e-Blood

THROMBOSIS AND HEMOSTASIS

Rare and low-frequency variants and their association with plasma levels of fibrinogen, FVII, FVIII, and vWF

Jennifer E. Huffman,[1-3] Paul S. de Vries,[4] Alanna C. Morrison,[5] Maria Sabater-Lleal,[6] Tim Kacprowski,[7] Paul L. Auer,[8] Jennifer A. Brody,[9] Daniel I. Chasman,[10,11] Ming-Huei Chen,[12] Xiuqing Guo,[13,14] Li-An Lin,[5,15] Riccardo E. Marioni,[16-18] Martina Müller-Nurasyid,[19-21] Lisa R. Yanek,[22] Nathan Pankratz,[23] Megan L. Grove,[5] Moniek P. M. de Maat,[24] Mary Cushman,[25] Kerri L. Wiggins,[9] Lihong Qi,[26] Bengt Sennblad,[6,27] Sarah E. Harris,[16,17] Ozren Polasek,[28] Helene Riess,[29,30] Fernando Rivadeneira,[4,31] Lynda M. Rose,[10] Anuj Goel,[32] Kent D. Taylor,[13,14] Alexander Teumer,[33] André G. Uitterlinden,[4,31] Dhananjay Vaidya,[22,34] Jie Yao,[13] Weihong Tang,[35] Daniel Levy,[3] Melanie Waldenberger,[21,29,36] Diane M. Becker,[22,37] Aaron R. Folsom,[35] Franco Giulianini,[10] Andreas Greinacher,[38] Albert Hofman,[4] Chiang-Ching Huang,[8] Charles Kooperberg,[39] Angela Silveira,[6] John M. Starr,[16,40] Konstantin Strauch,[19,41] Rona J. Strawbridge,[6] Alan F. Wright,[1] Barbara McKnight,[42] Oscar H. Franco,[4] Neil Zakai,[25] Rasika A. Mathias,[22,43] Bruce M. Psaty,[9,44-46] Paul M. Ridker,[10,11] Geoffrey H. Tofler,[47] Uwe Völker,[7] Hugh Watkins,[32] Myriam Fornage,[5,15] Anders Hamsten,[6] Ian J. Deary,[16,48] Eric Boerwinkle,[5,49] Wolfgang Koenig,[30,50] Jerome I. Rotter,[13,14] Caroline Hayward,[1] Abbas Dehghan,[4] Alex P. Reiner,[39,44] Christopher J. O'Donnell,[2] and Nicholas L. Smith[44,46,51]

[1]Medical Research Council Human Genetics Unit, Medical Research Council Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom; [2]The Framingham Heart Study, Cardiovascular Epidemiology and Human Genomics Branch, National Heart, Lung, and Blood Institute, Framingham, MA; [3]The Framingham Heart Study, Framingham, MA, and the Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD; [4]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; [5]Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX; [6]Cardiovascular Genetics and Genomics Group, Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institute, Stockholm, Sweden; [7]Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt University Greifswald, Greifswald, Germany; [8]Joseph J. Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, WI; [9]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA; [10]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA; [11]Harvard Medical School, Boston, MA; [12]Department of Neurology, Boston University School of Medicine, Boston, MA; [13]The Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute, and [14]Department of Pediatrics, Harbor-University of California at Los Angeles (UCLA) Medical Center, Torrance, CA; [15]Brown Foundation Institute of Molecular Medicine, The University of Texas Health Science Center at Houston, Houston, TX; [16]Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom; [17]Centre for Genomic and Experimental Medicine, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom; [18]Queensland Brain Institute, University of Queensland, Brisbane, QLD, Australia; [19]Institute of Genetic Epidemiology, Helmholtz Centre Munich-German Research Center for Environmental Health, Neuherberg, Germany; [20]Department of Medicine I, Ludwig-Maximilians-University Munich, Munich, Germany; [21]German Centre for Cardiovascular Research, partner site Munich Heart Alliance, Munich, Germany; [22]GeneSTAR Research Program, Division of General Internal Medicine, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; [23]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN; [24]Department of Hematology, Erasmus MC, Rotterdam, The Netherlands; [25]Departments of Medicine and Pathology, University of Vermont, Colchester, VT; [26]University of California at Davis, Davis, CA; [27]Science for Life Laboratory, Karolinska Institute, Stockholm, Sweden; [28]Department of Public Health, Faculty of Medicine, University of Split, Split, Croatia; [29]Institute of Epidemiology II, Helmholtz Centre Munich-German Research Center for Environmental Health, Neuherberg, Germany; [30]Department of Internal Medicine II-Cardiology, University of Ulm Medical School, Ulm, Germany; [31]Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands; [32]Radcliffe Department of Medicine, Division of Cardiovascular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, United Kingdom; [33]Study of Health in Pomerania/ Clinical and Epidemiological Research Department, Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany; [34]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; [35]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN; [36]Research Unit of Molecular Epidemiology, Helmholtz Centre Munich-German Research Center for Environmental Health, Neuherberg, Germany; [37]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; [38]Institute of Immunology and Transfusion Medicine, Department of Transfusion Medicine, University Medicine Greifswald, Greifswald, Germany; [39]Fred Hutchinson Cancer Research Center, Seattle, WA; [40]Alzheimer Scotland Dementia Research Centre, Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom; [41]Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University, Munich, Germany; [42]Department of Biostatistics, University of Washington, Seattle, WA; [43]Division of Allergy and Clinical Immunology, Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD; [44]Department of Epidemiology, and [45]Department of Health Services, University of Washington, Seattle, WA; [46]Group Health Research Institute, Group Health Cooperative, Seattle, WA; [47]Royal North Shore Hospital, University of Sydney, Sydney, NSW, Australia; [48]Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom; [49]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; [50]German Heart Centre Munich, Munich Technical University, Munich, Germany; and [51]Seattle Epidemiologic Research and Information Center, Veterans Affairs Office of Research and Development, Seattle, WA

Submitted February 11, 2015; accepted May 27, 2015. Prepublished online as *Blood* First Edition paper, June 23, 2015; DOI 10.1182/blood-2015-02-624551.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

This article contains a data supplement.

Key Points

- Twelve independent, novel, low-frequency (n = 2) and rare (n = 10) genetic variants were associated with fibrinogen, FVII, FVIII, or vWF.
- Nine were within previously associated genes, and 3 novel candidate genes (*KCNT1*, *HID1*, and *KATNB1*) were confined to cohorts of African ancestry.

Fibrinogen, coagulation factor VII (FVII), and factor VIII (FVIII) and its carrier von Willebrand factor (vWF) play key roles in hemostasis. Previously identified common variants explain only a small fraction of the trait heritabilities, and additional variations may be explained by associations with rarer variants with larger effects. The aim of this study was to identify low-frequency (minor allele frequency [MAF] ≥0.01 and <0.05) and rare (MAF <0.01) variants that influence plasma concentrations of these 4 hemostatic factors by meta-analyzing exome chip data from up to 76 000 participants of 4 ancestries. We identified 12 novel associations of low-frequency (n = 2) and rare (n = 10) variants across the fibrinogen, FVII, FVIII, and vWF traits that were independent of previously identified associations. Novel loci were found within previously reported genes and had effect sizes much larger than and independent of previously identified common variants. In addition, associations at *KCNT1*, *HID1*, and *KATNB1* identified new candidate genes related to hemostasis for follow-up replication and functional genomic analysis. Newly identified low-frequency and rare-variant associations accounted for modest amounts of trait variance and therefore are unlikely to increase predicted trait heritability but provide new information for understanding individual variation in hemostasis pathways. (*Blood*. 2015;126(11):e19-e29)

Introduction

Fibrinogen, coagulation factor VII (FVII) and factor VIII (FVIII) and its carrier protein von Willebrand factor (vWF) play key roles in hemostasis. Plasma levels of these hemostatic factors are associated with risk of arterial and venous thrombosis, and fibrinogen is also a marker of inflammation.[1-6] Previous genome-wide association studies (GWASs) mainly interrogated common genetic variation and identified variants of modest effect across these phenotypes,[4,7-14] with the largest studies identifying 23 loci for fibrinogen,[9] 5 each for FVII[13] and FVIII,[13] and 8 for vWF.[13] Nonetheless, the associated variants still explain little about the trait heritabilities.[9,12,15] An additional proportion of the missing heritability may be attributed to association with rare variants, which are not captured by the conventional genome-wide marker arrays or imputation panels that have been used for GWASs.[15] In addition, investigating rare genetic variation is important to understanding individual variation in the biology underlying hemostasis pathways.

