the SNP in the thrombomodulin gene, and the SNP in the PAI-1 gene. For all 5 failed confirmations, the p-values testing for no difference in risk estimates were 0.09 or greater and statistical power was limited (power ranged from 38–72%) and may have contributed to null findings.

We meta-analyzed data from a case-control study and a cohort study to replicate, to explore, and to confirm associations identified in a previously reported case-control study. Although the number of incident VT events was nearly the same in the original and meta-analytic studies, statistical power in the meta-analysis to detect risk estimates of the magnitude originally reported was less than 80%. Original risk estimates are likely to be overestimates of any true effect size, which hampered power. Other reasons for failed replication or confirmations may have been related to heterogeneity attributable to non-genetic factors or to original findings that were false positives.

In summary, this meta-analysis replicated the association of factor V 45888 (rs4524) with the risk of incident VT. The factor XI 22771 (rs2289252) variant was not associated with a statistically significant increase in risk in this meta-analysis but was recently replicated in 2 large case-control studies. Additionally, the meta-analysis confirmed associations between genetic variation in the 3 fibrinogen genes in addition to the well-characterized factor V Leiden and prothrombin 20210A variants. Replication and confirmation efforts that are sufficiently powered are needed to fully describe associations between variation in hemostasis genes and risk of incident VT.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The case-control study reported in this article was supported by the National Health Lung and Blood Institute grants HL73410, HL60739, HL68639, HL43201, HL74745, HL68986, and HL95080; by a grant from the Leducq Foundation, Paris, France for the development of Transatlantic Networks of Excellence in Cardiovascular Research. The CHS/LITE/ARIC studies were funded by National Heart, Lung, and Blood Institute (NHLBI) grant R01 HL59367 (LITE), contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022 (ARIC), and contracts N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, NHLBI grant U01 HL080295, with additional contribution from the National Institute of Neurological Disorders and Stroke (CHS). A full list of participating CHS investigators and institutions can be found at <http://www.chs-nhlbi.org>. The authors thank the staff and participants of the ARIC and CHS studies for their important contributions.

The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. Dr NL Smith had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Smith NL, Hindorff LA, Heckbert SR, Lemaitre RN, Marcianti KD, Rice K, Lumley T, Bis JC, Wiggins KL, Rosendaal FR, Psaty BM. Association of genetic variations with nonfatal venous thrombosis in postmenopausal women. *JAMA*. 2007; 297:489–498. [PubMed: 17284699]
2. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B*. 1995; 57:289–300.
3. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A, O'Leary DH, Psaty BM, Rautaharju P, Tracy RP, Weiler PG. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol*. 1991; 1:263–276. [PubMed: 1669507]
4. Cushman M, Tsai AW, White RH, Heckbert SR, Rosamond WD, Enright P, Folsom AR. Deep vein thrombosis and pulmonary embolism in two cohorts: the longitudinal investigation of thromboembolism etiology. *Am J Med*. 2004; 117:19–25. [PubMed: 15210384]
5. Bezemer ID, Bare LA, Doggen CJ, Arellano AR, Tong C, Rowland CM, Catanese J, Young BA, Reitsma PH, Devlin JJ, Rosendaal FR. Gene variants associated with deep vein thrombosis. *Jama*. 2008; 299:1306–1314. [PubMed: 18349091]
6. Uitte de Willige S, de Visser MC, Houwing-Duistermaat JJ, Rosendaal FR, Vos HL, Bertina RM. Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen gamma' levels. *Blood*. 2005; 106:4176–4183. [PubMed: 16144795]

Table 1

Replication of associations between selected variants and haplotypes across hemostasis genes and the risk of incident venous thrombosis.

