Cross-ancestry investigation of venousc genomic predictors

Thibord, F.; Klarin, D.; Brody, J.A.; Chen, M.H.; Levin, M.G.; Chasman, D.I.; ... ; Smith, N.L.

Circulation

ORIGINAL RESEARCH ARTICLE

Cross-Ancestry Investigation of Venous Thromboembolism Genomic Predictors

Florian Thibord, PhD*; Derek Klarin, MD*; Jennifer A. Brody, BA; Ming-Huei Chen, PhD; Michael G. Levin, MD; Daniel I. Chasman, PhD; Ellen L. Goode, PhD, MPH; Kristian Hveem, MD, PhD; Maris Teder-Laving, MS; Angel Martinez-Perez, MS; Dylan Aïssi, PhD; Delphine Daian-Bacq, PhD; Kaoru Ito, MD, PhD; Pradeep Natarajan, MD, MMSc; Pamela L. Lutsey, PhD; Girish N. Nadkarni, MD, MPH; Paul S. de Vries, PhD; Gabriel Cuellar-Partida, PhD; Brooke N. Wolford, PhD; Jack W. Pattee, PhD; Charles Kooperberg, PhD; Sigrid K. Braekkan, PhD; Ruifang Li-Gao, PhD; Noemie Saut, PhD; Corriene Sept, AM; Marine Germain, MS; Renae L. Judy, MS; Kerri L. Wiggins, MS, RD; Darae Ko, MD, MSc; Christopher J. O'Donnell, MD, MPH; Kent D. Taylor, PhD; Franco Giulianini, PhD; Mariza De Andrade, PhD; Therese H. Nøst, PhD; Anne Boland, PhD; Jean-Philippe Empana, PhD; Satoshi Koyama, PhD, MD; Thomas Gilliland, MD; Ron Do, PhD; Jennifer E. Huffman, PhD; Xin Wang, PhD; Wei Zhou, PhD; Jose Manuel Soria, PhD; Juan Carlos Souto, MD, PhD; Nathan Pankratz, PhD; Jeffery Haessler, MS; Kristian Hindberg, PhD; Frits R. Rosendaal, MD, PhD; Constance Turman, MS; Robert Olaso, PhD; Rachel L. Kember, MS, PhD; Traci M. Bartz, MS; Julie A. Lynch, PhD, RN, MBA; Susan R. Heckbert, MD, PhD; Sebastian M. Armasu, MS; Ben Brumpton, PhD; David M. Smadja, MD, PhD; Xavier Jouven, MD, PhD; Issei Komuro, MD, PhD; Katharine R. Clapham, MD; Ruth J.F. Loos, PhD; Cristen J. Willer, PhD; Maria Sabater-Lleal, PhD; James S. Pankow, PhD; Alexander P. Reiner, MD, MSc; Vania M. Morelli, MD, PhD; Paul M Ridker, MD, MPH; Astrid van Hylckama Vlieg, PhD; Jean-François Deleuze, PhD; Peter Kraft, PhD; Daniel J. Rader, MD; Global Biobank Meta-Analysis Initiative; Estonian Biobank Research Team; 23andMe Research Team; Biobank Japan; CHARGE Hemostasis Working Group; Kyung Min Lee, PhD; Bruce M. Psaty, MD, PhD; Anne Heidi Skogholt, PhD; Joseph Emmerich, MD, PhD; Pierre Suchon, MD, PhD; Stephen S. Rich, PhD; Ha My T. Vy, PhD; Weihong Tang, MD, PhD; Rebecca D. Jackson, MD; John-Bjarne Hansen, MD, PhD; Pierre-Emmanuel Morange, MD, PhD; Christopher Kabrhel, MD, MPH; David-Alexandre Trégouët, PhD†; Scott M. Damrauer, MD†; Andrew D. Johnson, PhD†; Nicholas L. Smith, PhD†

BACKGROUND: Venous thromboembolism (VTE) is a life-threatening vascular event with environmental and genetic determinants. Recent VTE genome-wide association studies (GWAS) meta-analyses involved nearly 30 000 VTE cases and identified up to 40 genetic loci associated with VTE risk, including loci not previously suspected to play a role in hemostasis. The aim of our research was to expand discovery of new genetic loci associated with VTE by using cross-ancestry genomic resources.

METHODS: We present new cross-ancestry meta-analyzed GWAS results involving up to 81 669 VTE cases from 30 studies, with replication of novel loci in independent populations and loci characterization through in silico genomic interrogations.

RESULTS: In our genetic discovery effort that included 55 330 participants with VTE (47 822 European, 6320 African, and 1188 Hispanic ancestry), we identified 48 novel associations, of which 34 were replicated after correction for multiple testing. In our combined discovery-replication analysis (81 669 VTE participants) and ancestry-stratified meta-analyses (European, African, and Hispanic), we identified another 44 novel associations, which are new candidate VTE-associated loci requiring replication. In total, across all GWAS meta-analyses, we identified 135 independent genomic loci significantly associated with VTE risk. A genetic risk score of the significantly associated loci in Europeans identified a 6-fold increase in risk for those in the top 1% of scores compared with those with average scores. We also identified 31 novel transcript

Correspondence to: Nicholas L. Smith, PhD, Department of Epidemiology, University of Washington, Seattle, WA 98195. Email nlsmith@u.washington.edu

*F. Thibord and D. Klarin contributed equally.

†D.-A. Trégouët , S.M. Damrauer, A.D. Johnson, and N.L. Smith contributed equally.

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/CIRCULATIONAHA.122.059675.

For Sources of Funding and Disclosures, see page 1239.

Circulation is available at www.ahajournals.org/journal/circ

associations in transcriptome-wide association studies and 8 novel candidate genes with protein quantitative-trait locus Mendelian randomization analyses. In silico interrogations of hemostasis and hematology traits and a large phenome-wide association analysis of the 135 GWAS loci provided insights to biological pathways contributing to VTE, with some loci contributing to VTE through well-characterized coagulation pathways and others providing new data on the role of hematology traits, particularly platelet function. Many of the replicated loci are outside of known or currently hypothesized pathways to thrombosis.

CONCLUSIONS: Our cross-ancestry GWAS meta-analyses identified new loci associated with VTE. These findings highlight new pathways to thrombosis and provide novel molecules that may be useful in the development of improved antithrombosis treatments.

Key Words: genetics ■ genome-wide association study ■ meta-analysis ■ venous thromboembolism ■ venous thrombosis

Clinical Perspective

What Is New?

- Our venous thromboembolism (VTE) genetic analyses revealed 135 loci associated with VTE, of which 92 were novel. Although novel VTE-associated variants were typically noncoding and displayed small odds ratios, they point at novel biological pathways involved in VTE.
- In particular, a large number of novel VTE variants are shared with platelet traits and located in loci with known roles in hematopoiesis or megakaryocyte development, which suggests that platelet generation, turnover, or reactivity may be a feature of VTE pathogenesis.

What Are the Clinical Implications?

- These results constitute a valuable resource for thrombosis researchers and for the discovery of new VTE therapeutic targets.
- A genetic risk score constructed from the European-specific results and applied to the UK Biobank participants of European ancestry explained ~5% of the phenotypic variance, and displayed a strong predictive ability with an area under the curve equal to 0.62.

Nonstandard Abbreviations and Acronyms

AA	African ancestry
AUC	area under the curve
CBC	complete blood count
EA	European ancestry
GRS	genetic risk score
GWAS	genome-wide association study
HIS	Hispanic ancestry
LD	linkage disequilibrium
MAF	minor allele frequency
MR	Mendelian randomization
OR	odds ratio
PAI-1	plasminogen activator inhibitor 1
PheWAS	phenome-wide association study
pQTL	protein quantitative trait loci
QTL	quantitative trait loci
TWAS	transcriptome-wide association study
VTE	venous thromboembolism
VWF	von Willebrand factor

Venous thrombosis is a vascular event resulting from an imbalance in the regulation of hemostasis, with subsequent pathologic coagulation and vascular thrombosis formation. Clinically, venous thrombosis can manifest as deep vein thrombosis, when occurring in the deep veins primarily of the legs and trunk, or as a pulmonary embolism, when the thrombus embolizes and obstructs the pulmonary arteries. Collectively, these events are known as venous thromboembolism (VTE), a life-threatening condition with an incidence of 1 to 2 events per 1000 person-years.[1–3] VTE is a complex disease with both environmental and genetic determinants. Family studies, candidate-gene approaches, and early genome-wide association studies (GWAS) primarily identified genetic risk factors in loci with well characterized effects on coagulation (*F2*, *F5*, *F11*, *FGG*, *ABO*, *SERPINC1*, *PROCR*, *PROC*, *PROS1*), supporting current therapeutic strategies that mainly target the coagulation cascade.[4–8] In recent years, larger GWAS meta-analyses revealed unanticipated loci, such as *SLC44A2*,[9] which was later characterized as a choline transporter involved in platelet activation,[10] and in the adhesion and activation of neutrophils.[11,12] Thus, genetic associations with VTE in larger and more diverse populations may uncover new biological pathways and molecular events contributing to the disease and potentially help identify novel targets for treatment. Most recently, 2 large efforts involving up to 30 000 VTE cases, led by the INVENT (International Network Against Venous Thrombosis) consortium[13] and the MVP (Million Veteran Program),[14] identified up to 43 genetic loci associated with VTE. To expand discovery of novel VTE risk loci, we conducted a large, cross-ancestry GWAS meta-analysis involving >80 000 VTE cases, along with a replication of novel loci and their characterization through downstream analyses.

METHODS

The data that support the findings of this study will be available through dbGaP (database of Genotypes and Phenotypes).

Design and Study Participants

The study design (Figure 1) included a cross-ancestry discovery meta-analysis of GWAS summary data from 4 consortia/studies (INVENT-2019, MVP, FinnGen, Estonian Genome Project) followed by a replication of discovery loci that exceeded the genome-wide significance threshold ($P<5.0\times10^{-8}$). The replication population involved 12 studies, limiting data to nonoverlapping studies with our discovery.[15] The combined discovery and replication data (when available) were then meta-analyzed, and ancestry-stratified meta-analyses were performed for African-ancestry (AA), European-ancestry (EA), and Hispanic-ancestry (HIS) participants to enable further downstream ancestry-specific analyses, such as fine mapping. Participants from studies provided written informed consent for use of their genetic and health information for analysis, and the studies were individually approved by the appropriate institutional review boards (Supplemental Material).

Study-Specific GWAS

Each study performed association analyses and provided summary data for meta-analysis. Genotyping arrays, imputation panels, and analyses performed by each participating study are detailed in Table S1. Additional study specifics are available as Supplemental Material.

Discovery, Replication, and Combined GWAS Meta-Analyses

All GWAS meta-analyses were conducted with METAL,[16] using a fixed-effects inverse-variance weighted model. All variants were included and there was no lower minor allele frequency (MAF) limit beyond study-specific minor allele count. Genome-wide significant variants ($P<5.00\times10^{-8}$) were kept if a concordant effect direction was observed in 2 or more studies and grouped into the same locus if they were within 1 Mb. We used the closest gene to the lead variant to refer to each locus, except at known loci where the causal gene has been previously identified and is different from the closest gene (such as *PROCR* or *PROS1*). We defined a locus as novel if a genetic association with VTE has not been previously observed in the region according to our review of peer-reviewed published reports.

Discovery Meta-Analysis

For the discovery cross-ancestry GWAS meta-analysis, we meta-analyzed data from 4 consortia/studies: INVENT-2019, MVP, FinnGen, and EGP. Participants were EA, AA, and HIS adult men and women with VTE (either deep vein thrombosis or pulmonary embolism cases) and controls. At each locus with a genome-wide significant signal, the lead variant was extracted and tested in an independent replication meta-analysis.

Replication

The replication GWAS meta-analysis consisted of the remaining 10 participating studies, as well as 2 external collaborators (GBMI[15] and 23andMe[17]). Replicating variants from the discovery were defined as those that had concordant effect direction in the discovery and the replication, and reached a Bonferroni-corrected *P* value threshold in the replication population corresponding to the number of variants tested for replication with a 1-sided hypothesis: *P* value threshold = [(0.05*2)/number of variants tested for replication] in the replication analysis.

Combined GWAS Meta-Analysis and Stratification by Ancestry

We performed a combined, cross-ancestry GWAS meta-analysis of discovery and replication data (when available) using participating studies with genome-wide summary data. We included variants with MAF≥0.01 to maintain adequate statistical power by reducing the number of low-powered tests because replication was not available. We estimated the heterogeneity associated with each variant using the Cochran Q test and the corresponding I^2 statistic. We assessed the genomic inflation with the lambda genomic control.[18] We report on variants exceeding the genome-wide

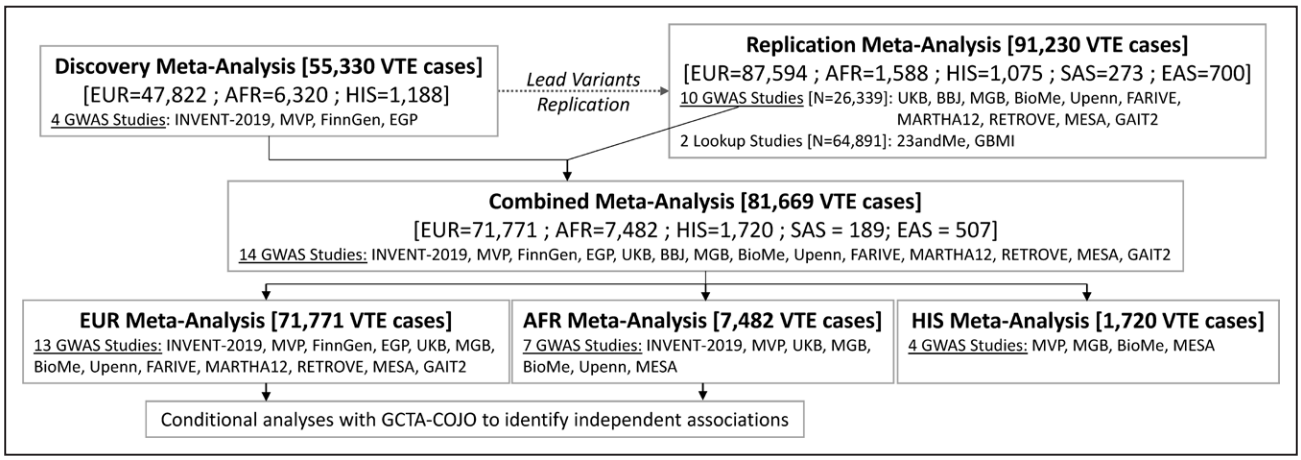

Figure 1. Analyses workflow.