The aim of this study was to identify low-frequency and rare variants, analyzed individually or at the level of the gene, that influence plasma concentrations of fibrinogen, FVII, FVIII, and vWF. To this end, we meta-analyzed phenotype-genotype associations of low-frequency (minor allele frequency [MAF], 0.01-0.05) and rare (MAF < 0.01) exonic variants in 76 000 individuals of European (EUR), African (AFR), Hispanic (HIS), or East Asian (ASI) ancestry from 16 studies within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium.[16] We restricted our analyses to variants that were predicted to alter the coding sequence of the gene product to enhance the likelihood of identifying causal variants and to reduce the burden of multiple testing.

Methods

Setting and participating cohorts

This study was organized within the CHARGE Consortium Hemostasis Working Group and included 16 cohorts of EUR, AFR, ASI, or HIS ancestry. Descriptions and ancestry composition of participating cohorts are found in the supplemental Data available on the *Blood* Web site).

Hemostatic factors

Hemostasis phenotypes included plasma measures of fibrinogen, FVII, FVIII, and vWF. Fibrinogen (g/L) was available in all 16 studies; FVII activity (% or IU/mL × 100) and FVII antigen (% or IU/mL × 100) were available in 7 studies; FVIII activity (% or IU/mL × 100) was available in 5 studies; and vWF antigen (% or IU/mL × 100) was available in 8 studies. Methods used by each study are noted in Table 1.

Genotype calling and quality control

Fourteen studies were genotyped by using the HumanExome BeadChip v1.0 (Illumina, Inc., San Diego, CA) whereas one was genotyped by using BeadChip v1.1 and another by using BeadChip v1.2. Variant calling and quality control procedures are described in the supplemental Data and in previously published articles.[17,18] Prior to analysis, individual studies recoded variants to additive coding by using the minor allele derived from the CHARGE joint calling.

Statistical analysis

In each study fibrinogen measures were natural-log (ln) transformed. For untransformed FVII, FVIII, or vWF, participants with values 3 standard deviations above or below the population mean were removed prior to cohort-level analysis. Study-specific regression analyses were adjusted for sex, age, study design variables, and population substructure by using principal components. MAF thresholds were defined by using the ancestry-specific allele frequencies derived from the CHARGE joint calling.[17] Variant annotation was performed centrally within CHARGE by using dbNSFP v2.0.[19,20] All association analyses were performed by using the R package seqMeta (http://cran.r-project.org/web/packages/seqMeta/index.html). Details of the genotyping chip and version of statistical software used by each study are provided in supplemental Table 1.

Main association testing. *Single-variant tests.* We investigated low-frequency and rare variants individually by using standard single-variant association analyses. From among the functional variants on the array (defined as missense, stop-gain, stop-loss, or splice-site changes), we selected variants with an MAF <5% and an expected minor allele count of ≥5 in the total meta-analysis sample for single-variant association of autosomal chromosomes. Because commonly occurring variations on the X chromosome have not previously been investigated for some of the phenotypes, no upper MAF threshold was used when testing for associated variants on this chromosome. The Y and mitochondrial chromosomes were not interrogated. Bonferroni-corrected *P* value thresholds of statistical significance were based on the number of single-variant tests performed, and they varied by ancestry: 2.5×10^{-7} (ALL cohorts), 2.6×10^{-7} (EUR + AFR cohorts), 2.9×10^{-7} (EUR only), 3.3×10^{-7} (AFR only), 1.7×10^{-6} (ASI only), and 4.7×10^{-7} (HIS only) (see supplemental Data).

Gene-based tests. Analytical methods that aggregate the effect of multiple rare variants across a gene were used to test for association. This resulted in a *P* value for a gene rather than for a single variant. Both unidirectional and random effects tests were used; unidirectional texts are more powerful when rare variant effects within a region are in the same direction, and random effects tests are more powerful when rare variants affect a phenotype in opposite directions or when many variants have null effects.

Table 1. Study participant characteristics and phenotype assay or measure

Factor and study acronym	Ancestry	No. of participants in study	% Female	Mean age, y	Trait		Assay/measure
					Mean	SD	
Fibrinogen (g/L)							
ARIC[49]	EUR	10 757	53.1	54.3	2.90	1.21	Clauss
	AFR	3 643	61.9	53.5	3.13	1.23	
CARDIA[50]	EUR	2 041	52.5	30.5	2.51	1.23	Immunonephelometry
	AFR	1 709	56.9	29.4	2.66	1.23	
CHS[51]	EUR	4 034	56.2	72.8	3.13	1.22	Clauss
	AFR	757	62.2	72.7	3.35	1.23	
FHS[52,53]	EUR	6 711	54.3	46.0	3.24	0.68	Clauss
GeneSTAR[54]	EUR	1 091	51.2	41.2	3.51	0.98	Modified Clauss
	AFR	641	61.9	40.6	3.80	1.12	
KORA S4[55,56]	EUR	2 687	53.1	47.9	2.60	0.58	Immunonephelometry
Korcula[57]	EUR	748	64.3	56.4	4.55	1.52	Clauss
LBC 1921[58,59]	EUR	466	57.4	79.1	3.59	0.86	Clauss
LBC 1936[58,60]	EUR	973	49.2	69.6	3.27	0.63	Clauss
MESA[61]	EUR	2 483	52.1	62.7	3.35	0.7	Immunonephelometry on the BN II nephelometer
	AFR	1 638	53.8	62.2	3.60	0.79	
	ASI	764	50.8	62.4	3.29	0.61	
	HIS	1 431	51.5	61.0	3.59	0.75	
PROCARDIS[62]	EUR	1 404	36.8	60.9	4.06	0.96	Immunonephelometric
RS-I-1[63-65]	EUR	1 114	59.0	70.2	2.70	1.26	Prothrombin time
RS-I-3[63-65]	EUR	972	46.7	72.4	3.96	0.89	Prothrombin time
SCARF[66]	EUR	697	17.5	53.2	3.47	0.79	Immunonephelometric
SHIP[67]	EUR	5 940	52.3	47.9	2.99	0.71	Clauss
WGHS[7,68]	EUR	22 411	100	54.7	3.59	0.78	Mass-based immunoturbidimetric assay
WHI[69-71]	EUR	1 204	100	69.6	3.06	0.86	Clauss
FVII (% antigen or % activity)*							
ARIC	EUR	10 544	52.9	54.3	118.3	26.7	Clotting assay (% activity)
	AFR	3 574	61.9	53.6	116.7	28.4	
CARDIA	EUR	997	52.5	30.6	83.7	21.5	Clotting assay (% activity)
	AFR	637	55.6	29.2	84.2	26.2	
CHS	EUR	4 063	56.2	72.8	125.9	29.5	Clotting assay (% activity)
	AFR	760	62.1	72.6	113.0	26.4	
FHS	EUR	2 620	55.3	53.9	100.3	16.3	ELISA (% antigen)
RS-I	EUR	670	59.0	70.6	107.5	19.1	Clotting assay (% activity)
SCARF	EUR	698	17.5	53.2	139.9	35.8	ELISA (% antigen)
WHI	EUR	809	100	69.9	146.0	52.5	Clotting assay (% activity)
FVIII (% activity)							
ARIC	EUR	10 708	53.0	54.3	124.1	30.6	Clotting assay
	AFR	3 618	61.7	53.5	144.8	41.7	
CARDIA	EUR	998	52.6	30.6	89.8	31.7	Clotting assay
	AFR	632	55.6	29.2	103.5	38.7	
CHS	EUR	4 009	56.2	72.8	120.8	36.7	Clotting assay
	AFR	191	63.9	72.6	138.3	43.9	
MESA	EUR	2 483	52.1	62.7	156.9	64.6	Clotting assay
	AFR	1 638	53.8	62.2	178.0	74.6	
	ASI	764	7.7	62.4	157.9	57.2	
	HIS	1 418	51.5	61.0	161.8	63.4	
RS-I	EUR	1 832	52.0	68.6	115.7	46.1	Clotting assay
vWF (% antigen)							
ARIC	EUR	10 736	53.1	54.3	110.7	39.1	ELISA
	AFR	3 625	61.8	53.5	131.4	51.1	
CARDIA	EUR	1 002	52.6	30.6	89.9	36.4	ELISA
	AFR	636	55.7	29.2	94.3	44.4	
FHS	EUR	2 621	55.3	53.9	125.3	45.0	ELISA
GeneSTAR	EUR	991	52.5	42.6	78.7	46.1	ELISA
	AFR	582	62.2	42.6	76.8	42.5	
LBC 1921	EUR	150	57.3	86.6	149.7	45.9	ELISA
LBC 1936	EUR	706	47.9	72.5	122.6	37.8	ELISA
MESA	EUR	443	54.7	62.7	135.2	54.5	ELISA
	AFR	193	64.8	62.2	156.1	64.8	
RS-I	EUR	1 587	49.9	73.1	135.9	54.1	ELISA

Full cohort descriptions can be found in the supplemental Data.