Variant identifier (location)* and RS number		Original Study (n=349 incident VT events)				Meta-Analysis (n=361 incident VT events)				P-value risks do not differ	
		Variant	MAF ^J (%)	OR (95%CI)	P-value	Original Status	MAF ^J (%)	RR (α-based CI)	P-value		Updated Status
Replication of New Findings											
FV	45888/rs4524	K858R	27.0	0.74 (0.60–0.90)	0.003	New ²	27.6	0.73 (0.58–0.92)	0.002	Rpl	0.94
FXI	22771(I)/rs2289252	C>T	39.5	1.31 (1.10–1.55)	0.002	New ²	40.5	1.02 (0.84–1.23)	0.43	F	0.05
PROC	11310(E)/rs5937	T>C	31.5	1.37 (1.15–1.64)	<0.001	New ²	33.1	1.06 (0.87–1.30)	0.27	F	0.05
Exploration of Other Associations											
FII	3696(I)/rs3136520	C>T	2.3	1.82 (1.16–2.86)	0.009	New	3.5	0.69 (0.17–2.77)	0.43	NA	0.06
FV	3578(I)/rs3753305	C>G	42.5	1.20 (1.01–1.42)	0.04	New	41.3	1.08 (0.74–1.57)	0.55	NA	0.50
FV	Haplotype 2 (low)	-	18.3	0.75 (0.60–0.95)	0.01	New	20.2	0.93 (0.67–1.30)	0.53	NA	0.19
FV	46058/rs9332695	T915S	1.3	2.01 (1.12–3.63)	0.02	New	1.6 ³	0.22 (0.01–4.97) ³	0.15 ³	NA	0.04
FIX	12806(I)/rs4149755	T>A	5.6	1.49 (1.08–2.05)	0.02	New	6.6	1.11 (0.70–1.75)	0.51	NA	0.19
FX	8946(I)/rs693335	C>G	40.4	0.80 (0.68–0.96)	0.01	New	41.6	0.98 (0.79–1.33)	0.83	NA	0.14
FXI	3543(I)/rs3822057	C>A	48.6	0.83 (0.70–0.99)	0.04	New	49.0	1.02 (0.86–1.22)	0.81	NA	0.10
TF	Haplotype other	-	0.4	2.55 (1.17–5.57)	0.03	New	0.8	1.00 (0.29–3.47)	1.00	NA	0.11
PROS	Haplotype other	-	2.0	1.85 (1.12–3.05)	0.008	New	2.0	0.82 (0.55–1.22)	0.14	NA	0.005
TFPI	1502(P)/rs2192824	G>A	38.8	1.25 (1.05–1.48)	0.01	New	39.9	0.81 (0.62–1.07)	0.03	NA	<0.001
TFPI	Haplotype 3	-	23.5	0.73 (0.59–0.90)	0.003	New	20.5	1.19 (0.83–1.70)	0.15	NA	0.003
TFPI	Haplotype other	-	0.3	0.40 (0.18–0.91)	0.01	New	0.2	0.71 (0.19–2.61)	0.44	NA	0.35
PAI-1	5878(I)/rs2227667	A>G	20.8	1.26 (1.03–1.55)	0.02	New	22.2	1.13 (0.83–1.54)	0.24	NA	0.44
PAI-1	Haplotype 3	-	11.5	1.38 (1.08–1.77)	0.007	New	11.5	1.27 (0.80–1.76)	0.19	NA	0.41
TAFI	36326(I)/rs17844078	T>C	4.4	0.52 (0.30–0.88)	0.02	New	3.7	1.24 (0.59–2.62)	0.39	NA	0.02
Confirmation of Hypothesized Associations											
FII	rs1799963	G>A	1.5	1.86 (1.05–3.26)	0.03	Rpl	1.3	3.34 (2.09–5.34)	<0.001	Conf	0.14
FV	38592/rs6025	R534Q	2.3	3.75 (2.56–5.15)	<0.001	Rpl	2.2	3.07 (2.07–4.57)	<0.001	Conf	0.52
FGA	251(P)/rs2070006	G>A	38.3	1.25 (1.05–1.29)	0.01	Rpl	38.8	1.27 (1.10–1.47)	0.003	Conf	0.89
FGA	6534/rs6050	T331A	26.0	1.24 (1.03–1.50)	0.02	Rpl	25.1 ⁴	1.66 (1.35–2.04) ⁴	<0.001 ⁴	Conf	0.06
FGB	1643(P)/rs1800788	C>T	21.3	1.31 (1.08–1.60)	0.008	Rpl	18.8	1.30 (1.09–1.55)	0.008	Conf	0.94

Variant identifier (location)* and RS number		Original Study (n=349 incident VT events)				Meta-Analysis (n=361 incident VT events)				P-value risks do not differ
		Variant	MAF ¹ (%)	OR (95%CI)	P-value	Original Status	MAF ¹ (%)	RR (α-based CI)	P-value	Updated Status
Gene										
FGG	129(P)/rs2066854	A>T	24.8	1.23 (1.02–1.49)	0.03	Rpl	23.8	1.42 (1.21–1.68)	<0.001	Conf
PROC	2583(U)/rs1799810	A>T	42.0	1.33 (1.12–1.57)	0.001	Rpl ²	44.8	1.12 (0.97–1.30)	0.09	NC
PROC	4515(P)/rs2069910	C>T	44.7	0.82 (0.69–0.98)	0.03	Rpl	47.2	0.85 (0.69–1.05)	0.11	NC
PROC	4919(P)/rs2069915	G>A	40.7	0.78 (0.65–0.93)	0.005	Rpl ²	39.6	0.96 (0.83–1.12)	0.34	NC
THBD	5110(U)/rs1042580	A>G	36.0	1.24 (1.04–1.47)	0.02	Inc	37.0	1.02 (0.88–1.19)	0.42	NC
PAI-1	664(P)/rs2227631	A>G	42.7	1.19 (1.01–1.41)	0.04	Inc	45.9	1.03 (0.89–1.19)	0.36	NC

Abbreviations: CI, confidence interval; Conf, confirmation of previous replication; E, exon; F, failed replication; I, intron; Inc, inconsistent associations in literature; MAF, minor allele frequency; NA, no association; NC, no confirmation; OR, odds ratio; P, promoter; PAI-1, plasminogen activator inhibitor 1; Rpl, replication of other report(s) in the literature; RR, relative risk; RS No., reference single nucleotide polymorphism accession identifier; TAFI, thrombin activatable fibrinolysis inhibitor; TPPI, tissue factor pathway inhibitor; U, untranslated region.

¹ Among controls in the original study and equally weighted across the 2 studies in the meta-analysis using controls from the case-control study and all subjects from the cohort study.

² Declared significant after adjusting for multiple testing using the false discovery rate statistic.

³ Estimates based on data from the cohort study only since genotyping data from the case-control study were missing.

⁴ Estimates based on data from the case-control study only since genotyping data from the cohort study were missing.