Workflow of genetic analyses conducted for this study. AFR indicates African ancestry; BBJ, Biobank Japan; EAS, East Asian ancestry; EUR, European ancestry; GAIT2, Genetic Analysis of Idiopathic Thrombophilia; GBMI, Global Biobank Meta-Analysis Initiative; HIS, Hispanic ancestry; INVENT-2019, International Network Against Venous Thrombosis (2019 meta-analysis); MARTHA12, Marseille Thrombosis Association study of 2010-2012; MESA, Multi-Ethnic Study of Atherosclerosis; MGB, Mass General Brigham Biobank; MVP, Million Veteran Program; RETROVE, Riesgo de Enfermedad TROmboembólica VEnosa study; SAS, South Asian ancestry; UKB, UK Biobank; Upenn, Penn Medicine Biobank; and VTE, venous thromboembolism.

threshold ($P<5.00\times10^{-8}$) and view these as candidate novel loci associated with VTE and needing future replication.

We then stratified the analyses by ancestry and limited strata to EA, AA, and HIS because the remaining ancestries had too few VTE events to be informative. We estimated heterogeneity and genomic inflation; the linkage disequilibrium (LD) score intercept was computed for EA analysis, using the recommended Hapmap3 variants.[19] We report all additional ancestry-specific variants exceeding the genome-wide threshold ($P<5.00\times10^{-8}$) and view these as ancestry-specific candidate loci associated with VTE and needing future replication.

Ancestry-Stratified Analyses: Conditional Analyses and Fine-Mapping

To estimate the presence of independent signals, we performed conditional analyses with GCTA-COJO[20] at each locus with significant signals in EA, AA, and HIS GWAS meta-analyses. The TOPMed (Trans-Omics for Precision Medicine) ancestry-specific sequence data were used as reference panels.[21] Conditional analyses were performed at each locus, using a window that encompassed at least the genome-wide significant variants present in the locus with an additional buffer of ±100 kb. A stepwise joint regression model was used to identify secondary signals with joint P values $<5.00\times10^{-8}$ and a LD $r^2<0.2$ with selected variants. In addition, for each locus and for each ancestry-specific GWAS meta-analysis, we produced forest plots, Manhattan plots, and regional association plots to visually inspect the local genetic architecture (Figures S1–S9).[22,23] Additional information is found in the Supplemental Material.

Genetic Risk Score

We constructed an ancestry-specific genetic risk score (GRS) derived from the genome-wide significant lead variants observed in the EA meta-analysis and evaluated it for UK Biobank EA participants. The GRS for AA and HIS were not constructed because of a lack of availability of a large-scale dataset with accessible genotype data for other ancestries. The EA GRS was calculated for each individual as a summation of log(odds ratio [OR])-weighted genotypes. We then performed logistic regression to measure the association of the GRS with VTE status, while correcting for age, sex, and the top 10 genetic principal components. The predictive ability of the score was estimated by calculating the area under the curve (AUC), using the *pROC* R library.[24] Additional information is available in the Supplemental Material.

Transcriptome-Wide Association Studies

We performed a transcriptome-wide association study (TWAS) with the FUSION pipeline[25] using the EA meta-analysis results. We first performed a series of single-tissue TWAS using gene expression from expression quantitative trait loci (eQTL) datasets relevant to blood and thrombosis disorders: whole blood, peripheral blood, liver, lung, and spleen.[26–28] All associations reaching a Bonferroni-corrected significance threshold corresponding to the number of genes tested (N=14 219, $P<3.52\times10^{-6}$) were deemed statistically significant. Additional details are available in the Supplemental Material.

Protein QTL Mendelian Randomization

Using the combined, cross-ancestry VTE GWAS meta-analysis results, we performed a proteome Mendelian randomization (MR) analysis with high-confidence genomic instruments corresponding to protein QTL (pQTL) for 1216 circulating plasma proteins that passed consistency and pleiotropy filters, as previously described.[29] Additional information is available in the Supplemental Material. To account for multiple testing, associations passing the Bonferroni-corrected threshold (N=1256, $P<3.98\times10^{-5}$) were considered statistically significant.

Association of VTE Loci With Hemostasis and Hematology Traits

We conducted a series of in silico investigations involving hemostasis and hematology traits to better characterize the VTE-associated variants from the GWAS meta-analyses. To better understand if novel VTE-associated variants operate through hemostasis pathways, we extracted associations from published GWAS of 10 hemostatic traits: fibrinogen[30]; fibrin D-dimer[31]; coagulation factors VII,[32] VIII,[33] and XI[34]; von Willebrand factor (VWF) [33]; tissue plasminogen activator[35]; PAI-1 (plasminogen-activator inhibitor 1)[36]; activated partial thromboplastin time; and prothrombin time.[37] Because each variant association was investigated in 10 hemostasis traits, we set a P value threshold of 0.005 (0.05/10 traits tested for each lead variant of a locus) to separate associations of interest from other associations.

Similarly, we extracted associations with complete blood count (CBC) measures using summary data from nearly 750 000 individuals on 15 leukocyte, erythrocyte, and platelet traits.[38] Given the large sample size and high statistical power of these analyses, we used a more stringent threshold of interest that was a Bonferroni correction corresponding to the number of look-ups performed ($P<1.92\times10^{-5}$). We further performed colocalization analyses with the *coloc*[39] R library for significant associations, using the discovery, combined, EA, and AA VTE meta-analyses. Additional information is available in the Supplemental Material.

Phenome-Wide Association Testing

To explore associations between VTE-associated variants and other traits agnostically, we performed a phenome-wide association study (PheWAS) using the Medical Research Council Integrative Epidemiology Unit infrastructure and the associated *ieugwasr* R library.[40] Lead variants identified in our VTE meta-analyses were queried in 2 sources of GWAS (using the PheWAS codes "ukb-a" and "ukb-d"), which correspond to 1500 UKB analyses performed by the Neale laboratory (https://gwas.mrcieu.ac.uk/datasets/) on 337 000 individuals of British ancestry. We then retrieved associations reaching genome-wide significance ($P<5.00\times10^{-8}$) for each of the 1500 traits investigated.

RESULTS

Discovery Cross-Ancestry Meta-Analysis and Replication

The primary cross-ancestry discovery analysis included 55 330 participants among 3 ancestry groups with VTE

(47 822 EA, 6320 AA, and 1188 HIS) and 1 081 973 participants without VTE (918 195 EA, 118 144 AA, and 45 634 HIS). Over the 22 autosomal and X chromosomes, 35.5 million variants were analyzed, and the observed lambda was 1.06. We identified 10 493 variants reaching genome-wide significance, corresponding to 85 loci, of which 48 have not been identified in previous genetic studies of VTE (Table S2).

We tested lead variants from these 85 loci for replication in 91 230 cases and 3 322 939 controls from the independent replication data. After meta-analyzing the results of these 85 tests in the replication population, we identified 83 variants with a concordant effect direction between the discovery and the replication, of which 68 were replicated at the 1-sided Bonferroni-corrected significance threshold ($P<0.1/83=0.0012$; Table 1, Figure 2, Table S2). The successfully replicated signals corresponded to 34 known and 34 novel loci. Among the 34 novel loci that replicated, heterogeneity was minimal (heterogeneity $P>0.05$), ORs ranged between 0.84 to 0.98 and 1.03 to 1.18, and MAFs were all ≥0.021. The majority of variants were gene-centric (4 exonic, 16 intronic, and 3 in 3' or 5' untranslated regions or immediately downstream), 3 were linked to intronic noncoding RNA, and 8 were considered intergenic. Among the 17 variants and their associated loci that failed replication, 14 were novel and remain candidate loci that merit additional replication, whereas 3 were known loci.

Combined Cross-Ancestry GWAS Meta-Analysis and Ancestry-Stratified Results

Combined

The combined, cross-ancestry meta-analysis of the studies with genome-wide markers included 81 669 individuals with VTE and 1 426 717 individuals without VTE. We analyzed 19.1 million common variants (MAF≥0.01) and observed a lambda of 1.16, which is slightly elevated but expected for large-scale meta-analyses of polygenic traits.[41] We identified 16 550 variants reaching genome-wide significance in 111 loci, of which 41 were not observed in the discovery analysis (Table S3, Figure 2). Of these 41 additional loci, 1 corresponded to a common variant at the known *SERPINC1* locus (rs6695940) which encodes antithrombin, 4 were previously identified in the INVENT-2019[13] or MVP[14] meta-analyses at the *PEPD*, *ABCA5*, *MPHOSPH9*, and *ARID4A* loci, and 1 was a known pathogenic missense variant located in *SERPINA1* (rs28929474, p.Glu366Lys).[42] The remaining 35 loci were novel associations and are presented in Table 2. Among these 35 candidate loci, all had ORs with ranges of 0.93 to 0.97 and 1.03 to 1.15 and had a minimum MAF of 0.021. The majority of the variants were gene-centric (18 intronic and 3 in 3' untranslated regions), 3 were intronic in noncoding RNA, and 11 were considered intergenic.

European Ancestry

The EA meta-analysis, which included 71 771 participants with VTE and 1 059 740 participants without VTE, had a lambda of 1.22. Because population stratification might be introduced by founder effects in Finnish participants from FinnGen, we did a sensitivity analysis by removing this cohort, and observed a similar genomic factor of 1.19. We also observed an LD score intercept of 1.07, indicating an inflation mainly caused by polygenic architecture, and possibly slight residual stratification. Of the 11.1 million variants analyzed, 16 867 were genome-wide significant and clustered into 100 regions, of which 7 did not overlap with loci identified in the discovery or combined meta-analysis (Table 2, Figure 2, Table S4). For these 7 additional candidate loci, the ORs ranged from 0.94 to 0.97 and 1.04 to 1.07, and the minimum MAF was 0.058. Conditional analyses were performed at each of the 100 significant loci and revealed a subset of 21 loci with multiple independent signals (Table S5) and included 3 of the novel loci.

African Ancestry

The AA meta-analysis included 7482 participants with VTE and 129 975 participants without VTE from 7 cohorts and had a lambda of 1.05. Here, 17.1 million variants were analyzed, of which 752 were genome-wide significant and located within 13 loci, of which 2 corresponded to novel ancestry-specific signals at *RBFOX1* (OR=0.56; MAF=0.04) and *COL6A2* (OR=2.16; MAF=0.011; Table 2, Figure 2, Table S6). Conditional analyses were performed at each of the 13 significant loci and revealed 3 loci with additional independent signals (Table S7).

Hispanic Ancestry

The HIS meta-analysis included 1720 participants with VTE and 57 367 participants without VTE from 4 cohorts and had a lambda of 1.02. We analyzed 11.1 million variants, of which 58 were genome-wide significant, all located at the *ABO* locus with rs2519093 as lead variant (OR=1.49, MAF=0.15, $P=3.08\times10^{-15}$). The conditional analysis revealed a secondary signal at this locus (Table S7).

Comparison of Ancestry-Specific and Cross-Ancestry Meta-Analysis Results

We then investigated the lead variants from the AA and EA meta-analyses at the 11 loci (all known) identified in both analyses. At 5 loci, none of the AA lead variants were available in the EA analyses, because of their low frequency in EA (MAF<0.0006 for all 5 lead variants in non-Finnish Europeans according to gnomAD[43]). At the remaining 6 loci, the lead variants from the AA analysis were also genome-wide significant in the EA analysis, and shared similar effect sizes.

Table 1. Sixty-Eight Lead Variants From the Discovery That Meta-Analysis Successfully Replicated