ARIC, Atherosclerosis Risk in Communities Study; CARDIA, Coronary Artery Risk Development in Young Adults; CHS, Cardiovascular Health Study; ELISA, enzyme-linked immunosorbent assay; FHS, Framingham Heart Study; GeneSTAR, Genetic Study of Atherosclerosis Risk; KORA S4, Kooperative Gesundheitsforschung in der Region Augsburg; Korcula, Croatia-Korcula study; LBC 1921, Lothian Birth Cohort 1921; LBC 1936, Lothian Birth Cohort 1936; MESA, Multi-Ethnic Study of Atherosclerosis; PROCARDIS, Precocious Coronary Artery Disease Study; RS-I, Rotterdam Study-I; SCARF, Stockholm Coronary Artery Risk Factors; SD, standard deviation; SHIP, Study of Health in Pomerania; WGHS, Women's Genome Health Study; WHI, Women's Health Initiative.

Table 2. Single-variant meta-analysis results for hemostatic factors fibrinogen, FVII, FVIII, and vWF

Factor and variant	AA Change*†	Gene	Ancestry	No. of participants in study	MAF‡	β	*P*
Fibrinogen							
rs201909029 (new)	K178N (K148N)	*FGB*	ALL	76 316	7.7E-04	−0.139	**3.2E-13**
			EUR	65 733	8.8E-04	−0.139	**5.2E-13**
			AFR	8 388	6.0E-05	−0.163	4.3E-01
			ASI	764	0	NA	NA
			HIS	1 431	3.5E-04	−0.117	5.5E-01
rs6054	P265L (P235L)	*FGB*	ALL	76 316	4.2E-03	−0.111	**1.8E-43**
			EUR	65 733	4.7E-03	−0.111	**3.7E-42**
			AFR	8 388	1.2E-03	−0.104	2.6E-02
			ASI	764	1.3E-03	−0.130	3.0E-01
			HIS	1 431	0	NA	NA
rs145051028 (new)	S245F (S219F)	*FGG*	ALL	76 316	1.6E-04	−0.239	**4.8E-09**
			EUR	65 733	0	NA	NA
			AFR	8 388	1.5E-03	−0.239	**4.8E-09**
			ASI	764	0	NA	NA
			HIS	1 431	0	NA	NA
rs148685782	A108G (A82G)	*FGG*	ALL	76 316	3.3E-03	−0.238	**9.2E-152**
			EUR	65 733	3.8E-03	−0.239	**2.3E-150**
			AFR	8 388	4.2E-04	−0.165	3.4E-02
			ASI	764	0	NA	NA
			HIS	1 431	3.5E-04	−0.347	7.7E-02
rs10479001	A225V	*PDLIM4*	ALL	76 316	5.5E-02	0.013	**1.3E-08**
			EUR	65 733	4.5E-02	0.018	**4.3E-11**
			AFR	8 388	1.4E-01	−0.001	8.3E-01
			ASI	764	0	NA	NA
			HIS	1 431	5.5E-02	0.019	2.3E-01
rs1800961	T117I	*HNF4A*	ALL	76 316	2.7E-02	−0.020	**2.3E-10**
	T139I		EUR	65 733	3.0E-02	−0.020	**5.5E-10**
	T169I		AFR	8 388	5.9E-03	0.012	5.5E-01
			ASI	764	1.1E-02	−0.031	4.8E-01
			HIS	1 431	4.2E-02	−0.038	4.0E-02
rs151272083 (new)	R865Q	*KCNT1*	ALL	76 316	2.2E-03	0.007	5.3E-01
	R877Q		EUR	65 733	2.4E-03	0.017	1.3E-01
	R891Q		AFR	8 388	7.2E-04	−0.330	**2.7E-07**
	R910Q		ASI	764	0	NA	NA
			HIS	1 431	0	NA	NA
rs141869748 (new)	I193T	*HID1*	ALL	76 316	1.6E-04	−0.216	4.2E-07
	I421T		EUR	65 733	0	NA	NA
			AFR	8 388	1.3E-03	−0.252	**4.0E-08**
			ASI	764	0	NA	NA
			HIS	1 431	1.1E-03	0.008	9.4E-01
FVII							
rs150525536 (new)	R117Q	*F7*	ALL	25 372	9.5E-04	−31.44	**1.8E-17**
	R70Q		EUR	20 401	9.8E-05	−13.92	2.2E-01
	R139Q		AFR	4 971	4.4E-03	−33.56	**9.7E-18**
rs121964926 (new)	R342Q	*F7*	ALL	25 372	1.2E-03	−25.02	**1.3E-14**
	R295Q		EUR	20 401	4.2E-04	−0.52	9.3E-01
	R364Q		AFR	4 971	4.4E-03	−38.08	**2.8E-21**
rs3093248 (new)	E423K	*F7*	ALL	25 372	7.5E-04	−22.00	2.8E-07
	E376K		EUR	20 401	2.5E-05	−62.77	2.3E-02
	E445K		AFR	4 971	3.7E-03	−20.99	1.3E-06
FVIII							
rs7962217	G2705R	*VWF*	ALL	28 291	4.6E-02	5.16	**2.5E-13**
			EUR	20 030	5.5E-02	4.84	**4.0E-11**
			AFR	6 079	1.6E-02	8.58	7.9E-03
			ASI	764	7.2E-03	17.63	3.0E-01
			HIS	1 418	5.8E-02	10.21	2.7E-02

Only SNPs that were still significant after conditional analyses are included in the table. SNPs that achieved genome-wide significance threshold (ALL, $P = 2.50E-07$; EUR, $P = 2.88E-07$; AFR, $P = 3.30E-07$; ASI, $P = 1.70E-06$; and HIS, $P = 4.67E-07$) are shown in **bold**.

ALL, all ancestries (only EUR + AFR for FVII and vWF); NA, not applicable.

*AA change, amino acid change of SNP.

†Amino acid position in parentheses is for the mature protein for *FGB* (position 30) and *FGG* (position 26).

‡MAF, minor allele frequency from CHARGE joint calling.

Table 2. (continued)

Factor and variant	AA Change*†	Gene	Ancestry	No. of participants in study	MAF‡	β	*P*
rs41276738 (new)	R854Q	*VWF*	ALL	28 291	4.0E-03	−16.89	**2.2E-13**
			EUR	20 030	5.3E-03	−15.96	**9.2E-12**
			AFR	6 079	9.9E-04	−49.57	3.8E-04
			ASI	764	0	NA	NA
			HIS	1 418	1.1E-03	−19.47	5.5E-01
rs141041254 (new)	E2377K	*STAB2*	ALL	28 291	8.7E-04	26.81	**2.1E-08**
			EUR	20 030	1.2E-03	28.06	**7.6E-09**
			AFR	6 079	2.5E-04	−11.70	6.6E-01
			ASI	764	0	NA	NA
			HIS	1 418	0	NA	NA
rs1800291	D1260E	*F8*	ALL	28 291	2.7E-01	−1.73	**8.2E-08**
			EUR	20 030	1.7E-01	−2.15	**5.0E-09**
			AFR	6 079	3.5E-01	−0.54	4.5E-01
			ASI	764	4.7E-02	7.29	1.8E-01
			HIS	1 418	2.5E-01	0.28	8.9E-01
rs142508811 (new)	D413D	*KATNB1*	ALL	28 291	2.7E-04	39.36	4.8E-04
	D410D		EUR	20 030	1.8E-04	1.08	9.4E-01
	(predicted to alter splicing)		AFR	6 079	6.6E-04	86.35	**2.8E-07**
			ASI	764	0	NA	NA
			HIS	1 418	0	NA	NA
vWF							
rs141041254 (new)	E2377K	*STAB2*	ALL	23 272	8.2E-04	33.65	**2.4E-07**
			EUR	18 236	9.9E-04	35.21	**1.1E-07**
			AFR	5 036	2.0E-04	−11.56	7.5E-01

Only SNPs that were still significant after conditional analyses are included in the table. SNPs that achieved genome-wide significance threshold (ALL, $P = 2.50E-07$; EUR, $P = 2.88E-07$; AFR, $P = 3.30E-07$; ASI, $P = 1.70E-06$; and HIS, $P = 4.67E-07$) are shown in **bold**.

ALL, all ancestries (only EUR + AFR for FVII and vWF); NA, not applicable.