rsID	CHR:POS:EA:NEA	EAF.Disc	OR.Disc	P.Disc	OR.Repl	P.Repl	Locus. Context	Locus.Gene
rs9442580	1:9339467:T:C	0.1551	1.06	1.83E-08	1.03	9.70E-05	Intergenic	*H6PD;SPSB1**
rs3767812	1:118155620:A:G	0.2437	1.05	9.64E-11	1.06	1.03E-20	Intronic	*TENT5C**
rs6025	1:169519049:T:C	0.0259	3.02	8.40E-811	3.59	9.29E-3103	Exonic	*F5* (p.Q534Q)
rs2842700	1:207282149:A:C	0.1092	1.11	5.95E-17	1.12	1.19E-25	Intronic	*C4BPA*
rs3811444	1:248039451:T:C	0.3324	0.96	5.70E-09	0.95	1.53E-20	Exonic	*TRIM58** (p.T374M)
rs7600986	2:68636923:A:T	0.2819	1.06	3.54E-12	1.05	9.18E-19	Intergenic	*PLEK;FBXO48*
rs182293241	2:128029746:A:G	0.0195	1.89	1.83E-27	1.55	0.0001063	Intronic	*ERCC3*
rs6719550	2:188272460:T:C	0.6639	1.04	7.56E-09	1.05	1.93E-17	Intronic	*CALCRL**
rs715	2:211543055:T:C	0.7022	0.95	3.51E-09	0.95	1.43E-17	UTR3	*CPS1**
rs13412535	2:224874874:A:G	0.2047	1.06	3.05E-10	1.08	1.10E-36	Intronic	*SERPINE2**
rs13084580	3:39188182:T:C	0.1076	1.09	2.89E-15	1.08	9.10E-22	Exonic	*CSRNP1* (p.G18S)
rs562281690	3:90177913:T:G	0.0024	2.01	6.45E-15	2.40	8.68E-31	Intergenic	*EPHA3;NONE*
rs62282204	3:138584405:T:C	0.5784	0.96	1.87E-08	0.98	6.73E-05	Intergenic	*PIK3CB;LINC01391**
rs7613621	3:169191186:A:G	0.4467	1.04	3.21E-09	1.03	5.33E-09	Intronic	*MECOM**
rs710446	3:186459927:T:C	0.5799	0.96	5.92E-11	0.96	1.41E-16	Exonic	*KNG1* (p.I581I)
rs6797948	3:194784705:T:C	0.7983	1.06	2.99E-11	1.05	7.59E-16	Intergenic	*LINC01968;XXYLT1**
rs6826579	4:83785031:T:C	0.7914	1.05	2.38E-08	1.03	2.44E-07	Intronic	*SEC31A**
rs17010957	4:86719165:T:C	0.8581	1.06	3.99E-09	1.05	1.00E-11	Intronic	*ARHGAP24**
rs2066864	4:155525695:A:G	0.2585	1.23	1.98E-172	1.23	1.94E-284	UTR3	*FGG*
rs3756011	4:187206249:A:C	0.3903	1.23	7.48E-198	1.24	9.26e-398	Intronic	*F11*
rs16867574	5:38708554:T:C	0.6673	0.95	2.78E-11	0.95	5.67E-16	ncRNA_intronic	*OSMR-AS1*
rs38032	5:96321887:T:C	0.6049	1.04	8.74E-09	1.03	1.49E-09	Intronic	*LNPEP**
rs9268881	6:32431606:A:T	0.5727	0.96	4.17E-10	0.97	6.73E-09	Intergenic	*HLA-DRA;HLA-DRB5**
rs145294670	6:34622561:A:AG	0.1385	1.06	6.11E-10	1.04	6.89E-06	Intronic	*ILRUN**
rs9390460	6:147694334:T:C	0.4957	0.95	2.49E-13	0.95	1.01E-20	Intronic	*STXBP5*
rs67694436	8:6654220:T:C	0.3486	0.96	3.94E-08	0.98	0.0001105	Intergenic	*AGPAT5;XKR5**
rs2685417	8:27807434:C:G	0.2562	1.06	1.57E-14	1.06	2.84E-25	Intronic	*SCARA5*
rs6993770	8:106581528:A:T	0.7142	1.08	4.48E-25	1.09	3.55E-48	Intronic	*ZFPM2*
rs35208412	9:99194509:A:AT	0.8298	1.09	1.56E-08	1.04	5.54E-06	Intergenic	*ZNF367;HABP4**
rs505922	9:136149229:T:C	0.6334	0.74	1.11E-425	0.69	1.55E-1043	Intronic	*ABO*
rs1887091	10:14535113:T:C	0.4936	0.96	4.77E-08	0.98	0.001107	Intergenic	*MIR1265;FAM107B**
rs17490626	10:71218646:C:G	0.1136	0.80	1.02E-79	0.80	3.23E-160	Intronic	*TSPAN15*
rs16937003	10:80938499:A:G	0.0287	1.15	1.07E-08	1.11	2.11E-11	Intronic	*ZMIZ1**
rs2274224	10:96039597:C:G	0.4414	1.04	2.55E-09	1.03	1.29E-10	Exonic	*PLCE1** (p.R1267P)
rs10886430	10:121010256:A:G	0.8897	0.89	7.34E-25	0.88	2.76E-64	Intronic	*GRK5*
rs11032074	11:32993887:A:G	0.7792	1.05	5.37E-09	1.03	3.24E-06	Intronic	*QSER1*
rs1799963	11:46761055:A:G	0.0136	2.05	2.19E-135	2.09	6.86E-420	UTR3	*F2*
rs141687379	11:56666822:A:G	0.9953	0.52	3.56E-31	0.64	1.06E-42	Intronic	*FADS2B*
rs174551	11:61573684:T:C	0.6583	1.07	1.65E-19	1.07	4.90E-35	Intronic	*FADS1*
rs35257264	11:126296816:T:C	0.0212	1.21	2.88E-14	1.18	2.28E-24	Intronic	*ST3GAL4**
rs1558519	12:6153738:A:G	0.6175	0.93	7.73E-24	0.92	1.42E-55	Intronic	*VWF*
rs7311483	12:9053661:T:C	0.3589	0.96	2.74E-09	0.97	2.73E-07	Intergenic	*A2ML1;PHC1**
rs6580981	12:54723028:A:G	0.5081	0.96	3.71E-09	0.95	2.26E-23	Intronic	*COPZ1**
rs3184504	12:111884608:T:C	0.4520	1.05	1.18E-11	1.04	3.30E-12	Exonic	*SH2B3** (p.T178T)

(*Continued*)

Table 1. Continued

rsID	CHR:POS:EA:NEA	EAF.Disc	OR.Disc	P.Disc	OR.Repl	P.Repl	Locus. Context	Locus.Gene
rs3211752	13:113787459:A:G	0.5527	0.95	1.69E-12	0.94	3.49E-25	Intronic	*F10*
rs57035593	14:92268096:T:C	0.3202	1.07	1.08E-20	1.07	2.64E-38	Intronic	*TC2N*
rs8013957	14:103140254:T:C	0.3699	1.04	5.33E-09	1.03	2.23E-07	Intronic	*RCOR1**
rs55707100	15:43820717:T:C	0.0270	0.87	2.90E-08	0.84	2.49E-27	Exonic	*MAP1A** (p.P2349L)
rs59442804	15:60899031:G:GAAAT	0.6438	0.96	4.67E-08	0.97	5.42E-10	ncRNA_intronic	*RORA-AS1**
rs12443808	16:30996871:C:G	0.4668	1.06	3.89E-14	1.03	1.85E-07	UTR5	*HSD3B7**
rs56943275	16:81898152:T:G	0.2446	1.08	4.15E-13	1.07	1.20E-26	Intronic	*PLCG2*
rs28634651	16:88553198:T:C	0.6191	1.06	9.20E-13	1.04	7.62E-14	Intronic	*ZFPM1**
rs6503222	17:1977862:A:G	0.6188	1.05	1.59E-12	1.04	5.21E-06	Intronic	*SMG6*
rs7225756	17:6893691:A:G	0.4877	0.96	3.57E-08	0.98	1.20E-06	ncRNA_intronic	*ALOX12-AS1**
rs62054822	17:43927708:A:G	0.8028	0.95	6.39E-09	0.95	7.11E-19	ncRNA_intronic	*MAPT-AS1**
rs142140545	17:64191540:CTATT:C	0.1169	0.93	2.27E-08	0.95	7.83E-07	Intergenic	*CEP112;APOH**
rs59277920	19:6077231:A:G	0.8210	0.94	1.47E-09	0.96	8.52E-06	Intronic	*RFX2**
rs8110055	19:10739143:A:C	0.2000	0.89	5.36E-44	0.89	6.50E-70	Intronic	*SLC44A2*
rs34783010	19:46180414:T:G	0.2132	0.95	3.25E-09	0.96	4.87E-10	Intronic	*GIPR**
rs1688264	19:49209560:T:G	0.5341	0.96	2.07E-10	0.96	3.02E-15	downstream	*FUT2**
rs1654425	19:55538980:T:C	0.1468	0.91	2.65E-18	0.94	4.21E-14	Exonic	*GP6* (p.S192S)
rs79388863	20:23168500:A:G	0.1521	0.92	1.74E-18	0.92	4.48E-27	Intergenic	*LINC00656;NXT1*
rs6060288	20:33772243:A:G	0.3417	1.12	8.19E-54	1.13	1.52E-102	Intronic	*MMP24-AS1-EDEM2*
rs4820093	22:33160208:T:C	0.2693	1.05	1.04E-08	1.04	5.39E-14	Intronic	*SYN3**
rs9611844	22:43115776:C:G	0.1286	1.10	2.09E-21	1.07	7.54E-20	Intronic	*A4GALT*
rs3002416	23:39710195:T:C	0.3638	0.95	2.20E-18	0.93	2.23E-23	Intergenic	*MIR1587;BCOR*
rs6048	23:138633280:A:G	0.7215	1.07	1.09E-25	1.08	1.59E-46	Exonic	*F9* (p.T156T)
rs2084408	23:154346709:T:G	0.3764	0.94	5.36E-19	0.94	6.27E-09	Intronic	*BRCC3*

Results from the discovery are in presented in columns suffixed with "Disc," whereas results from the replication are in columns suffixed with "Repl." CHR indicates chromosome; EA, effect allele; EAF, effect allele frequency; NEA, noneffect allele; OR, odds ratio; P, *P* value; and POS, position (hg19 build). *Indicates novel genetic associations.

Across the discovery, combined, EA, AA, and HIS meta-analyses, we identified 135 distinct loci (Figure 2). A summary of each locus, including LD patterns between lead variants from each meta-analysis as well as independent signals and association test results across all meta-analyses, is available in Table S8.

Genetic Risk Score

Using the 100 lead variants identified in the EA meta-analysis, we developed a GRS that was applied to independent UKB EA participants, which included 18516 cases and 92929 controls (Figure 3A and 3B). The GRS was significantly associated with VTE status (OR, 1.55 [95% CI, 1.53–1.58]), and the phenotypic variance explained by the score was estimated at 0.051. To assess the predictive ability of the score, we first calculated the AUC of the base model, which included the age, sex, and 10 genetic principal components, and obtained AUC_{base}=0.516 (95% CI, 0.511–0.520). After adding the GRS to the model, the AUC reached AUC_{GRS}=0.620 (95% CI, 0.616–0.625), an increase of Δ-AUC=0.104 over the base model. Compared with individuals with a score in the middle stratum (45%–55%), participants with a GRS in the top 1% had a significantly higher risk (OR, 6.07 [95% CI, 5.33–6.91]), whereas individuals in the bottom 1% had a significant risk reduction (OR, 0.52 [95% CI, 0.42–0.65]; Figure 3, Table S9).

Gene Prioritization With TWAS and pQTL MR

Transcriptome-Wide Association Study

Across the 6 single-tissue and 3 cross-tissues datasets analyzed, we identified 166 significant ($P<3.52\times10^{-6}$) and conditionally independent associations with a high posterior probability of colocalization (>0.75) between gene expression and VTE risk (Table S10). These associations involved 108 genes, of which 77 were mapped

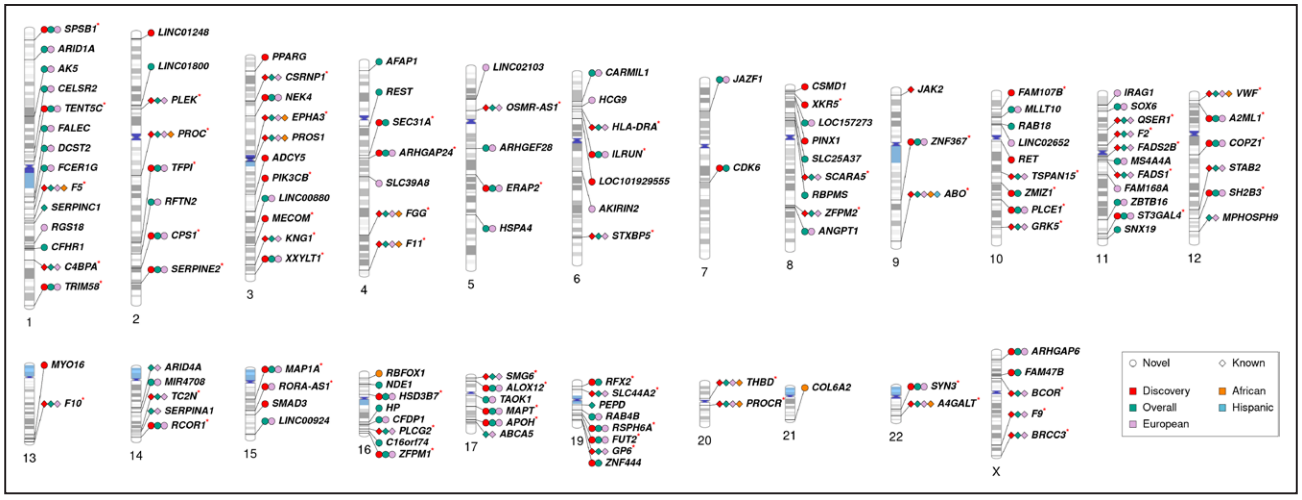

Figure 2. Genetic loci associated with VTE.

This figure presents the 135 loci significantly associated with VTE identified across all 5 meta-analyses: the discovery (in red), the overall meta-analysis (in green), and the analysis restricted to individuals of European ancestry (in purple), African ancestry (in orange), and Hispanic ancestry (in blue). Novel loci are represented with circles and known loci with diamonds. Loci with replication evidence are indicated with a red asterisk. VTE indicates venous thromboembolism.

to 46 genome-wide significant GWAS loci, leaving an additional 31 novel candidate genes that mapped outside of genome-wide significant GWAS loci (Table S11). At 33 GWAS loci, an associated gene matched the gene closest to the lead variant, supporting a role as a causal gene, whereas associated genes at the remaining 13 GWAS loci indicate genes for further investigation.

Protein QTL MR

We performed agnostic MR of 1216 plasma circulating pQTL using the combined VTE meta-analysis results and identified 23 proteins with a significant association ($P<3.98\times10^{-5}$, Figure 4, Table S12). For 13 proteins, the gene coordinates matched a genome-wide significant GWAS locus and included 5 of the novel GWAS loci.

Association of VTE-Associated Variants With Hemostasis and Hematology Traits

The association of any lead or conditionally independent variant at the 135 GWAS loci with hemostasis traits is presented in Figure 5A and Table S13. Across the 10 traits, we observed 83 signals shared with VTE. Among the 92 novel (replicated and candidate) loci reported above (see "Discovery Cross-Ancestry Meta-Analysis and Replication" and "Combined Cross-Ancestry GWAS Meta-Analysis and Ancestry-Stratified Results"), 18 (19%) had a variant associated with 1 or more of the 10 hemostasis traits (Figure S10A).

Next, we investigated associations of the 135 GWAS loci with hematology traits, presented in Figure 5B and Table S14. Across all 15 CBC measures, we identified 375 shared signals, and among the 92 novel loci, we observed at least 1 association at 55 (59%) novel (replicated and candidate) loci (Figure S10B).

Rates of colocalization with VTE signals (colocalized signals/total shared signals) were similar for hemostatic factors (48/83=58%, Figure S11A) and hematology traits (214/375=57%, Figure S11B). At shared loci, we also examined the effect directions of both VTE risk and the studied trait levels. For each hemostatic factor, the observed directions of effect were mostly consistent and agreed with our current biological knowledge, with the exception of factor VII, which shared 4 loci with the same effect direction than VTE, 4 with an opposite direction, and 1 with 2 independent variants that displayed the same direction for the first and an opposite direction for the second. Hematology traits displayed less consistent directions of effect with VTE across shared loci.

Phenome-Wide Association Studies

We performed a PheWAS of lead and conditionally independent variants at the 135 significantly associated loci across 1500 publicly available phenotypes involving EA UKB participants (Table S15). For each trait, only genome-wide significant variants were retrieved, and we restricted our analyses on traits sharing at least 10 loci with VTE (Figure 6, Table S16), which might indicate common biological pathways. Hematology traits, in particular platelet traits, shared the most loci with VTE (for example, 33 for platelet count), consistent with our observations from the larger CBC GWAS (N~750 000) sample (Figure 5B). Several traits correspond to height and weight measurements, as well as enzymes mainly produced by the liver (such as albumin, sex hormone–binding globulin, or insulin growth factor-1), and plasma lipid-related traits (apoA and apoB, high-density lipoprotein cholesterol, or triglycerides). Blood pressure (systolic and diastolic), glycated hemoglobin, calcium, cystatin C,

Table 2. Additional 44 Candidate Novel Loci Identified in the Overall, European, and African Ancestry Meta-Analyses

rsID	CHR:POS:EA:NEA	EAF	EFFECT	SE	OR	*P* value	Locus.Context	Locus.Gene
Novel loci identified in the overall meta-analysis								
rs551176418	1:27107263:T:TC	0.9248	0.0759	0.0132	1.08	9.61E-09	UTR3	*ARID1A*
rs6695572	1:77945635:A:G	0.1938	0.0424	0.0072	1.04	4.28E-09	Intronic	*AK5*
rs3832016	1:109818158:CT:C	0.7627	−0.0449	0.0066	0.96	8.95E-12	UTR3	*CELSR2*
rs1267881263	1:150496127:CA:C	0.5468	0.0426	0.0076	1.04	2.36E-08	Intergenic	*FALEC;ADAMTSL4*
rs905938	1:154991389:T:C	0.7448	−0.0346	0.0063	0.97	3.70E-08	Intronic	*DCST2*
rs3557	1:161188893:T:G	0.9182	0.0654	0.0106	1.07	7.70E-10	UTR3	*FCER1G*
rs143410348	1:196809316:T:TAA	0.5434	0.0415	0.0074	1.04	2.44E-08	Intergenic	*CFHR1;CFHR4*
rs78475244	2:65086804:T:C	0.0542	−0.0713	0.0128	0.93	2.52E-08	ncRNA_intronic	*LINC01800*
rs78872368	2:198545250:C:G	0.1919	−0.0412	0.0071	0.96	7.27E-09	Intergenic	*RFTN2;MARS2*
rs900399	3:156798732:A:G	0.6205	0.0382	0.0060	1.04	1.46E-10	Intergenic	*LINC02029;LINC00880*
rs9654093	4:7903763:C:G	0.1504	0.0492	0.0081	1.05	1.03E-09	Intronic	*AFAP1*
rs781656	4:57778645:A:G	0.1963	0.0389	0.0070	1.04	2.26E-08	Intronic	*REST*
rs7730244	5:72957088:T:C	0.5245	−0.0328	0.0057	0.97	1.04E-08	Intronic	*ARHGEF28*
rs147133967	5:132426851:G:GTT	0.0810	−0.0659	0.0110	0.94	2.43E-09	Intronic	*HSPA4*
rs214059	6:25536937:T:C	0.4331	0.0357	0.0055	1.04	1.01E-10	Intronic	*CARMIL1*
rs2394251	6:29943688:G:C	0.7331	−0.0405	0.0063	0.96	1.43E-10	ncRNA_intronic	*HCG9*
rs1513275	7:28259233:T:C	0.7453	0.0449	0.0070	1.05	1.40E-10	ncRNA_intronic	*JAZF1-AS1*
rs10099512	8:9178821:C:G	0.1105	0.0608	0.0105	1.06	6.98E-09	Intergenic	*LOC101929128;LOC157273*
rs2048528	8:23373680:A:G	0.3089	−0.0347	0.0060	0.97	5.77E-09	Intergenic	*ENTPD4;SLC25A37*
rs2915595	8:30402817:A:G	0.2391	0.0365	0.0065	1.04	2.52E-08	Intronic	*RBPMS*
rs4236786	8:108291878:C:G	0.2492	0.0353	0.0064	1.04	3.93E-08	Intronic	*ANGPT1*
rs1243187	10:21907016:T:C	0.6920	−0.0341	0.0061	0.97	2.53E-08	Intronic	*MLLT10*
rs4272700	10:27881308:A:T	0.2726	0.0395	0.0064	1.04	7.75E-10	Intergenic	*RAB18;MKX*
rs2030291	11:16251251:A:T	0.6077	−0.0325	0.0056	0.97	8.19E-09	Intronic	*SOX6*
rs4354705	11:60088159:C:G	0.3635	0.0315	0.0058	1.03	4.83E-08	Intergenic	*MS4A4A;MS4A6E*
rs2846027	11:114003415:T:C	0.3112	−0.0344	0.0061	0.97	1.42E-08	Intronic	*ZBTB16*
rs7107568	11:130779668:T:C	0.5610	−0.0303	0.0056	0.97	4.71E-08	Intronic	*SNX19*
rs2127869	14:65794352:T:C	0.3350	−0.0340	0.0062	0.97	4.68E-08	Intergenic	*LINC02324;MIR4708*
rs7183672	15:96101018:A:G	0.6432	−0.0358	0.0062	0.96	7.34E-09	Intergenic	*LINC00924;LOC105369212*
rs71376077	16:15738114:C:G	0.9728	0.1408	0.0249	1.15	1.57E-08	Intronic	*NDE1*
rs7197453	16:72079127:C:G	0.3572	0.0315	0.0057	1.03	3.19E-08	Intergenic	*DHODH;HP*
rs77246010	16:75429853:T:C	0.4489	0.0408	0.0069	1.04	4.12E-09	Intronic	*CFDP1*
rs8049403	16:85778651:A:G	0.0214	0.1365	0.0248	1.15	3.91E-08	Intronic	*C16orf74*
rs71138827	17:27833678:A:AGATT	0.4288	0.0336	0.0058	1.03	5.89E-09	Intronic	*TAOK1*
rs2545774	19:41287674:T:C	0.2528	−0.0378	0.0065	0.96	6.80E-09	Intronic	*RAB4B*
Additional novel loci identified in the European meta-analysis								
rs4540639	1:192104320:C:G	0.4675	0.0346	0.0060	1.04	6.88E-09	Intergenic	*LINC01680;RGS18*
rs35225200	4:103146888:A:C	0.9190	−0.0645	0.0115	0.94	1.89E-08	Intergenic	*BANK1;SLC39A8*
rs112367053	5:28379046:T:G	0.6662	0.0586	0.0107	1.06	4.07E-08	Intergenic	*LINC02103;LSP1P3*
rs2754251	6:88385949:A:G	0.0584	0.0715	0.0129	1.07	2.65E-08	Intronic	*AKIRIN2*
rs10763665	10:28771491:C:G	0.5783	−0.0342	0.0062	0.97	3.13E-08	ncRNA_intronic	*LINC02652*
rs7122100	11:10732560:A:C	0.2411	0.0410	0.0075	1.04	4.93E-08	Intergenic	*IRAG1;CTR9*
rs1145656	11:73305859:A:C	0.8171	−0.0442	0.0079	1.05	2.00E-08	upstream	*FAM168A*
Additional novel loci identified in the African meta-analysis								
rs76668186	16:6686083:A:T	0.9597	−0.5776	0.1056	0.56	4.52E-08	Intronic	*RBFOX1*
rs114102448	21:47523605:A:G	0.0114	0.9527	0.1725	2.60	4.11E-08	Intronic	*COL6A2*

CHR indicates chromosome; EA, effect allele; EAF, effect allele frequency; NEA, noneffect allele; OR, odds ratio; POS, position (hg19 build); and SE, standard error of effect.

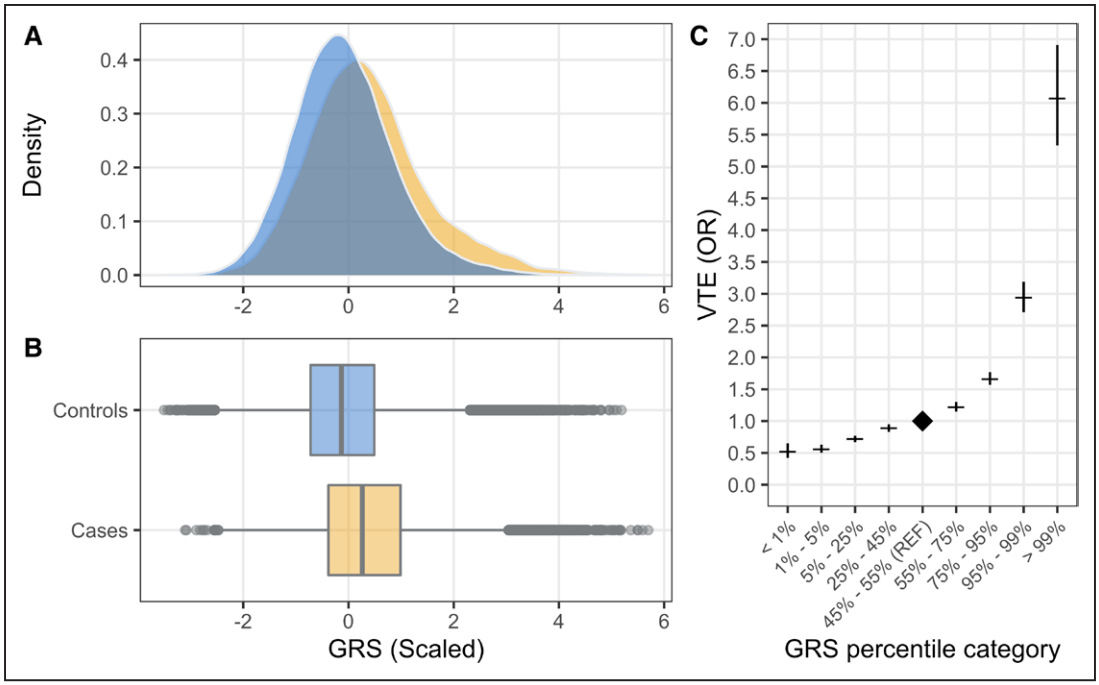

Figure 3. GRS analysis.
Distribution of the GRS in VTE cases (in green) and controls (in purple) as a density plot (**A**) and a boxplot (**B**). **C**, Presentation of the VTE risk as ORs and associated 95% CIs (*y* axis) for different percentiles ranges of the GRS score (*x* axis) relative to the middle range (45%–55%). GRS indicates genetic risk score; OR, odds ratio; REF, reference; and VTE, venous thromboembolism.

and C-reactive protein levels were among additional traits sharing at least 10 loci with VTE. Few traits had a consistent direction of effect with respect to VTE risk across shared loci (Figure 6). For example, out of 10 loci shared between bilirubin levels and VTE, 9 (90%) were associated with an increase of both bilirubin levels and VTE risk. For albumin levels, glycated hemoglobin, and systolic blood pressure, an opposite direction of effect between these traits and VTE risk was observed at >75% of shared loci.

DISCUSSION

We identified 135 independent genomic loci and 39 additional genes from TWAS and pQTL associated with an increased or decreased risk of VTE. This reflects a substantial increase in the number of validated and candidate loci for VTE risk beyond past genetic mapping efforts.[13,14] Although the novel VTE associated variants were typically noncoding and displayed small effect sizes, they may provide valuable insights into genetic loci not previously suspected to play a role in VTE. Our results highlight genetic variation across the rare-to-common allele frequency spectrum in multiple ancestry groups and add new evidence of biologic predictors of VTE pathogenesis for further investigation. The in silico interrogations provide valuable clues about the putative causal gene at each locus and additional insights to biological pathways shared with VTE.

Biological Insights

Novel Replicated Loci

Our strongest evidence supports 34 loci with novel VTE associations. Except for *TFPI* and *SERPINE2*, the novel genetic loci were not in established VTE pathophysiology pathways. A subset of these 34 loci (12 loci, 35%) was associated with plasma levels of the hemostasis traits interrogated, and most (26 loci, 76%) were associated with a hematology trait. This contrast should be interpreted with caution because statistical power for the hemostasis traits was much smaller than for the hematology traits.

Although most of the novel associations reported had an OR in the range of 0.90 to 0.98 and 1.03 to 1.10, we were able to identify and replicate 3 uncommon variants with larger estimated effects: an intronic variant (MAF=0.021) in the glycosyltransferase *ST3GAL4* ($OR_{discovery}$, OR=1.21, $OR_{replication}$=1.18), which was also associated with increased VWF and factor VIII levels; an intronic variant (MAF=0.029) in the transcriptional coactivator *ZMIZ1* ($OR_{discovery}$=1.15, $OR_{replication}$=1.11); and an exonic variant (MAF=0.027) in *MAP1A* (p.Pro2349Leu, $OR_{discovery}$=0.87, $OR_{replication}$=0.84), which was also associated with decreased levels of VWF and fibrinogen, and had a protective effect against VTE.