*AA change, amino acid change of SNP.

†Amino acid position in parentheses is for the mature protein for *FGB* (position 30) and *FGG* (position 26).

‡MAF, minor allele frequency from CHARGE joint calling.

All gene-based tests were again restricted to include only functional single nucleotide variants. Random effects (sequence kernel association test [SKAT][21]) and unidirectional[22] (T5) gene tests were performed using only variants with an MAF <5%. The T5 burden was defined as the total number of rare alleles among variants in the gene with an MAF <5%.[23] All genes were required to contain more than 1 variant to be included in the analysis and to have a cumulative MAF greater than the frequency such that the meta-analysis sample size would have an expected minor allele count of 5. A Bonferroni-corrected, gene-based *P* value threshold of 1.9×10^{-6} was used for gene-based tests (0.05/26 965 genes).

Meta-analysis. Meta-analyses of single variants and gene-based analyses were performed by using seqMeta v1.3. The primary analysis was to meta-analyze all ancestries together, with a secondary set of ancestry-specific analyses performed to complement and inform the results of the primary analysis. All significant nonsynonymous variants were re-annotated by using an updated version of dbNSFP (v.3.0).[19,20,24,25]

Conditional analyses. To test for independence of the new discoveries from variants previously demonstrated to be associated with the phenotype at that locus, conditional analyses were performed and meta-analyzed. These analyses were undertaken for EUR and AFR ancestry cohorts only, and in some cases, the single nucleotide polymorphisms (SNPs) that were conditioned on differed between ancestry groups, generally because of the conditional SNP being monomorphic in 1 population. A description of conditional analyses undertaken is included in supplemental Table 3.

Results

Single-variant and gene-based tests for all 4 hemostatic factors identified significantly associated loci for all phenotypes. The Q-Q plots for all association analyses are found in supplemental Figures 1-3. Functional annotations for all significant nonsynonymous single variants can be found in supplemental Table 2.

Fibrinogen

Exome array genotyping and fibrinogen measures were available for 76 316 participants across 16 cohorts and 4 ancestry groups.

Single-variant testing. Associations for 6 rare or low-frequency variants that exceeded array-wide significance were observed within 4 genes: 2 fibrinogen structural genes (*FGB* and *FGG*) and 2 other genes (*PDLIM4* and *HNF4A*) (Table 2 and supplemental Figure 4).

Two rare variants within *FGB*, rs6054 (Pro235Leu; MAF, 0.0042; $P = 1.8 \times 10^{-43}$) and rs201909029 (Lys148Asn; MAF, 0.00077; $P = 3.2 \times 10^{-13}$) were associated with lower fibrinogen levels. Both variants had similar effect sizes (–0.111 and −0.139 ln[g/L]) and the magnitude and direction of the association was similar for both variants in all ancestry groups (Table 2). Fibrinogen levels were lower by 10.5% and 13.0%, respectively, per copy of the minor allele when other model factors were fixed (see supplemental Data). The rs6054 association has been reported previously,[10] but the rs201909029 variant association is new. Two rare variants within *FGG* were also associated with fibrinogen levels: rs148685782 (Ala82Gly; MAF, 0.0033; $P = 9.2 \times 10^{-152}$) and rs145051028 (Ser219Phe; MAF, 0.00016; $P = 4.8 \times 10^{-09}$). In this study, rs148685782 had an effect size of −0.238 ln(g/L), which translates to a 21.1% lower fibrinogen level per copy of the minor allele. The direction and magnitude of the effect was similar across all ancestry groups in which it was polymorphic (Table 2). The *FGG* Ala82Gly variant has previously been associated with low plasma fibrinogen levels.[26-28] The rs145051028 variant has an effect size of −0.239 ln(g/L) or a 21.3% lower level of fibrinogen per copy of the minor allele and was polymorphic only in AFR ancestry cohorts. This association has not been previously reported.

To determine whether the newly and previously identified associations within the fibrinogen gene cluster were independent of one another, 3 separate conditional analyses were undertaken: (1) adjustment for previously associated common variants in *FGB* (rs4220 and rs6056),[10] (2) adjustment for the significant rare variants in *FGG* (rs148685782 and rs145051028; AFR only), and (3) adjustment for the most significant rare variant in *FGB* (rs6054) (supplemental Table 3). Results demonstrated independence of all variants from one another (Table 3). In total, the rare variants within the fibrinogen gene cluster explained ~1.3% and ~0.12% of the trait variance in the EUR and AFR populations, respectively. The majority of the variance in the EUR population (~0.9%) was attributed to *FGG* rs148685782.

The association of low-frequency variants within the *PDLIM4* and *HNF4A* genes supports prior reported associations. The *PDLIM4* SNP was in high linkage disequilibrium with previously reported *IRF1* SNP rs11242111 (r^2, 0.85; D′, 1 within 1000 Genomes Map Pilot 1 v.3, CEU) on chromosome 5,[9] and the *HNF4A* SNP rs1800961 has been previously reported, although it was just below the genome-wide significance threshold in that study.[10] The effect size for each was 10-fold smaller than those for *FGB* and *FGG*.

Single variants in *KCNT1* and in *HID1*, located in regions not previously reported to be associated with fibrinogen levels, reached array-wide significance in the exploratory AFR only analysis of fibrinogen (Table 2 and supplemental Figure 4). *KCNT1* rs151272083 (MAF, 0.00072; $P = 2.7 \times 10^{-07}$) codes for an Arg891Gln change (also reported as the same amino acid change at position 865, 877, or 910 because of transcriptional variation) and was predicted to decrease fibrinogen by 0.330 ln(g/L) or approximately 28.1% per copy of the minor allele in the AFR population. This SNP was also polymorphic in EUR populations but did not reach statistical significance, and the estimated effect was 20-fold smaller (β, 0.017; $P = .13$). *HID1* rs141869748 (Ile421Thr/Ille193Thr; MAF, 0.0013; $P = 4.0 \times 10^{-08}$) was associated with 0.252 ln(g/L) lower fibrinogen (22.3% decrease per copy of the minor allele) in the AFR population. This SNP was monomorphic in the EUR and ASI populations, and its estimated effect in the HIS population, although small, was not in the same direction despite a similar MAF (MAF, 0.0011; β, 0.008; $P = .94$).

When we further explored these characteristics of the novel associations in the AFR population, we found no evidence for heterogeneity across studies (P_{het}, 0.07 [rs151272083] and 0.91 [rs141869748]; supplemental Figure 5), and we confirmed that carriers of the variant allele in AFR cohorts had lower mean plasma fibrinogen levels than noncarriers (supplemental Table 5). The variants explained approximately 0.7% (rs151272083) and 0.4% (rs141869748) of the trait variance.

Gene-based testing. SKAT and T5 tests yielded gene-level associations with all 4 genes described earlier: *FGB, FGG, PDLIM4*, and *HNF4A* (Table 4). Gene-based testing did not identify other genes that contributed to plasma-level variation in fibrinogen.

FVII

Exome array genotyping and coagulation FVII measures were available for 25 372 participants across 7 studies of EUR and AFR participants.

Single-variant testing. Five exome-wide significant coding rare-variant associations were observed in *F7* as well as nearby genes *MCF2L* and *PROZ*. When conditioning on the common, previously reported coding variant rs6046 in *F7*,[13] 3 previously unreported rare variants within *F7* remained exome-wide significant, whereas the variants in *MCF2L* and *PROZ* were no longer significant (Table 3). The minor alleles of *F7* variants rs150525536 (Arg117Gln; MAF, 0.0010; $P_{cond} = 1.0 \times 10^{-22}$), rs121964926 (Arg342Gln; MAF, 0.0015; $P_{cond} = 1.5 \times 10^{-14}$), and rs3093248 (Glu423Lys; MAF, 0.00085; $P_{cond} = 1.4 \times 10^{-07}$) were all associated with significantly lower plasma FVII levels (Table 2 and supplemental Figure 4). The three variants explained ~0.06% of the trait variance in EUR participants and 4.5% of the trait variance in AFR participants. For all identified variants, the MAF was lower in EUR than in AFR populations but the direction of effect was the same even if the magnitude varied (Table 2). Sensitivity analyses that removed the 2 studies with FVII antigen rather than activity measured did not have an impact on the findings.

Gene-based testing. SKAT and T5 tests yielded gene-level associations with *F7* (Table 4). No other gene was associated with plasma levels of FVII.

FVIII and vWF

As reported by our prior GWASs, association results for plasma levels of FVIII and vWF were similar, so they will be presented together.[13] FVIII measures were available from 28 291 participants from 5 cohorts across all ancestry groups, whereas vWF was available in 23 272 EUR and AFR participants from 8 cohorts.