Variants associated with hemostasis traits provide clues that the causal gene at these loci might directly or indirectly perturb the coagulation cascade. For instance, *XXYLT1* encodes a xylosyltransferase

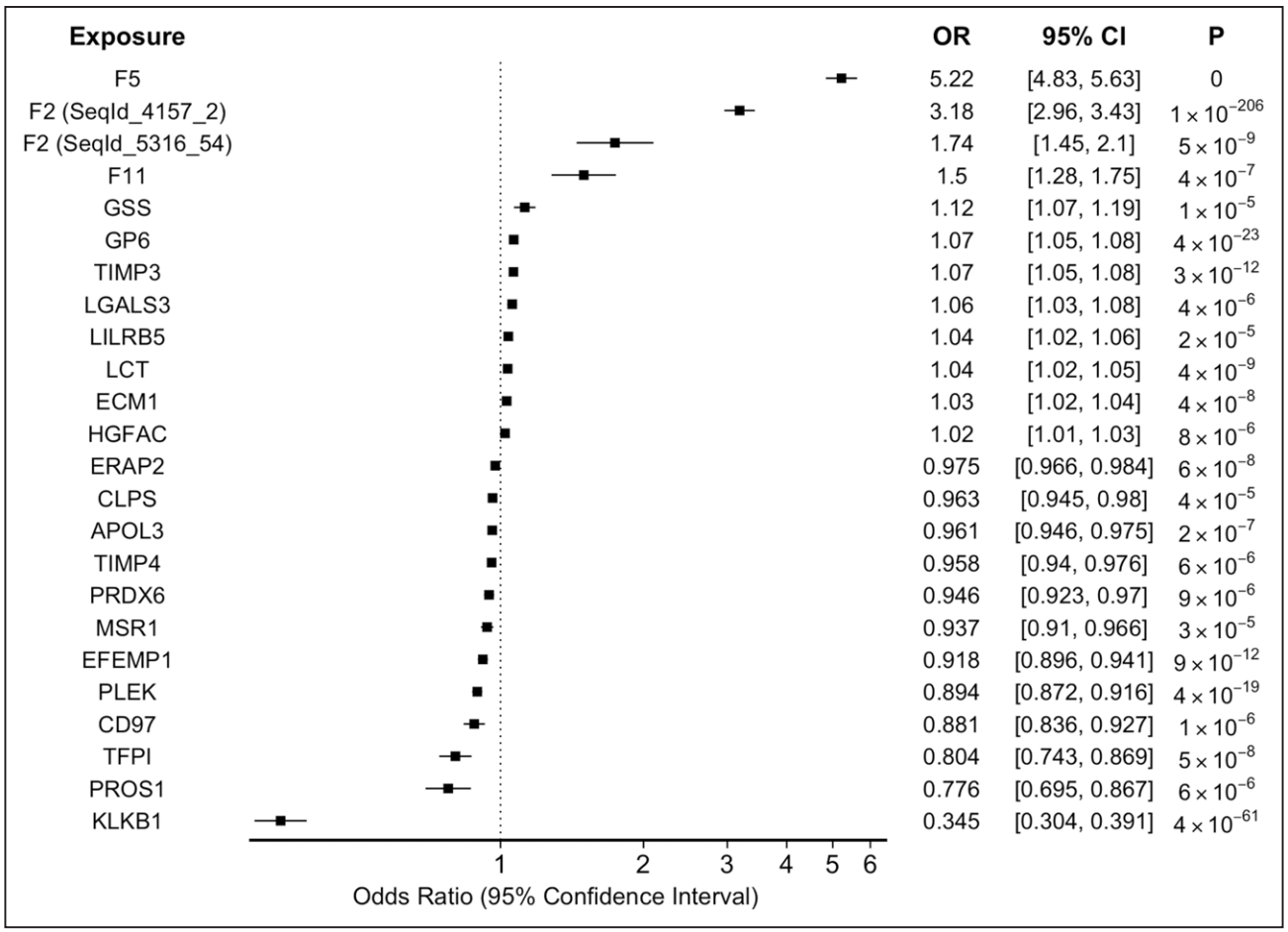

Figure 4. Significant associations of protein quantitative trait loci Mendelian randomization.
Twenty-three proteins significantly associated with venous thromboembolism, of 1216 plasma protein analyzed, using the combined venous thromboembolism summary statistics. OR indicates odds ratio.

known to interact with coagulation factors[44] and had a nearby variant ($OR_{discovery}$=1.06, $OR_{replication}$=1.06) also associated with decreased factor VII levels. Another example is *FUT2*, a fucosyltransferase gene with a downstream variant ($OR_{discovery}$=0.96, $OR_{replication}$=0.96) that was also associated with decreased VWF levels. In addition, some variants were associated with several hematology traits, suggesting common genetic regulatory pathways affecting hematopoiesis, such as the replicated *RCOR1* signal on chromosome 14, and the candidate gene *REST* on chromosome 4 identified in the combined meta-analysis, 2 genes that form the transcriptional repressor CoREST, known to mediate hematopoiesis.[45]

Among the 34 loci, 17 (50%) had TWAS evidence linking transcript expression with a gene in the locus, and 3 were linked to protein measures. These results may help to prioritize biologically relevant genes for further investigations. At the *COPZ1* locus, the lead variant was associated with several CBC measures, including platelet count and red blood cell count, and the TWAS revealed an association with *NFE2*, known to regulate erythroid and megakaryocyte maturation.[46]

Other Replicated and Nonreplicated Loci

Replicated variants included 2 rare variants at the known *EPHA3* (intergenic, MAF=0.0024, OR=2.40) and *FADS2B* (intronic, MAF=0.0047, OR=0.64) loci. Among the 17 failed replications, 7 reached nominal significance ($P<0.05$), suggesting that these variants might need a larger replication sample to be validated. See the Supplemental Material for more details.

Novel Candidate Loci

Across the multiple interrogation approaches, we identified several scores of candidate loci with evidence to support their association with VTE, although not yet replicated. This included 35 candidates from the combined GWAS, 7 candidates from the EA GWAS, and 2 candidates from the AA GWAS. Interestingly, the 2 variants (MAF 0.04 and 0.011) in the AA population were not present in EA participants and were associated with nearly 2-fold changes in risk of VTE. However, these 2 variants were detected in only a subset of studies, which included only 882 AA VTE cases out of 7482, warranting additional investigations to confirm these 2 signals in *RBFOX1* (an RNA-binding protein)

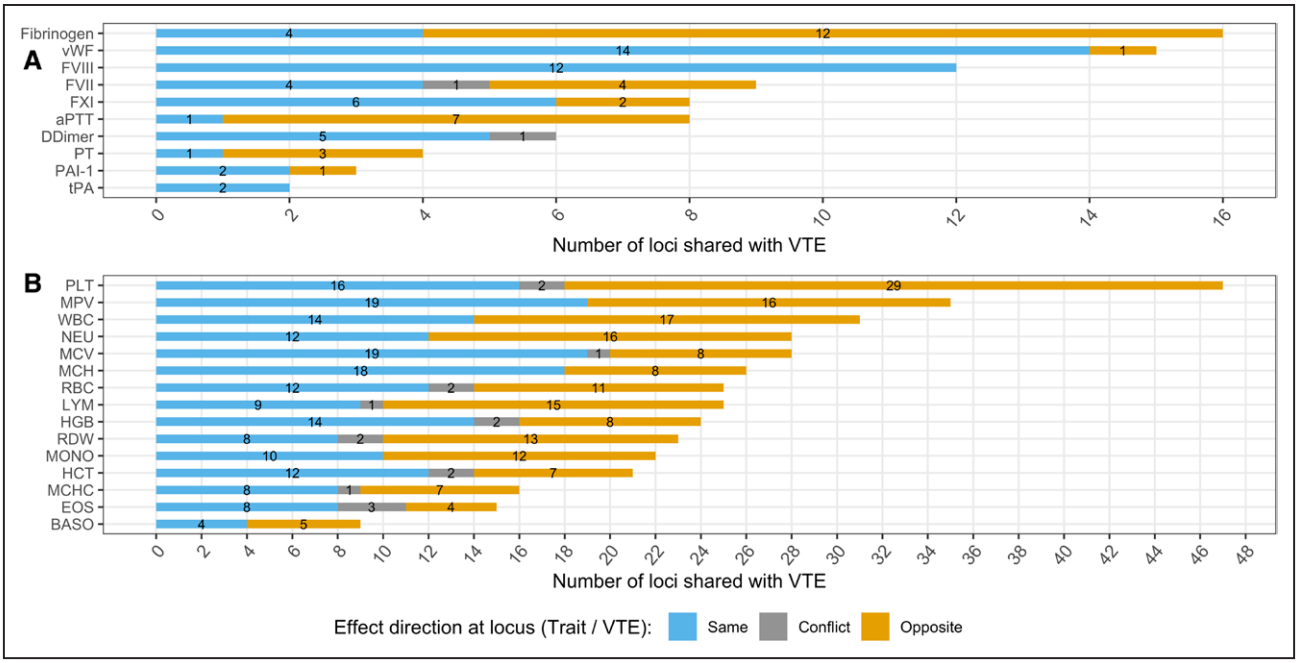

Figure 5. VTE genetic loci shared with hemostatic factors and blood traits.
A, Number of VTE loci shared with each of the 10 hemostatic factors investigated. Loci with shared variants that had an opposite effect direction between the trait and VTE are indicated in orange, whereas those that had the same effect direction are presented in blue. Loci with multiple independent shared variants and conflictual effect directions are indicated in gray. **B**, The same analysis with complete blood count traits: platelet count (PLT), mean platelet volume (MPV), red blood cell count (RBC), mean corpuscular volume (MCV), hematocrit (HCT), mean corpuscular hemoglobin (MCH), MCH concentration (MCHC), hemoglobin concentration (HGB), red cell distribution width (RDW), white blood cell count (WBC), monocyte count (MONO), neutrophil count (NEU), eosinophil count (EOS), basophil count (BASO), and lymphocyte count (LYM). aPTT indicates activated partial thromboplastin time; PT, prothrombin time; and VTE, venous thromboembolism.

and *COL6A2* (a collagen-generating gene that contains several domains similar to *VWF* type A domains). For the remaining candidate GWAS loci, we saw attributes and associations similar to those with the replicated loci. With additional replication resources in the future, these candidates may become fully replicated genetic associations.

In addition, the conditional analyses revealed independently associated variants mapping to distinct genes that may be of interest for further investigations, such as *BRD3* at the *ABO* locus, a chromatin reader known to associate with the hematopoietic transcription factor *GATA1*.[47] At the *EPHA3* locus, we also noted that the lead GWAS variant and the conditionally independent variant mapped upstream and downstream of *PROS2P*, a protein S pseudogene that might be of interest.

At these candidate loci, genes prioritized by the TWAS may also provide putative genes at these loci. For example, *ZBTB7B*, a zinc-finger protein that represses the expression of extracellular matrix genes such as fibronectin and collagen,[48] was identified by TWAS at the GWAS candidate locus *DCST2*. The 31 candidate genes identified in the TWAS as well as the additional 8 from the pQTL MR analyses, although lacking a significant genetic association at these loci, might indicate relevant genes for future investigations. For instance, SYK is a critical platelet-activation protein, and tyrosine kinase inhibitors of SYK have been explored for platelet inhibition.[49,50]

Clinical Implications

The GRS provided VTE risk discrimination in our EA population, and those at the extremes of the score distribution experienced multifold risk differences. We were not able to integrate or to compare nongenetic risk factors with the GRS.

Current anticoagulation therapy to prevent or treat VTE operates through the modulation of proteins produced in the liver (coumarin-based therapies) or through direct inhibition of coagulation factors IIa (thrombin) and Xa. Although the safety profile of anticoagulation treatments has evolved, bleeding remains a life-threatening off-target outcome. New approaches to preventing thrombosis while minimizing bleeds are in development, including a focus on contact (intrinsic) pathway proteins factor XI, factor XII, prekallikrein, and high-molecular-weight kininogen.[51] Agnostic interrogations such as these may lead to discovery of novel proteins that "break the inexorable link between antithrombotic therapy and bleeding risk."[52]

The hematology traits investigations and the PheWAS established that CBC measures share a large number of loci with VTE, and platelet phenotypes in particular are

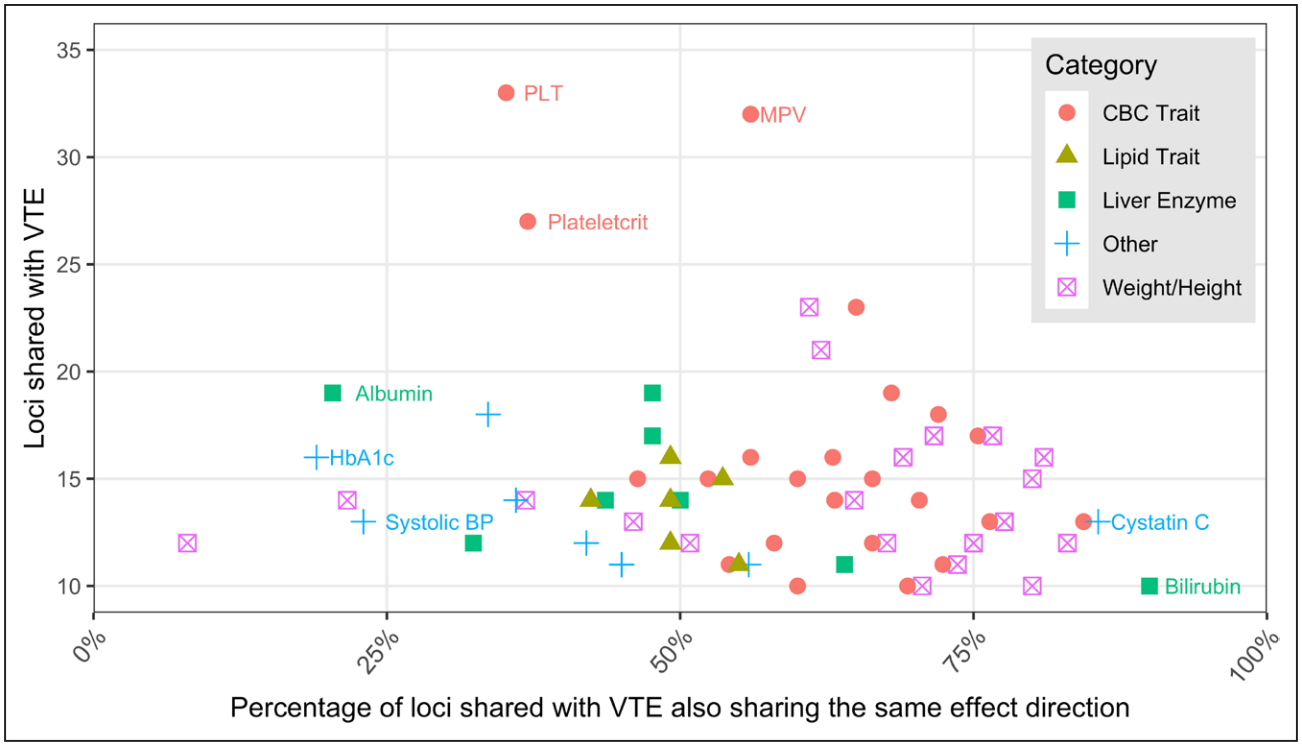

Figure 6. PheWAS traits sharing at least 10 loci with VTE.
The PheWAS traits sharing at least 10 loci with VTE. Shape and color represent 1 of 5 categories: complete blood count (CBC) traits, lipid traits, liver enzyme, height and weight traits, or other (if the trait did not fit in one of the aforementioned categories). The *x* axis indicates the number of loci shared between VTE and the PheWAS trait, whereas the *y* axis indicates the proportion of loci where the direction of effect was the same between the PheWAS trait and VTE. As a result, traits close to 100% have the same direction of effect as VTE at most shared loci, whereas traits close to 0% have an opposite direction than VTE at most shared loci. BP indicates blood pressure; HbA1c, glycated hemoglobin; MPV, mean platelet volume; PheWAS, phenome-wide association study; PLT, platelet count; and VTE, venous thromboembolism.

the most frequent traits shared with VTE variants: 51 loci were associated with platelet count, mean platelet volume, plateletcrit, or platelet distribution width in the PheWAS, and 35 of these loci are novel, which represents more than a third of all novel genetic associations. Several loci associated with VTE harbor genes with known roles in hematopoiesis and megakaryocyte development, or platelet turnover,[45,46,53–60] or platelet aggregation (Supplemental Material).[10,61–71] Altered platelet generation, turnover, or reactivity may be a feature of VTE pathogenesis. For one, past prospective studies[72] and case-control studies[73,74] suggest that enlarged platelets, as measured by mean platelet volume, are associated with VTE and VTE outcomes. Studies of platelet function measures with VTE have been less conclusive, which may relate to the limitations of these studies in assessing comprehensive and standardized platelet reactivity mechanisms.[75–77] Collectively, these results suggest that treatments inhibiting platelet activation such as aspirin might be beneficial in the prevention of VTE, although previous studies and trials on aspirin and combinations with anticoagulants offered mixed results.[78] Different antiplatelets, such as more targeted thrombin, PAR1 or PAR4 inhibitors, or intracellular PDE platelet signaling inhibitors like cilostazol, could be worthwhile for further study in VTE prevention.

Strengths and Limitations

The major strength of this genetic discovery effort is the large sample size of the populations contributing to the genetic variation interrogations. We increased statistical power compared with previous VTE GWAS meta-analysis efforts and increased our ability to detect new associations, many of which were replicated, and less common genetic variations. The cross-ancestry meta-analyses also increased discovery potential where allele frequencies were more common in some populations compared with others.

Several limitations deserve mention. Case ascertainment varied by study, and some studies provided validated VTE events whereas others relied on information from electronic health records. Further, some studies included only hospitalized VTE events and did not capture events in the outpatient setting. These differences may have introduced some bias if case ascertainment and hospitalization status have genetic determinants. We included all VTE cases and did not stratify by provoked status to increase statistical power. Many of the studies had not classified the VTE events as provoked and unprovoked. In addition, although the cross-ancestry approach provided benefits, the numbers of VTE cases were not evenly distributed by ancestry, thus reducing

our ability to detect ancestry-specific VTE variants in the underrepresented ancestry groups with more modest case counts. Because of the diversity of imputation panels used by the participating studies, genetic variants had variable coverage across studies, which weakened our power to detect associations. Another limitation of our approach that used summary GWAS statistics from meta-analyses is the absence of participant-specific genotype-level information. This required us to rely on LD information extracted from external datasets, which can result in variants being missed and LD patterns not accurately captured. This may have introduced some bias in analyses that relied on LD, such as the conditional analyses and the TWAS. Further, in silico work was performed using external datasets such as the hemostatic factors and hematology traits summary statistics, where the size (and statistical power) of the datasets varied greatly. Although different significance thresholds were used for significance, this may have biased the detection of significant associations to those traits that had large sample sizes. In addition, the pQTL MR analyses relied in some cases on a single genetic instrument, such as the *KLKB1* analysis, and these results should be considered hypothesis-generating.

Conclusions

These cross-ancestry GWAS meta-analyzes identified 34 loci that replicated discovery findings. Some of the novel loci may contribute to VTE through well-characterized coagulation pathways, whereas others provide new data on the role of hematology traits, particularly platelet function. Many of the replicated loci are outside of known or currently hypothesized pathways to thrombosis. We also provided a list of 44 new candidate loci including candidates from the combined cross-ancestry GWAS, from the EA GWAS, from the AA GWAS, and also 39 candidate genes from the TWAS and pQTL MR. These findings highlight new pathways to thrombosis and provide novel molecules that may be useful in the development of antithrombosis treatment that reduces bleeding adverse occurrences.

ARTICLE INFORMATION

Received March 1, 2022; accepted August 9, 2022.

Affiliations

Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, Framingham, MA (F.T., M.-H.C., A.D.J.). Framingham Heart Study, Boston University, Framingham, MA (F.T., M.-H.C., D. Ko, A.D.J.). Division of Vascular Surgery, Stanford University School of Medicine, CA (D. Klarin). Veterans Affairs Palo Alto Healthcare System, CA (D. Klarin). Cardiovascular Health Research Unit, Department of Medicine (J.A.B., K.L.W., B.M.P.), Department of Health Systems and Population Heath (B.M.P.), University of Washington, Seattle. Departments of Medicine (M.G.L., D.J.R), Surgery (R.L.J., S.M.D), Psychiatry (R.L.K.), and Genetics (D.J.R , S.M.D.), Perelman School of Medicine, University of Pennsylvania, Philadelphia. Division of Preventive Medicine (D.I.C., F.G., P.M.R.), Division of Pulmonary and Critical Care Medicine (K.R.C.), Brigham and Women's Hospital, Boston, MA. Department of Medicine (P.N., T.G., K.R.C., C.J.O.), Emergency Medicine (C. Kabrhel), Harvard Medical School (D.I.C., P.M.R.), Boston, MA. Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN (E.L.G., M.D.A., S.M.A.). HUNT Research Center, Department of Public Health and Nursing, Norwegian University of Science and Technology, Levanger (K. Hveem). K.G. Jebsen Centre for Genetic Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim (K. Hveem, T.H.N., B.B., A.H.S.). Institute of Genomics, University of Tartu, Estonia (M.T.-L.). Genomics of Complex Disease Unit, Institut d'Investigació Biomèdica Sant Pau, Barcelona, Spain (A.M.-P., J.M.S., J.C.S., M.S.-L.). Bordeaux Population Health Research Center, University of Bordeaux, France (D.A., M.G., D.-A.T.). UMR1219, Institut national de la santé et de la recherche médicale, Bordeaux, France (D.A., M.G., D.-A.T.). Centre National de Recherche en Génomique Humaine, CEA, Université Paris-Saclay, Evry, France (D.D.-B., A.B., R.O., J.-F.D.). Laboratory of Excellence on Medical Genomics, GenMed, Paris, France (D.D.-B., M.G., A.B., R.O., J.-F.D., D.-A.T.). Laboratory for Cardiovascular Genomics and Informatics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan (K.I., S.K.). Cardiovascular Research Center (P.N., S.K., T.G.), Analytic and Translational Genetics Unit, Department of Medicine (W.Z.), Emergency Medicine (C. Kabrhel), Massachusetts General Hospital, Boston. Program in Medical and Population Genetics and the Cardiovascular Disease Initiative, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge (P.N., S.K., T.G., K.R.C.). Division of Epidemiology and Community Health, School of Public Health (P.L.L., J.S.P., W.T.), Division of Biostatistics (J.W.P.), Department of Laboratory Medicine and Pathology (N.P.), University of Minnesota, Minneapolis. Charles Bronfman Institute for Personalized Medicine (G.N.K., R.D., R.J.F.L., H.M.T.V.), Department of Genetics and Genomic Sciences (G.N.K., R.D., H.M.T.V.), Division of Nephrology, Department of Medicine (G.N.K.), BioMe Phenomics Center (R.D.), Icahn School of Medicine at Mount Sinai, New York, NY. Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, University of Texas Health Science Center at Houston (P.S.d.V.). 23andMe, Inc, Sunnyvale, CA (G.C.-P., X.W.). Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor (B.N.W., C.J.W.). Center for Innovative Design & Analysis and Department of Biostatistics & Informatics, Colorado School of Public Health, Aurora (J.W.P.). Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA (C. Kooperberg, J.H., A.P.R.). Thrombosis Research Center, UiT—Arctic University of Norway, Tromsø (S.K.B., K. Hindberg, V.M.M., J.-B.H.). Division of Internal Medicine, University Hospital of North Norway, Tromsø (S.K.B., V.M.M., J.-B.H.). Clinical Epidemiology, Leiden University Medical Center, Netherlands (R.L.-G., F.R.R., A.v.H.V.). Hematology Laboratory, La Timone University Hospital of Marseille, France (N.S., P.S., P.-E.M.). Department of Epidemiology, Harvard T.H. Chan Harvard School of Public Health, Boston, MA (C.S., C.T., P.K.). Section of Cardiovascular Medicine, Boston University School of Medicine, MA (D.K.). Cardiology Section, Department of Medicine (C.J.O.), MAVERIC (J.E.H.), Veterans Affairs Boston Healthcare System, MA. Institute for Translational Genomics and Population Sciences, Lundquist Institute for Biomedical Innovation, Torrance, CA (K.D.T.). Integrative Epidemiology of Cardiovascular Diseases, Université Paris Cité, Paris Cardiovascular Research Center, France (J.-P.E.). Department of Cardiology, Assistance Publique – Hôpitaux de Paris, Hopital Européen Georges Pompidou, Paris, France (J.-P.E.). Unit of Thrombosis and Hemostasis, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain (J.C.S.). Cardiovascular Health Research Unit, Departments of Biostatistics and Medicine (T.M.B.), Department of Epidemiology (S.R.H., A.P.R., B.M.P., N.L.S.), University of Washington, Seattle. Veterans Affairs Informatics & Computing Infrastructure, Veterans Affairs Salt Lake City Healthcare System, UT (J.A.L., K.M.L.). Epidemiology, University of Utah, Salt Lake City (J.A.L.). Hematology Department and Biosurgical Research Laboratory (Carpentier Foundation), European Georges Pompidou Hospital, Assistance Publique Hôpitaux de Paris, France (D.M.S.). Innovative Therapies in Haemostasis, Institut national de la santé et de la recherche médicale (D.M.S.), Department of Vascular Medicine, Paris Saint-Joseph Hospital Group (J.E.), Université de Paris, France. Integrative Epidemiology of Cardiovascular Diseases (X.J.), Paris Cardiovascular Research Center, Inserm U970, Institut national de la santé et de la recherche médicale (X.J.), Université Paris Descartes, Sorbonne Paris Cité, France. Department of Cardiovascular Medicine, University of Tokyo, Japan (I.K.). Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet, Center for Molecular Medicine, Stockholm, Sweden (M.S.-L.). Centre D'Etude du Polymorphisme Humain, Fondation Jean Dausset, Paris, France (J.-F.D.). UMR1153, Institut national de la santé et de la recherche médicale CRESS, Paris, France (J.E.). C2VN, Institut national de la santé et de la recherche médicale, institut national de recherche pour l'agriculture, l'alimentation et l'environnement, Aix-Marseille University, France (P.S., P.-E.M.). Center for Public Health Genomics, University of Virginia, Charlottesville (S.S.R.). College of Medicine, Ohio State University, Columbus (R.D.J.). Corporal Michael J. Crescenz Philadelphia Veterans Affairs Medical Center, PA (S.M.D.). Kaiser Permanente Washington Health Research Institute, Seattle (N.L.S.). Seattle Epidemiologic Research and Information Center, Department of Veterans Affairs Office of Research and Development, Seattle, WA (N.L.S.).

Acknowledgments

The INVENT Consortium acknowledges all the participants across studies that provided their health information to support these analyses. The views expressed in this article are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute, the National Institutes of Health, the Department of Veterans Affairs, or the US Department of Health and Human Services.

Sources of Funding

The INVENT Consortium is supported in part by National Heart, Lung, and Blood Institute grants HL134894 and HL154385 and by the GENMED Laboratory of Excellence on Medical Genomics ANR-10-LABX-0013, a research program managed by the National Research Agency (ANR) as part of the French Investment for the Future. This research is based on data from the Million Veteran Program, Office of Research and Development, Veterans Health Administration, and was supported by award MVP#003. The Analysis Commons was funded by the National Heart, Lung, and Blood Institute grant R01HL131136. Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. This work was supported by National Heart, Lung, and Blood Institute Intramural Research Program funding (F.T., M.-H.C., and A.D.J.). S.M.D. is supported by the Veterans Administration grant IK2-CX001780. Study-specific acknowledgments and funding can be found in the Supplemental Material.

Disclosures

Dr Psaty serves on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. Dr Ridker has received investigator-initiated research grant support for unrelated projects from the National Heart, Lung, and Blood Institute, Operation Warp Speed, Novartis, Kowa, Amarin, and Pfizer; and has served as a consultant on unrelated issues to Novo Nordisk, Flame, Agepha, Uppton, Novartis, Jansen, Health Outlook, Civi Biopharm, Alnylam, and SOCAR. Dr Natarajan reports investigator-initiated grants from Amgen, Apple, AstraZeneca, Boston Scientific, and Novartis, and personal fees from Allelica, Apple, AstraZeneca, Blackstone Life Sciences, Foresite Labs, Novartis, and Roche/Genentech; is a co-founder of TenSixteen Bio; is a scientific advisory board member of Esperion Therapeutics, geneXwell, and TenSixteen Bio; and reports spousal employment at Vertex, all unrelated to the present work. Dr Chasman has received research funding for unrelated projects from Pfizer. Dr Ko initiated a research grant from Boston Scientific and received a consulting fee from Eagle Pharmaceutical, both unrelated to the current work. Dr O'Donnell is employed by Novartis Institute of Biomedical Research. Dr Cuellar-Partida and Dr Wang are employed by and hold stock or stock options in 23andMe, Inc. The spouse of Dr Willer works at Regeneron Pharmaceuticals. Dr Clapham reports fees from Tectonics Therapeutics. Dr Li-Gao is a contractor for Metabolon, Inc. Dr Do reports receiving grants from AstraZeneca, and grants and nonfinancial support from Goldfinch Bio; being a scientific co-founder, consultant, and equity holder for Pensieve Health (pending); and being a consultant for Variant Bio, all not related to this work. Dr Klarin is a scientific advisor and received consulting fees from Bitterroot Bio, Inc, unrelated to the current research. Dr Damrauer receives research support from RenalytixAI and personal consulting fees from Calico Labs, outside the scope of the current research. Dr Damrauer is named as a co-inventor on a government-owned US Patent application related to the use of genetic risk prediction for venous thromboembolic disease filed by the US Department of Veterans Affairs in accordance with federal regulatory requirements. The other authors report no conflicts.