Single-variant testing. Genome-wide significant rare and low-frequency variants are presented in Table 2, and cluster plots for the associated SNPs are found in supplemental Figure 4. Five novel low-frequency and rare variant associations were found for FVIII and vWF levels, most within loci with previous FVIII/vWF associations.[13]

Low-frequency variant rs7962217 (Gly2705Arg; MAF, 0.046; $P = 2.5 \times 10^{-13}$) and rare variant rs41276738 (Arg854Gln; MAF, 0.0040; $P = 2.2 \times 10^{-13}$) in *VWF* were significantly associated with lower plasma levels of FVIII but not vWF ($P = .96$ and $P = .03$, respectively). Only the association of rs7962217 has been previously reported,[29] and conditioning on the most significant common *VWF* variants associated with FVIII levels (rs1063856 and rs62643635[13]) did not materially alter these results (Table 3). Ancestry-specific analyses yielded effects with the same direction and similar magnitudes, although the MAFs varied by up to 2 orders of magnitude (Table 2).

A single rare variant in *STAB2* rs141041254 (Glu2377Lys; MAF, 0.00087) was significantly associated with FVIII ($P = 2.1 \times 10^{-08}$), and vWF levels ($P = 2.4 \times 10^{-07}$) and the new signal remained unchanged when adjusting for rs2271637, the most highly associated *STAB2* common variant on the array. In the 2 ancestries in which the variant was polymorphic (AFR and EUR), the direction and the magnitude of the effects diverged (Table 2). This association has not been previously reported.

For FVIII and vWF levels, 11 significant single-variant associations were observed with rare or low-frequency variants within *ABO* and surrounding genes on chromosome 9. However, after conditioning on common variants tagging the major ABO blood types (A1, A2, B, and O), none of the 11 associations identified in this region remained. A description of these conditional analyses is presented in the supplemental Data and supplemental Table 4.

In exploratory analyses for the FVIII phenotype only, there was a significant association with a common variant on the X chromosome in *F8*, the gene encoding FVIII. This coding variant, rs1800291 (Asp1260Glu; MAF, 0.27; $P = 8.2 \times 10^{-08}$), had an MAF and effect direction that varied across ancestry groups (Table 2).

For the FVIII phenotype only, a rare variant in *KATNB1*, a gene not previously associated with FVIII levels, achieved array-wide

Table 3. Single-variant test meta-analysis results for conditional analyses of hemostatic factors fibrinogen, FVII, FVIII, and vWF

Factor and variant (gene)	Ancestry	No. of participants included in analysis*	*P*			
			UNCOND†	COND1	COND2	COND3
Fibrinogen						
rs201909029 *(FGB)*	ALL	46 841	**1.97E-10**	**1.35E-09**	**2.27E-10**	**3.44E-10**
	EUR	40 091	**2.69E-10**	**1.83E-09**	**3.10E-10**	**4.68E-10**
	AFR	6 750	4.25E-01	4.24E-01	4.21E-01	4.25E-01
rs6054 *(FGB)*	ALL	46 841	**1.00E-41**	**6.72E-39**	**2.67E-42**	
	EUR	40 091	**4.86E-41**	**3.40E-38**	**5.46E-42**	
	AFR	6 750	7.66E-02	7.25E-02	1.97E-01	
rs145051028 *(FGG)*	ALL	46 841	2.93E-06	2.67E-06		2.90E-06
	EUR	40 091	NA	NA	NA	NA
	AFR	6 750	2.93E-06	2.67E-06		2.90E-06
rs148685782 *(FGG)*	ALL	46 841	**3.24E-144**	**6.52E-137**		**2.49E-143**
	EUR	40 091	**1.03E-143**	**2.16E-136**		**8.02E-143**
	AFR	6 750	9.46E-02	9.52E-02		9.43E-02
FVII						
rs150525536 (*F7*)	ALL	20 549	**8.29E-20**	**1.02E-22**		
	EUR	16 338	2.23E-01	1.20E-01		
	AFR	4 211	**3.45E-20**	**7.56E-23**		
rs121964926 (*F7*)	ALL	20 549	**5.71E-14**	**1.49E-14**		
	EUR	16 338	9.25E-01	5.80E-01		
	AFR	4 211	**1.75E-20**	**1.95E-20**		
rs3093248 (*F7*)	ALL	20 549	2.54E-06	**1.35E-07**		
	EUR	16 338	NA	NA		
	AFR	4 211	2.54E-06	**1.35E-07**		
FVIII						
rs7962217 (*VWF*)	ALL	25 477	**6.60E-11**	**1.64E-09**		
	EUR	20 030	**8.69E-10**	**1.39E-08**		
	AFR	5 447	1.18E-02	2.35E-02		
rs41276738 (*VWF*)	ALL	25 477	**1.56E-11**	**9.85E-14**		
	EUR	20 030	**1.52E-10**	**1.41E-12**		
	AFR	5 447	5.96E-03	3.47E-03		
rs141041254 (*STAB2*)	ALL	25 477	**7.37E-09**	**4.11E-09**		
	EUR	20 030	**4.03E-09**	**2.22E-09**		
	AFR	5 447	9.17E-01	9.20E-01		
vWF						
rs141041254 (*STAB2*)	ALL	22 636	**6.82E-08**	**3.29E-08**		
	EUR	18 236	**2.85E-08**	**1.34E-08**		
	AFR	4 400	7.46E-01	7.49E-01		

SNPs achieving genome-wide significance threshold (ALL, $P = 2.57E-07$; EUR, 2.88E-07; AFR, 3.30E-07) are shown in **bold**.
*Only EUR and AFR cohorts were asked to run conditional analyses and not all cohorts participated.
†UNCOND, unadjusted analyses; a description of conditional (COND) analyses is provided in supplemental Table 3.

significance in the AFR population. This variant, rs142508811, was rare in both EUR and AFR populations and was monomorphic in ASI and HIS populations; the estimated effect size was 80-fold larger in AFR than in EUR populations. Across the studies with AFR populations, there was no evidence of heterogeneity ($P_{het,}$ 0.74); a forest plot for these associations is presented in supplemental Figure 5. Levels of FVIII in carriers of the variant allele had a higher mean FVIII than noncarriers (supplemental Table 5).

For the FVIII phenotype, the 5 variants explained approximately 0.9% of the phenotype variation in both EUR and AFR populations. For the vWF phenotype, the *STAB2* variant explained 0.2% and 0% in EUR and AFR populations, respectively.

Gene-based testing. For FVIII levels, *ABO*, *VWF*, and *STAB2* yielded gene-wide significant associations with SKAT testing, whereas *ABO* and *VWF* were significant with T5 testing (Table 4). For vWF levels, *ABO* and *STAB2* yielded gene-wide significant associations with SKAT testing, whereas *ABO* was significant with T5 testing; the *VWF* gene did not achieve significance for vWF. No new associations were identified through gene-based testing.

Discussion

We identified 12 novel associations of low-frequency (n = 2) and rare (n = 10) variants across the fibrinogen, FVII, FVIII, and vWF traits that were independent of previously identified associations. Nine of the variants were within genes previously established as associated with the trait; findings for associations in 3 new candidate loci were detected in people of AFR ancestry, possibly because of monomorphic or much lower frequency characteristics of these variants in all other ancestries. These newly identified associations accounted for modest amounts of the variance explained and suggest that, at most, a small proportion of the missing heritability can be attributable to them. The gene-based tests did not reveal new loci.

Fibrinogen

Associations of rare variants with fibrinogen levels were found in gene regions previously associated with fibrinogen by common variant GWASs. The association of *FGB* rare variant rs6054 with

Table 4. Gene-based test meta-analysis results for hemostatic factors fibrinogen, FVII, FVIII, and vWF

Factor and gene	Ancestry	No. of participants included in analysis	*P*	
			SKAT5	T5
Fibrinogen				
FGB	ALL	76 316	**1.25E-45**	**5.59E-32**
	EUR	65 733	**2.03E-44**	**1.16E-36**
	AFR	8 388	4.50E-01	5.60E-01
	ASI	764	3.00E-01	2.98E-01
	HIS	1 431	9.37E-01	9.39E-01
FGG	ALL	76 316	**6.90E-99**	**7.25E-31**
	EUR	65 733	**2.49E-111**	**1.35E-61**
	AFR	8 388	**2.82E-09**	3.18E-04
	ASI	764	NA	NA
	HIS	1 431	5.65E-01	8.18E-01
FVII				
F7	ALL	25 372	**6.24E-35**	**2.36E-37**
	EUR	20 401	6.71E-05	**8.21E-07**
	AFR	4 971	**1.83E-35**	**3.03E-32**
FVIII				
ABO	ALL	28 291	**5.10E-18**	**5.71E-30**
	EUR	20 030	**1.90E-13**	**1.61E-17**
	AFR	6 079	1.91E-03	3.44E-04
	ASI	764	8.37E-01	9.56E-01
	HIS	1 418	3.48E-01	2.89E-02
VWF	ALL	28 291	**5.21E-21**	**1.61E-06**
	EUR	20 030	**2.20E-07**	1.47E-04
	AFR	6 079	8.13E-03	4.09E-01
	ASI	764	1.41E-01	8.01E-01
	HIS	1 418	2.27E-01	4.07E-01
STAB2	ALL	28 291	**3.49E-07**	2.56E-03
	EUR	20 030	**6.49E-07**	5.83E-03
	AFR	6 079	1.44E-01	8.23E-02
	ASI	764	1.78E-01	9.55E-02
	HIS	1 418	9.13E-01	3.09E-01
vWF				
ABO	*ALL*	23 272	**4.07E-19**	**3.69E-29**
	EUR	18 236	**2.84E-13**	**4.17E-18**
	AFR	5 036	2.89E-03	3.01E-04
STAB2	*ALL*	23 272	**2.99E-07**	8.07E-03
	EUR	18 236	**1.53E-06**	1.66E-01
	AFR	5 036	7.24E-04	6.46E-02

Genes that achieved genome-wide significance ($P < 1.85E-06$) are shown in **bold**.

lower fibrinogen has been previously reported.[10] Although the association of *FGB* rs201909029 is a novel finding in this context, it has been reported in mild hypofibrinogenemia cases[26] in clinical databases (MERIVALE II),[30] although it has not been reported to cause hemorrhage or thrombosis.[30] The rare *FGG* variant rs148685782 was associated with hypofibrinogenemia and hemorrhage[26-28] in multiple affected individuals. *FGG* rs145051028, which was associated with fibrinogen levels in AFR cohorts only, has not been reported in clinical databases or population studies. This may be a result of the low MAF and also a lack of studies that included AFR participants. Conditional analyses showed that the common and rare variant associations across the fibrinogen gene cluster were independent, an observation supported by their low r^2 for the pairwise linkage disequilibrium. Within the fibrinogen gene cluster, the 4 significant *FGB* and *FGG* rare variants explained two- to fourfold more trait variance than the common *FGB* rs4220 variant,[7,9,10,14,31] which had an effect size of 0.029 ln(g/L) or a 2.9% higher level of fibrinogen per copy of the minor allele in this study.

In exploratory ancestry-stratified analyses, the associations of *KCNT1* and *HID1* with fibrinogen in AFR participants were the only findings that identified new candidate loci tha influence fibrinogen regulation. These findings can only be considered hypothesis generating and require replication.

FVII

We identified 3 rare coding variants in the FVII protein structural gene *F7* associated with plasma levels of FVII, none of which were previously reported in the epidemiologic literature. rs150525536 was rare in the AFR population and had a 10-fold lower frequency in the EUR population. A previous case report of this variant was found in a male with an EUR ancestry homozygote who had severe FVII deficiency and who also carried another *F7* mutation (Arg212Gln).[32] Both mutations were thought to contribute to the phenotype. The mutation reported here is found in the first epidermal growth factor-like domain and is required for binding to tissue factor, its cofactor. It causes reduced binding to tissue factor and reduced clotting ability in a concentration-dependent manner as well as slower activation.[32] Variant rs121964926 was also more common among the AFR population than in the EUR population. It has been observed clinically in both asymptomatic and symptomatic individuals with FVII activity <5% from Germany and France as well as patients with reduced FVII activity from Costa Rica, Venezuela, and the United States.[33] Nothing has been reported regarding clinical consequences of the rs3093248 variant.

FVIII and vWF

The findings for the vWF trait consisted of a subset of the FVIII results. None of the associations between variants within the *ABO* gene region and FVIII/vWF were independent of established ABO blood group alleles. Two rare variants in *VWF*, rs7962217 and rs41276738, were associated with plasma FVIII levels. rs7962217 was associated with higher FVIII levels whereas rs41276738 was associated with lower levels and had an effect size similar to that of the strongest genetic predictor of FVIII levels, the O-deletion tagging SNP (rs657152). rs41276738 has been reported in patients with von Willebrand disease type 1[34,35] and type 2N,[36-43] but the association with vWF levels did not reach exome-wide significance, although its direction was consistent with the direction of effects on FVIII. The *STAB2* variant rs141041254 was associated with higher plasma levels of both FVIII and vWF. The effect size was more than 10-fold larger than that reported for the more common *STAB2* variant rs2271637 (β_{FVIII}, 1.95%; β_{vWF}, 2.47%). The common *F8* coding variant rs1800291 was associated with a much smaller effect on FVIII compared with the *ABO* O-deletion variant. It has been previously reported,[29,44,45] and in the European Association for Haemophilia and Allied Disorders (EAHAD) Coagulation Factor Variants Database, it is annotated as unlikely to be pathogenic. The *KATNB1* rs142508811 variant and FVIII association was restricted to the AFR population, although MAF and direction of effect were similar across the 2 polymorphic populations.

Clinical implications

Inferring causality of uncommon and rare variants with a phenotypic expression is challenging and requires strong statistical evidence combined with experimental data.[46] Inferring clinical implications from the causal variants requires additional evidence[47] not available in our approach. In this article, we identified rare variants associated with higher or lower phenotype levels in 4 hemostasis measures. Some of the variants have been found in patients with diseases related to blood clotting, which suggests that these genes and their uncommon and rare genetic variation may play a role in a clinical phenotype.[26-28,32-43] The distribution of the phenotypes within our research populations were within the extremes of a clinically

important range (FVII: 0.80-11.40 g/L [fibrinogen], 26% to 441% activity, and 2% to 297% antigen; FVIII: 14% to 500% activity; and vWF: 2% to 374% antigen). Furthermore, the magnitude of difference in the phenotype associated with the variant was mostly modest, although some were larger and were associated with a change equivalent to half the size of the estimated population mean for the phenotype of interest. Therefore, the magnitude of any clinically relevant effects of these variants would be expected to be small to modest. The findings from our study suggest that the contribution of the uncommon and rare variants to complex clinical phenotypes, such as arterial or venous thrombosis or hemorrhagic stroke, should be evaluated in large populations. This article identifies several variants which may be good potential candidates.

Limitations and strengths of the approach

We decided a priori to use all the phenotype-genotype association data for discovery to reduce false-negative findings,[48] but this approach provided us with no replication setting. Although these candidate variants are now well characterized, the rare allele frequencies will create challenges for replication in the absence of additional large phenotyped populations. However, our findings provide strong rationale for further functional genomic follow-up, and some of our observations confirm associations for several rare variants that have been reported in patients with the corresponding congenital clotting factor deficiencies. This investigation of low-frequency and rare variants on the 4 phenotypes was limited to the variants included on the BeadChip. Differing sample sizes for the meta-analysis between phenotypes likely affected our power to detect associations, but this may also be influenced by biological differences. In addition, we did not have the statistical power to test for differences in associations across the 4 ancestries. Although it was not an aim of our study, a subsequent effort with this objective would be worthwhile to better understand the genetic architecture of the phenotypes. Finally, although we enriched our variant population with those predicted to be causal, we cannot attribute causality to the variants with novel associations.

The quality of rare variant genotype calling was maximized by the joint clustering performed within CHARGE on thousands of samples.[17] By incorporating individuals of non-European ancestry in the primary analysis, we increased our power to detect association in which variants may be more frequent or genetic diversity greater in one ancestry group than another. It also allowed us to broadly look at ancestry-specific gene and rare-variant associations but was vastly underpowered to draw any strong conclusions.

In conclusion, in meta-analyses of 4 hemostatic factors and functionally enriched exonic variants, novel associations of low-frequency and rare variants were identified in 16 studies that included 4 ancestries. Novel variant associations were found within previously reported genes, and they had effect sizes that were often independent of and much larger than previously reported common variants. In addition, rare variant associations at *KCNT1*, *HID1,* and *KATNB1* identify new candidate genes related to hemostasis for follow-up replication and functional genomic analysis.

Acknowledgments

A complete list of acknowledgments for each cohort is found in the supplemental Data.