Supplemental Material

Supplemental Methods
Supplemental Discussion
Figures S1–S13
Tables S1–S16
References 79–123

REFERENCES

1. Silverstein MD, Heit JA, Mohr DN, Petterson TM, O'Fallon WM, Melton LJ. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Arch Intern Med.* 1998;158:585–593. doi: 10.1001/archinte.158.6.585
2. Ghanima W, Brodin E, Schultze A, Shepherd L, Lambrelli D, Ulvestad M, Ramagopalan S, Halvorsen S. Incidence and prevalence of venous thromboembolism in Norway 2010-2017. *Thromb Res.* 2020;195:165–168. doi: 10.1016/j.thromres.2020.07.011
3. Delluc A, Tromeur C, Le Ven F, Gouillou M, Paleiron N, Bressollette L, Nonent M, Salaun P-Y, Lacut K, Leroyer C, et al. EPIGETBO study group. Current incidence of venous thromboembolism and comparison with 1998: a community-based study in Western France. *Thromb Haemost.* 2016;116:967–974. doi: 10.1160/TH16-03-0205
4. Smith NL, Hindorff LA, Heckbert SR, Lemaitre RN, Marciante KD, Rice K, Lumley T, Bis JC, Wiggins KL, Rosendaal FR, et al. Association of genetic variations with nonfatal venous thrombosis in postmenopausal women. *JAMA.* 2007;297:489–498. doi: 10.1001/jama.297.5.489
5. Bezemer ID, Bare LA, Doggen CJM, Arellano AR, Tong C, Rowland CM, Catanese J, Young BA, Reitsma PH, Devlin JJ, et al. Gene variants associated with deep vein thrombosis. *JAMA.* 2008;299:1306–1314. doi: 10.1001/jama.299.11.1306
6. Heit JA, Armasu SM, Asmann YW, Cunningham JM, Matsumoto ME, Petterson TM, De Andrade M. A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. *J Thromb Haemost.* 2012;10:1521–1531. doi: 10.1111/j.1538-7836.2012.04810.x
7. Buil A, Trégouët D-A, Souto JC, Saut N, Germain M, Rotival M, Tiret L, Cambien F, Lathrop M, Zeller T, et al. C4BPB/C4BPA is a new susceptibility locus for venous thrombosis with unknown protein S-independent mechanism: results from genome-wide association and gene expression analyses followed by case-control studies. *Blood.* 2010;115:4644–4650. doi: 10.1182/blood-2010-01-263038
8. Tang W, Teichert M, Chasman DI, Heit JA, Morange P-E, Li G, Pankratz N, Leebeek FW, Paré G, de Andrade M, et al. A genome-wide association study for venous thromboembolism: the extended cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Genet Epidemiol.* 2013;37:512–521. doi: 10.1002/gepi.21731
9. Germain M, Chasman DI, de Haan H, Tang W, Lindström S, Weng L-C, de Andrade M, de Visser MCH, Wiggins KL, Suchon P, et al. Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet.* 2015;96:532–542. doi: 10.1016/j.ajhg.2015.01.019
10. Bennett JA, Mastrangelo MA, Ture SK, Smith CO, Loelius SG, Berg RA, Shi X, Burke RM, Spinelli SL, Cameron SJ, et al. The choline transporter Slc44a2 controls platelet activation and thrombosis by regulating mitochondrial function. *Nat Commun.* 2020;11:3479. doi: 10.1038/s41467-020-17254-w
11. Constantinescu-Bercu A, Grassi L, Frontini M, Salles-Crawley II, Woollard K, Crawley JT. Activated αIIbβ3 on platelets mediates flow-dependent NETosis via SLC44A2. *Elife.* 2020;9:e53353. doi: 10.7554/eLife.53353
12. Zirka G, Robert P, Tilburg J, Tishkova V, Maracle CX, Legendre P, van Vlijmen BJM, Alessi M-C, Lenting PJ, Morange P-E, et al. Impaired adhesion of neutrophils expressing Slc44a2/HNA-3b to VWF protects against NETosis under venous shear rates. *Blood.* 2021;137:2256–2266. doi: 10.1182/blood.2020008345
13. Lindström S, Wang L, Smith EN, Gordon W, van Hylckama Vlieg A, de Andrade M, Brody JA, Pattee JW, Haessler J, Brumpton BM, et al. Genomic and transcriptomic association studies identify 16 novel susceptibility loci for venous thromboembolism. *Blood.* 2019;134:1645–1657. doi: 10.1182/blood.2019000435
14. Klarin D, Busenkell E, Judy R, Lynch J, Levin M, Haessler J, Aragam K, Chaffin M, Haas M, Lindström S, et al; INVENT Consortium, Veterans Affairs' Million Veteran Program. Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. *Nat Genet.* 2019;51:1574–1579. doi: 10.1038/s41588-019-0519-3
15. Zhou W, Kanai M, Wu K-HH, Humaira R, Tsuo K, Hirbo JB, Wang Y, Bhattacharya A, Zhao H, Namba S, et . Global Biobank Meta-analysis Initiative: powering genetic discovery across human diseases. *medRxiv.* Preprint posted online November 21, 2021. doi:10.1101/2021.11.19.21266436
16. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26:2190–2191. doi: 10.1093/bioinformatics/btq340
17. Hinds DA, Buil A, Ziemek D, Martinez-Perez A, Malik R, Folkersen L, Germain M, Mälarstig A, Brown A, Soria JM, et al. Genome-wide association analysis of self-reported events in 6135 individuals and 252 827 controls identifies 8 loci associated with thrombosis. *Hum Mol Genet.* 2016;25:1867–1874. doi: 10.1093/hmg/ddw037
18. Devlin B, Roeder K. Genomic control for association studies. *Biometrics.* 1999;55:997–1004. doi: 10.1111/j.0006-341x.1999.00997.x
19. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, Daly MJ, Price AL, Neale BM; Schizophrenia Working Group of the Psychiatric Genomics Consortium. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet.* 2015;47:291–295. doi: 10.1038/ng.3211
20. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Madden PAF, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional

variants influencing complex traits. *Nat Genet.* 2012;44:369–375, S1-3. doi: 10.1038/ng.2213
21. Kowalski MH, Qian H, Hou Z, Rosen JD, Tapia AL, Shan Y, Jain D, Argos M, Arnett DK, Avery C, et al; NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium. Use of >100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium whole genome sequences improves imputation quality and detection of rare variant associations in admixed African and Hispanic/Latino populations. *PLoS Genet.* 2019;15:e1008500. doi: 10.1371/journal.pgen.1008500
22. Sabik OL, Farber CR. RACER: a data visualization strategy for exploring multiple genetic associations. *bioRxiv.* Preprint posted online December 14, 2018. doi:10.1101/495366
23. Myers TA, Chanock SJ, Machiela MJ. LDlinkR: An R package for rapidly calculating linkage disequilibrium statistics in diverse populations. *Front Genet.* 2020;11:157. doi: 10.3389/fgene.2020.00157
24. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf.* 2011;12:77. doi: 10.1186/1471-2105-12-77
25. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, Jansen R, de Geus EJC, Boomsma DI, Wright FA, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet.* 2016;48:245–252. doi: 10.1038/ng.3506
26. GTEx C. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science.* 2015;348:648–660. doi: 10.1126/science.1262110
27. Nuotio J, Oikonen M, Magnussen CG, Jokinen E, Laitinen T, Hutri-Kähönen N, Kähönen M, Lehtimäki T, Taittonen L, Tossavainen P, et al. Cardiovascular risk factors in 2011 and secular trends since 2007: the Cardiovascular Risk in Young Finns Study. *Scand J Public Health.* 2014;42:563–571. doi: 10.1177/1403494814541597
28. Wright FA, Sullivan PF, Brooks AI, Zou F, Sun W, Xia K, Madar V, Jansen R, Chung W, Zhou Y-H, et al. Heritability and genomics of gene expression in peripheral blood. *Nat Genet.* 2014;46:430–437. doi: 10.1038/ng.2951
29. Zhao H, Rasheed H, Nøst TH, Cho Y, Liu Y, Bhatta L, Bhattacharya A, Global Biobank Meta-Analysis Initiative, Hemani G, Smith GD, Brumpton BM, Zhou W, Neale BM, Gaunt TR, Zheng J. Proteome-wide Mendelian randomization in global biobank meta-analysis reveals multi-ancestry drug targets for common diseases. *medRxiv.* Preprint posted online January 11, 2022. doi:10.1101/2022.01.09.21268473
30. de Vries PS, Chasman DI, Sabater-Lleal M, Chen M-H, Huffman JE, Steri M, Tang W, Teumer A, Marioni RE, Grossmann V, et al. A meta-analysis of 120 246 individuals identifies 18 new loci for fibrinogen concentration. *Hum Mol Genet.* 2016;25:358–370.
31. Smith NL, Huffman JE, Strachan DP, Huang J, Dehghan A, Trompet S, Lopez LM, Shin S-Y, Baumert J, Vitart V, et al. Genetic predictors of fibrin D-dimer levels in healthy adults. *Circulation.* 2011;123:1864–1872. doi: 10.1161/CIRCULATIONAHA.110.009480
32. de Vries PS, Sabater-Lleal M, Huffman JE, Marten J, Song C, Pankratz N, Bartz TM, de Haan HG, Delgado GE, Eicher JD, et al; INVENT Consortium. A genome-wide association study identifies new loci for factor VII and implicates factor VII in ischemic stroke etiology. *Blood.* 2019;133:967–977. doi: 10.1182/blood-2018-05-849240
33. Sabater-Lleal M, Huffman JE, de Vries PS, Marten J, Mastrangelo MA, Song C, Pankratz N, Ward-Caviness CK, Yanek LR, Trompet S, et al; INVENT Consortium; MEGASTROKE Consortium of the International Stroke Genetics Consortium (ISGC). Genome-wide association transethnic meta-analyses identifies novel associations regulating coagulation factor VIII and von Willebrand factor plasma levels. *Circulation.* 2019;139:620–635. doi: 10.1161/CIRCULATIONAHA.118.034532
34. Sennblad B, Basu S, Mazur J, Suchon P, Martinez-Perez A, van Hylckama Vlieg A, Truong V, Li Y, Gådin JR, Tang W, et al. Genome-wide association study with additional genetic and post-transcriptional analyses reveals novel regulators of plasma factor XI levels. *Hum Mol Genet.* 2017;26:637–649. doi: 10.1093/hmg/ddw401
35. Huang J, Huffman JE, Yamakuchi M, Yamkauchi M, Trompet S, Asselbergs FW, Sabater-Lleal M, Trégouët D-A, Chen W-M, Smith NL, et al. Genome-wide association study for circulating tissue plasminogen activator levels and functional follow-up implicates endothelial STXBP5 and STX2. *Arterioscler Thromb Vasc Biol.* 2014;34:1093–1101. 10.1161/ATVBAHA.113.302088
36. Huang J, Sabater-Lleal M, Asselbergs FW, Tregouet D, Shin S-Y, Ding J, Baumert J, Oudot-Mellakh T, Folkersen L, Johnson AD, et al. Genome-wide association study for circulating levels of PAI-1 provides novel insights into its regulation. *Blood.* 2012;120:4873–4881. doi: 10.1182/blood-2012-06-436188
37. Tang W, Schwienbacher C, Lopez LM, Ben-Shlomo Y, Oudot-Mellakh T, Johnson AD, Samani NJ, Basu S, Gögele M, Davies G, et al. Genetic associations for activated partial thromboplastin time and prothrombin time, their gene expression profiles, and risk of coronary artery disease. *Am J Hum Genet.* 2012;91:152–162. doi: 10.1016/j.ajhg.2012.05.009
38. Chen M-H, Raffield LM, Mousas A, Sakaue S, Huffman JE, Moscati A, Trivedi B, Jiang T, Akbari P, Vuckovic D, et al. Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. *Cell.* 2020;182:1198–1213. doi: 10.1016/j.cell.2020.06.045
39. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* 2014;10:e1004383. doi: 10.1371/journal.pgen.1004383
40. Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, Bates P, Palmer T, Haberland V, Smith GD, et al. The MRC IEU OpenGWAS data infrastructure. *bioRxiv.* Preprint posted online August 10, 2020. doi:10.1101/2020.08.10.244293
41. Yang J, Weedon MN, Purcell S, Lettre G, Estrada K, Willer CJ, Smith AV, Ingelsson E, O'Connell JR, Mangino M, et al. Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet.* 2011;19:807. doi: 10.1038/ejhg.2011.39
42. Riis J, Nordestgaard BG, Afzal S. α1-Antitrypsin Z allele and risk of venous thromboembolism in the general population. *J Thromb Haemost.* 2022;20:115–125. doi: 10.1111/jth.15556
43. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, et al; Genome Aggregation Database Consortium. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581:434–443. doi: 10.1038/s41586-020-2308-7
44. Minamida S, Aoki K, Natsuka S, Omichi K, Fukase K, Kusumoto S, Hase S. Detection of UDP-D-xylose: alpha-D-xyloside alpha 1-->3xylosyltransferase activity in human hepatoma cell line HepG2. *J Biochem.* 1996;120:1002–1006. doi: 10.1093/oxfordjournals.jbchem.a021492
45. Saleque S, Kim J, Rooke HM, Orkin SH. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. *Mol Cell.* 2007;27:562–572. doi: 10.1016/j.molcel.2007.06.039
46. Schulze H, Shivdasani RA. Mechanisms of thrombopoiesis. *J Thromb Haemost.* 2005;3:1717–1724. doi: 10.1111/j.1538-7836.2005.01426.x
47. Lamonica JM, Deng W, Kadauke S, Campbell AE, Gamsjaeger R, Wang H, Cheng Y, Billin AN, Hardison RC, Mackay JP, et al. Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proc Natl Acad Sci USA.* 2011;108:E159–E168. doi: 10.1073/pnas.1102140108
48. Widom RL, Lee JY, Joseph C, Gordon-Froome I, Korn JH. The hcK-rox gene family regulates multiple extracellular matrix genes. *Matrix Biol.* 2001;20:451–462. doi: 10.1016/s0945-053x(01)00167-6
49. Perrella G, Montague SJ, Brown HC, Garcia Quintanilla L, Slater A, Stegner D, Thomas M, Heemskerk JWM, Watson SP. Role of tyrosine kinase Syk in thrombus stabilisation at high shear. *Int J Mol Sci* . 2022;23:493. doi: 10.3390/ijms23010493
50. Zheng TJ, Lofurno ER, Melrose AR, Lakshmanan HHS, Pang J, Phillips KG, Fallon ME, Kohs TCL, Ngo ATP, Shatzel JJ, et al. Assessment of the effects of Syk and BTK inhibitors on GPVI-mediated platelet signaling and function. *Am J Physiol Cell Physiol.* 2021;320:C902–C915. doi: 10.1152/ajpcell.00296.2020
51. Fredenburgh JC, Weitz JI. New anticoagulants: moving beyond the direct oral anticoagulants. *J Thromb Haemost.* 2021;19:20–29. doi: 10.1111/jth.15126
52. Mackman N, Bergmeier W, Stouffer GA, Weitz JI. Therapeutic strategies for thrombosis: new targets and approaches. *Nat Rev Drug Discov.* 2020;19:333–352. doi: 10.1038/s41573-020-0061-0
53. Han L, Madan V, Mayakonda A, Dakle P, Woon TW, Shyamsunder P, Nordin HBM, Cao Z, Sundaresan J, Lei I, et al. Chromatin remodeling mediated by ARID1A is indispensable for normal hematopoiesis in mice. *Leukemia.* 2019;33:2291–2305. doi: 10.1038/s41375-019-0438-4
54. Scheicher R, Hoelbl-Kovacic A, Bellutti F, Tigan A-S, Prchal-Murphy M, Heller G, Schneckenleithner C, Salazar-Roa M, Zöchbauer-Müller S, Zuber J, et al. CDK6 as a key regulator of hematopoietic and leukemic stem cell activation. *Blood.* 2015;125:90–101. doi: 10.1182/blood-2014-06-584417
55. Maslah N, Cassinat B, Verger E, Kiladjian J-J, Velazquez L. The role of LNK/SH2B3 genetic alterations in myeloproliferative neoplasms and other hematological disorders. *Leukemia.* 2017;31:1661–1670. doi: 10.1038/leu.2017.139