Authorship

Contribution: D.V., D.M.B., R.A.M., O.P., A.F.W., C.H., U.V., D.I.C., P.M.R., J.M.S., I.J.D., B.M., B.M.P., N.L.S., W.K., A.H., O.H.F., C.J.O., A.P.R., E.B., and J.I.R. designed the research; J.E.H., O.P., C.H., T.K., U.V., D.I.C., L.M.R., S.E.H., J.M.S., I.J.D., M.S.-L., A.R.F., J.A.B., K.L.W., K.D.T., K.S., A.H., O.H.F., D.L., C.J.O., P.L.A., M.F., N.P., X.G., and J.Y. performed the research; A.T., D.I.C., F.G., M.L.G., A.G., R.J.S., B.S., A.S., M.M.-N., M.W., W.K., M.P.M.d.M., F.R., A.G.U., P.L.A., L.Q., and C.K. contributed vital new reagents or analytical tools; L.R.Y., D.V., D.M.B., R.A.M., O.P., C.H., A.G., D.I.C., F.G., S.E.H., J.M.S., H.W., A.H., A.R.F., B.M.P., M.W., W.K., M.P.M.d.M., F.R., A.G.U., A.H., O.H.F., D.L., G.H.T., C.J.O., P.L.A., C.K., A.P.R., E.B., M.F., M.C., C.-C.H., and N.Z. collected data; L.R.Y., D.V., D.M.B., R.A.M., J.E.H., C.H., M.L.G., T.K., D.I.C., L.M.R., F.G., R.E.M., S.E.H., M.S.-L., A.C.M., J.A.B., K.D.T., B.M., B.M.P., N.L.S., H.R., P.S.d.V., A.D., M.-H.C., G.H.T., C.J.O., C.K., A.P.R., M.F., N.P., X.G., J.Y., W.T., and J.I.R. analyzed and interpreted data; L.R.Y., J.E.H., T.K., D.I.C., L.M.R., R.E.M., M.S.-L., A.C.M., J.A.B., M.M.-N., H.R., P.S.d.V., A.D., M.-H.C., P.L.A., L.-A.L., X.G., and J.Y. performed statistical analysis; and J.E.H., R.E.M., S.E.H., I.J.D., P.S.d.V., G.H.T., A.C.M., A.P.R., C.J.O., and N.L.S. wrote the manuscript. All co-authors were given the opportunity to revise and comment on the text and content of manuscript. This manuscript was approved by all relevant cohort Publication and Presentation committees prior to submission.

Conflict-of-interest disclosure: D.V. is a consultant for MBC, Inc. D.I.C. and P.M.R. have received funding for exome chip genotyping in the Women's Genome Health Study and collaborative scientific support from Amgen. B.M.P. serves on the data and safety management board for a clinical trial of a device funded by the manufacturer (ZOLL Lifecor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. O.H.F. is employed by ErasmusAGE, a center for aging research across the life course funded by Nestlé Nutrition (Nestec Ltd.), Metagenics Inc., and AXA. Nestlé Nutrition, Metagenics Inc., and AXA had no role in the design and conduct of the study, the collection, management, analysis, and interpretation of the data, or in the preparation, review, or approval of the manuscript. The remaining authors declare no competing financial interests

Correspondence: Nicholas L. Smith, 1730 Minor Ave, Suite 1360, Seattle, WA 98101; e-mail: nlsmith@u.washington.edu.

References

1. Folsom AR. Hemostatic risk factors for atherothrombotic disease: an epidemiologic view. *Thromb Haemost.* 2001;86(1):366-373.
2. Folsom AR, Cushman M, Heckbert SR, Ohira T, Rasmussen-Torvik L, Tsai MY. Factor VII coagulant activity, factor VII -670A/C and -402G/A polymorphisms, and risk of venous thromboembolism. *J Thromb Haemost.* 2007;5(8):1674-1678.
3. Smith A, Patterson C, Yarnell J, Rumley A, Ben-Shlomo Y, Lowe G. Which hemostatic markers add to the predictive value of conventional risk factors for coronary heart disease and ischemic stroke? The Caerphilly Study. *Circulation.* 2005;112(20):3080-3087.
4. Danesh J, Lewington S, Thompson SG, et al; Fibrinogen Studies Collaboration. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant

meta-analysis. *JAMA.* 2005;294(14):1799-1809.

5. Koster T, Blann AD, Briët E, Vandenbroucke JP, Rosendaal FR. Role of clotting factor VIII in effect of von Willebrand factor on occurrence of deep-vein thrombosis. *Lancet.* 1995;345(8943):152-155.
6. Spiel AO, Gilbert JC, Jilma B. von Willebrand factor in cardiovascular disease: focus on acute coronary syndromes. *Circulation.* 2008;117(11):1449-1459.
7. Danik JS, Paré G, Chasman DI, et al. Novel loci, including those related to Crohn disease, psoriasis, and inflammation, identified in a genome-wide association study of fibrinogen in 17 686 women: the Women's Genome Health Study. *Circ Cardiovasc Genet.* 2009;2(2):134-141.
8. Lovely RS, Yang Q, Massaro JM, et al. Assessment of genetic determinants of the association of γ′ fibrinogen in relation to cardiovascular disease. *Arterioscler Thromb Vasc Biol.* 2011;31(10):2345-2352.
9. Sabater-Lleal M, Huang J, Chasman D, et al; VTE Consortium; STROKE Consortium; Wellcome Trust Case Control Consortium 2 (WTCCC2); C4D Consortium; CARDIoGRAM Consortium. Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. *Circulation.* 2013;128(12):1310-1324.
10. Wassel CL, Lange LA, Keating BJ, 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.* 2011;117(1):268-275.
11. Johnsen JM, Auer PL, Morrison AC, et al; NHLBI Exome Sequencing Project. Common and rare von Willebrand factor (VWF) coding variants, VWF levels, and factor VIII levels in African Americans: the NHLBI Exome Sequencing Project. *Blood.* 2013;122(4):590-597.
12. Taylor KC, Lange LA, Zabaneh D, et al. A gene-centric association scan for Coagulation Factor VII levels in European and African Americans: the Candidate Gene Association Resource (CARe) Consortium. *Hum Mol Genet.* 2011;20(17):3525-3534.
13. Smith NL, Chen MH, Dehghan A, et al; Wellcome Trust Case Control Consortium. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation.* 2010;121(12):1382-1392.
14. Dehghan A, Yang Q, Peters A, et al; Wellcome Trust Case Control Consortium. Association of novel genetic Loci with circulating fibrinogen levels: a genome-wide association study in 6 population-based cohorts. *Circ Cardiovasc Genet.* 2009;2(2):125-133.
15. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747-753.
16. Psaty BM, O'Donnell CJ, Gudnason V, et al; CHARGE Consortium. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet.* 2009;2(1):73-80.
17. Grove ML, Yu B, Cochran BJ, et al. Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PLoS One.* 2013;8(7):e68095.
18. Peloso GM, Auer PL, Bis JC, et al; NHLBI GO Exome Sequencing Project. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am J Hum Genet.* 2014;94(2):223-232.
19. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Hum Mutat.* 2013;34(9):E2393-E2402.
20. Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat.* 2011;32(8):894-899.
21. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet.* 2011;89(1):82-93.
22. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet.* 2008;83(3):311-321.
23. Morris AP, Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. *Genet Epidemiol.* 2010;34(2):188-193.
24. Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet.* 2015;24(8):2125-2137.
25. Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res.* 2014;42(22):13534-13544.
26. Ivaskevicius V, Jusciute E, Steffens M, et al. gammaAla82Gly represents a common fibrinogen gamma-chain variant in Caucasians. *Blood Coagul Fibrinolysis.* 2005;16(3):205-208.
27. Brennan SO, Fellowes AP, Faed JM, George PM. Hypofibrinogenemia in an individual with 2 coding (gamma82 A→G and Bbeta235 P→L) and 2 noncoding mutations. *Blood.* 2000;95(5):1709-1713.
28. Wyatt J, Brennan SO, May S, George PM. Hypofibrinogenaemia with compound heterozygosity for two gamma chain mutations - gamma 82 Ala→Gly and an intron two GT→AT splice site mutation. *Thromb Haemost.* 2000;84(3):449-452.
29. Tang W, Cushman M, Green D, et al. Gene-centric approach identifies new and known loci for FVIII activity and VWF antigen levels in European Americans and African Americans. *Am J Hematol.* 2015;90(6):534-540.
30. Maghzal GJ, Brennan SO, George PM. Fibrinogen B beta polymorphisms do not directly contribute to an altered in vitro clot structure in humans. *Thromb Haemost.* 2003;90(6):1021-1028.
31. Schmelzer CH, Ebert RF, Bell WR. A polymorphism at B beta 448 of fibrinogen identified during structural studies of fibrinogen Baltimore II. *Thromb Res.* 1988;52(2):173-177.
32. Chaing S, Clarke B, Sridhara S, et al. Severe factor VII deficiency caused by mutations abolishing the cleavage site for activation and altering binding to tissue factor. *Blood.* 1994;83(12):3524-3535.
33. Herrmann FH, Wulff K, Auerswald G, et al; Greifswald Factor FVII Deficiency Study Group. Factor VII deficiency: clinical manifestation of 717 subjects from Europe and Latin America with mutations in the factor 7 gene. *Haemophilia.* 2009;15(1):267-280.
34. Goodeve A, Eikenboom J, Castaman G, et al. Phenotype and genotype of a cohort of families historically diagnosed with type 1 von Willebrand disease in the European study, Molecular and Clinical Markers for the Diagnosis and Management of Type 1 von Willebrand Disease (MCMDM-1VWD). *Blood.* 2007;109(1):112-121.
35. James PD, Notley C, Hegadorn C, et al. The mutational spectrum of type 1 von Willebrand disease: Results from a Canadian cohort study. *Blood.* 2007;109(1):145-154.
36. Corrales I, Catarino S, Ayats J, et al. High-throughput molecular diagnosis of von Willebrand disease by next generation sequencing methods. *Haematologica.* 2012;97(7):1003-1007.
37. Schneppenheim R, Brassard J, Krey S, et al. Defective dimerization of von Willebrand factor subunits due to a Cys→Arg mutation in type IID von Willebrand disease. *Proc Natl Acad Sci USA.* 1996;93(8):3581-3586.
38. Eikenboom JC, Reitsma PH, Peerlinck KM, Briët E. Recessive inheritance of von Willebrand's disease type I. *Lancet.* 1993;341(8851):982-986.
39. Mazurier C. von Willebrand disease masquerading as haemophilia A. *Thromb Haemost.* 1992;67(4):391-396.
40. Peerlinck K, Eikenboom JC, Ploos Van Amstel HK, et al. A patient with von Willebrand's disease characterized by a compound heterozygosity for a substitution of Arg854 by Gln in the putative factor-VIII-binding domain of von Willebrand factor (vWF) on one allele and very low levels of mRNA from the second vWF allele. *Br J Haematol.* 1992;80(3):358-363.
41. Gaucher C, Jorieux S, Mercier B, Oufkir D, Mazurier C. The "Normandy" variant of von Willebrand disease: characterization of a point mutation in the von Willebrand factor gene. *Blood.* 1991;77(9):1937-1941.
42. Kroner PA, Friedman KD, Fahs SA, Scott JP, Montgomery RR. Abnormal binding of factor VIII is linked with the substitution of glutamine for arginine 91 in von Willebrand factor in a variant form of von Willebrand disease. *J Biol Chem.* 1991;266(29):19146-19149.
43. Cacheris PM, Nichols WC, Ginsburg D. Molecular characterization of a unique von Willebrand disease variant. A novel mutation affecting von Willebrand factor/factor VIII interaction. *J Biol Chem.* 1991;266(21):13499-13502.
44. Viel KR, Machiah DK, Warren DM, et al. A sequence variation scan of the coagulation factor VIII (FVIII) structural gene and associations with plasma FVIII activity levels. *Blood.* 2007;109(9):3713-3724.
45. Scanavini D, Legnani C, Lunghi B, Mingozzi F, Palareti G, Bernardi F. The factor VIII D1241E polymorphism is associated with decreased factor VIII activity and not with activated protein C resistance levels. *Thromb Haemost.* 2005;93(3):453-456.
46. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. *Nature.* 2014;508(7497):469-476.
47. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405-423.
48. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet.* 2006;38(2):209-213.
49. The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol.* 1989;129(4):687-702.
50. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some

characteristics of the examined subjects. *J Clin Epidemiol.* 1988;41(11):1105-1116.

51. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol.* 1991;1(3):263-276.

52. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med.* 1975;4(4):518-525.

53. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J III. Factors of risk in the development of coronary heart disease–six year follow-up experience. The Framingham Study. *Ann Intern Med.* 1961;55:33-50.

54. Vaidya D, Yanek LR, Moy TF, Pearson TA, Becker LC, Becker DM. Incidence of coronary artery disease in siblings of patients with premature coronary artery disease: 10 years of follow-up. *Am J Cardiol.* 2007;100(9):1410-1415.

55. Holle R, Happich M, Löwel H, Wichmann HE; MONICA/KORA Study Group. KORA–a research platform for population based health research. *Gesundheitswesen.* 2005;67(Suppl 1): S19-S25.

56. Wichmann HE, Gieger C, Illig T; MONICA/KORA Study Group. KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen.* 2005; 67(Suppl 1):S26-S30.

57. Zemunik T, Boban M, Lauc G, et al. Genome-wide association study of biochemical traits in Korcula Island, Croatia. *Croat Med J.* 2009;50(1):23-33.

58. Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. *Int J Epidemiol.* 2012;41(6):1576-1584.

59. Deary IJ, Whiteman MC, Starr JM, Whalley LJ, Fox HC. The impact of childhood intelligence on later life: following up the Scottish mental surveys of 1932 and 1947. *J Pers Soc Psychol.* 2004; 86(1):130-147.

60. Deary IJ, Gow AJ, Taylor MD, et al. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr.* 2007;7:28.

61. Bild DE, Bluemke DA, Burke GL, et al. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol.* 2002;156(9):871-881.

62. Clarke R, Peden JF, Hopewell JC, et al; PROCARDIS Consortium. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med.* 2009;361(26): 2518-2528.

63. Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol.* 2009;24(9):553-572.

64. Hofman A, Darwish Murad S, van Duijn CM, et al. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol.* 2013;28(11):889-926.

65. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol.* 1991;7(4):403-422.

66. Samnegård A, Silveira A, Lundman P, et al. Serum matrix metalloproteinase-3 concentration is influenced by MMP-3 -1612 5A/6A promoter genotype and associated with myocardial infarction. *J Intern Med.* 2005;258(5):411-419.

67. Völzke H, Alte D, Schmidt CO, et al. Cohort profile: the study of health in Pomerania. *Int J Epidemiol.* 2011;40(2):294-307.

68. Ridker PM, Chasman DI, Zee RY, et al; Women's Genome Health Study Working Group. Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin Chem.* 2008;54(2): 249-255.

69. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials.* 1998;19(1):61-109.

70. Anderson CA, Boucher G, Lees CW, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet.* 2011;43(3): 246-252.

71. Hays J, Hunt JR, Hubbell FA, et al. The Women's Health Initiative recruitment methods and results. *Ann Epidemiol.* 2003;13(9 Suppl):S18-S77.

2015 126: e19-e29
doi:10.1182/blood-2015-02-624551 originally published online June 23, 2015

Rare and low-frequency variants and their association with plasma levels of fibrinogen, FVII, FVIII, and vWF

Jennifer E. Huffman, Paul S. de Vries, Alanna C. Morrison, Maria Sabater-Lleal, Tim Kacprowski, Paul L. Auer, Jennifer A. Brody, Daniel I. Chasman, Ming-Huei Chen, Xiuqing Guo, Li-An Lin, Riccardo E. Marioni, Martina Müller-Nurasyid, Lisa R. Yanek, Nathan Pankratz, Megan L. Grove, Moniek P. M. de Maat, Mary Cushman, Kerri L. Wiggins, Lihong Qi, Bengt Sennblad, Sarah E. Harris, Ozren Polasek, Helene Riess, Fernando Rivadeneira, Lynda M. Rose, Anuj Goel, Kent D. Taylor, Alexander Teumer, André G. Uitterlinden, Dhananjay Vaidya, Jie Yao, Weihong Tang, Daniel Levy, Melanie Waldenberger, Diane M. Becker, Aaron R. Folsom, Franco Giulianini, Andreas Greinacher, Albert Hofman, Chiang-Ching Huang, Charles Kooperberg, Angela Silveira, John M. Starr, Konstantin Strauch, Rona J. Strawbridge, Alan F. Wright, Barbara McKnight, Oscar H. Franco, Neil Zakai, Rasika A. Mathias, Bruce M. Psaty, Paul M. Ridker, Geoffrey H. Tofler, Uwe Völker, Hugh Watkins, Myriam Fornage, Anders Hamsten, Ian J. Deary, Eric Boerwinkle, Wolfgang Koenig, Jerome I. Rotter, Caroline Hayward, Abbas Dehghan, Alex P. Reiner, Christopher J. O'Donnell and Nicholas L. Smith

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.