56. Mancini E, Sanjuan-Pla A, Luciani L, Moore S, Grover A, Zay A, Rasmussen KD, Luc S, Bilbao D, O'Carroll D, et al. FOG-1 and GATA-1 act sequentially to specify definitive megakaryocytic and erythroid progenitors. *EMBO J.* 2012;31:351–365. doi: 10.1038/emboj.2011.390
57. Krosl J, Mamo A, Chagraoui J, Wilhelm BT, Girard S, Louis I, Lessard J, Perreault C, Sauvageau G. A mutant allele of the Swi/Snf member BAF250a determines the pool size of fetal liver hemopoietic stem cell populations. *Blood.* 2010;116:1678–1684. doi: 10.1182/blood-2010-03-273862
58. Ayoub E, Wilson MP, McGrath KE, Li AJ, Frisch BJ, Palis J, Calvi LM, Zhang Y, Perkins AS. EVI1 overexpression reprograms hematopoiesis via upregulation of Spi1 transcription. *Nat Commun.* 2018;9:4239. doi: 10.1038/s41467-018-06208-y
59. Fonseca-Pereira D, Arroz-Madeira S, Rodrigues-Campos M, Barbosa IAM, Domingues RG, Bento T, Almeida ARM, Ribeiro H, Potocnik AJ, Enomoto H, et al. The neurotrophic factor receptor RET drives haematopoietic stem cell survival and function. *Nature.* 2014;514:98–101. doi: 10.1038/nature13498
60. Gregory GD, Miccio A, Bersenev A, Wang Y, Hong W, Zhang Z, Poncz M, Tong W, Blobel GA. FOG1 requires NuRD to promote hematopoiesis and maintain lineage fidelity within the megakaryocytic-erythroid compartment. *Blood.* 2010;115:2156–2166. doi: 10.1182/blood-2009-10-251280
61. Keramati AR, Chen M-H, Rodriguez BAT, Yanek LR, Bhan A, Gaynor BJ, Ryan K, Brody JA, Zhong X, Wei Q, et al; NHLBI Trans-Omics for Precision (TOPMed) Consortium. Genome sequencing unveils a regulatory landscape of platelet reactivity. *Nat Commun.* 2021;12:3626. doi: 10.1038/s41467-021-23470-9
62. Mitsui T, Makino S, Tamiya G, Sato H, Kawakami Y, Takahashi Y, Meguro T, Izumino H, Sudo Y, Norota I, et al. ALOX12 mutation in a family with dominantly inherited bleeding diathesis. *J Hum Genet.* 2021;66:753–759. doi: 10.1038/s10038-020-00887-6
63. Fukami K. Structure, regulation, and function of phospholipase C isozymes. *J Biochem.* 2002;131:293–299. doi: 10.1093/oxfordjournals.jbchem.a003102
64. Johnson AD, Yanek LR, Chen M-H, Faraday N, Larson MG, Tofler G, Lin SJ, Kraja AT, Province MA, Yang Q, et al. Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. *Nat Genet.* 2010;42:608–613. doi: 10.1038/ng.604
65. Radomski A, Jurasz P, Sanders EJ, Overall CM, Bigg HF, Edwards DR, Radomski MW. Identification, regulation and role of tissue inhibitor of metalloproteinases-4 (TIMP-4) in human platelets. *Br J Pharmacol.* 2002;137:1330–1338. doi: 10.1038/sj.bjp.0704936
66. Moore SF, Smith NR, Blair TA, Durrant TN, Hers I. Critical roles for the phosphatidylinositide 3-kinase isoforms p110β and p110γ in thrombopoietin-mediated priming of platelet function. *Sci Rep.* 2019;9:1468. doi: 10.1038/s41598-018-37012-9
67. Kuijpers MJE, Mattheij NJA, Cipolla L, van Geffen JP, Lawrence T, Donners MMPC, Boon L, Lievens D, Torti M, Noels H, et al. Platelet CD40L modulates thrombus growth via phosphatidylinositol 3-kinase β, and not via CD40 and IκB kinase α. *Arterioscler Thromb Vasc Biol.* 2015;35:1374–1381. doi: 10.1161/ATVBAHA.114.305127
68. Rodriguez BAT, Bhan A, Beswick A, Elwood PC, Niiranen TJ, Salomaa V, FinnGen S, Trégouët D-A, Morange P-E, Civelek M, et al. A platelet function modulator of thrombin activation is causally linked to cardiovascular disease and affects PAR4 receptor signaling. *Am J Hum Genet.* 2020;107:211–221. doi: 10.1016/j.ajhg.2020.06.008
69. Antl M, von Brühl M-L, Eiglsperger C, Werner M, Konrad I, Kocher T, Wilm M, Hofmann F, Massberg S, Schlossmann J. IRAG mediates NO/cGMP-dependent inhibition of platelet aggregation and thrombus formation. *Blood.* 2007;109:552–559. doi: 10.1182/blood-2005-10-026294
70. Schinner E, Salb K, Schlossmann J. Signaling via IRAG is essential for NO/cGMP-dependent inhibition of platelet activation. *Platelets.* 2011;22:217–227. doi: 10.3109/09537104.2010.544151
71. van Geffen JP, Brouns SLN, Batista J, McKinney H, Kempster C, Nagy M, Sivapalaratnam S, Baaten CCFMJ, Bourry N, Frontini M, et al. High-throughput elucidation of thrombus formation reveals sources of platelet function variability. *Haematologica.* 2019;104:1256–1267. doi: 10.3324/haematol.2018.198853
72. Braekkan SK, Mathiesen EB, Njølstad I, Wilsgaard T, Størmer J, Hansen JB. Mean platelet volume is a risk factor for venous thromboembolism: the Tromsø Study, Tromsø, Norway. *J Thromb Haemost.* 2010;8:157–162. doi: 10.1111/j.1538-7836.2009.03498.x
73. Ghaffari S, Parvizian N, Pourafkari L, Separham A, Hajizadeh R, Nader ND, Javanshir E, Sepehrvand N, Tajlil A, Nasiri B. Prognostic value of platelet indices in patients with acute pulmonary thromboembolism. *J Cardiovasc Thorac Res.* 2020;12:56–62. doi: 10.34172/jcvtr.2020.09
74. Farah R, Nseir W, Kagansky D, Khamisy-Farah R. The role of neutrophil-lymphocyte ratio, and mean platelet volume in detecting patients with acute venous thromboembolism. *J Clin Lab Anal.* 2020;34:e23010. doi: 10.1002/jcla.23010
75. Puurunen MK, Hwang S-J, O'Donnell CJ, Tofler G, Johnson AD. Platelet function as a risk factor for venous thromboembolism in the Framingham Heart Study. *Thromb Res.* 2017;151:57–62. doi: 10.1016/j.thromres.2017.01.010
76. Sokol J, Skerenova M, Ivankova J, Simurda T, Stasko J. Association of genetic variability in selected genes in patients with deep vein thrombosis and platelet hyperaggregability. *Clin Appl Thromb Hemost.* 2018;24:1027–1032. doi: 10.1177/1076029618779136
77. Panova-Noeva M, Wagner B, Nagler M, Koeck T, Ten Cate V, Prochaska JH, Heitmeier S, Meyer I, Gerdes C, Laux V, et al. Comprehensive platelet phenotyping supports the role of platelets in the pathogenesis of acute venous thromboembolism - results from clinical observation studies. *EBioMedicine.* 2020;60:102978. doi: 10.1016/j.ebiom.2020.102978
78. Diep R, Garcia D. Does aspirin prevent venous thromboembolism?. *Hematology Am Soc Hematol Educ Program.* 2020;2020:634–641. doi: 10.1182/hematology.2020000150
79. Klarin D, Emdin CA, Natarajan P, Conrad MF, Kathiresan S. Genetic analysis of venous thromboembolism in UK Biobank identifies the ZFPM2 locus and implicates obesity as a causal risk factor. *Circ Cardiovasc Genet.* 2017;10:e001643. doi: 10.1161/CIRCGENETICS.116.001643
80. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015;12:e1001779. doi: 10.1371/journal.pmed.1001779
81. Mitt M, Kals M, Pärn K, Gabriel SB, Lander ES, Palotie A, Ripatti S, Morris AP, Metspalu A, Esko T, et al. Improved imputation accuracy of rare and low-frequency variants using population-specific high-coverage WGS-based imputation reference panel. *Eur J Hum Genet.* 2017;25:869–876. doi: 10.1038/ejhg.2017.51
82. Nagai A, Hirata M, Kamatani Y, Muto K, Matsuda K, Kiyohara Y, Ninomiya T, Tamakoshi A, Yamagata Z, Mushiroda T, et al; BioBank Japan Cooperative Hospital Group. Overview of the BioBank Japan Project: study design and profile. *J Epidemiol.* 2017;27:S2–S8. doi: 10.1016/j.je.2016.12.005
83. Smoller JW, Karlson EW, Green RC, Kathiresan S, MacArthur DG, Talkowski ME, Murphy SN, Weiss ST. An eMERGE Clinical Center at Partners Personalized Medicine. *J Pers Med.* 2016;6:E5. doi: 10.3390/jpm6010005
84. Antoni G, Morange P-E, Luo Y, Saut N, Burgos G, Heath S, Germain M, Biron-Andreani C, Schved J-F, Pernod G, et al. A multi-stage multi-design strategy provides strong evidence that the BAI3 locus is associated with early-onset venous thromboembolism. *J Thromb Haemost.* 2010;8:2671–2679. doi: 10.1111/j.1538-7836.2010.04092.x
85. Ibrahim-Kosta M, Suchon P, Couturaud F, Smadja D, Olaso R, Germain M, Saut N, Goumidi L, Derbois C, Thibord F, et al. Minor allele of the factor V K858R variant protects from venous thrombosis only in non-carriers of factor V Leiden mutation. *Sci Rep.* 2019;9:3750. doi: 10.1038/s41598-019-40172-x
86. Vázquez-Santiago M, Vilalta N, Cuevas B, Murillo J, Llobet D, Macho R, Pujol-Moix N, Carrasco M, Mateo J, Fontcuberta J, et al. Short closure time values in PFA-100® are related to venous thrombotic risk. Results from the RETROVE Study. *Thromb Res.* 2018;169:57–63. doi: 10.1016/j.thromres.2018.07.012
87. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Kronmal R, Liu K, et al. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol.* 2002;156:871–881. doi: 10.1093/aje/kwf113
88. Souto JC, Almasy L, Borrell M, Garí M, Martínez E, Mateo J, Stone WH, Blangero J, Fontcuberta J. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation.* 2000;101:1546–1551. doi: 10.1161/01.cir.101.13.1546
89. Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM, Coll I, Felices R, Stone W, Fontcuberta J, et al. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. Genetic Analysis of Idiopathic Thrombophilia. *Am J Hum Genet.* 2000;67:1452–1459. doi: 10.1086/316903

90. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC Investigators. *Am J Epidemiol.* 1989;129:687–702. doi: 10.1093/oxfordjournals.aje.a115184
91. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol.* 1991;1:263–276. doi: 10.1016/1047-2797(91)90005-w
92. Tell GS, Fried LP, Hermanson B, Manolio TA, Newman AB, Borhani NO. Recruitment of adults 65 years and older as participants in the Cardiovascular Health Study. *Ann Epidemiol.* 1993;3:358–366. doi: 10.1016/1047-2797(93)90062-9
93. Trégouët D-A, Heath S, Saut N, Biron-Andreani C, Schved J-F, Pernod G, Galan P, Drouet L, Zelenika D, Juhan-Vague I, et al. Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. *Blood.* 2009;113:5298–5303. doi: 10.1182/blood-2008-11-190389
94. McCarty CA, Chisholm RL, Chute CG, Kullo IJ, Jarvik GP, Larson EB, Li R, Masys DR, Ritchie MD, Roden DM, et al. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics.* 2011;4:13. doi: 10.1186/1755-8794-4-13
95. Milani L, Leitsalu L, Metspalu A. An epidemiological perspective of personalized medicine: the Estonian experience. *J Intern Med.* 2015;277:188–200. doi: 10.1111/joim.12320
96. Zhu T, Carcaillon L, Martinez I, Cambou J-P, Kyndt X, Guillot K, Vergnes M-C, Scarabin P-Y, Emmerich J. Association of influenza vaccination with reduced risk of venous thromboembolism. *Thromb Haemost.* 2009;102:1259–1264. doi: 10.1160/TH09-04-0222
97. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol.* 1979;110:281–290. doi: 10.1093/oxfordjournals.aje.a112813
98. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med.* 1975;4:518–525. doi: 10.1016/0091-7435(75)90037-7
99. Smith NL, Heckbert SR, Lemaitre RN, Reiner AP, Lumley T, Weiss NS, Larson EB, Rosendaal FR, Psaty BM. Esterified estrogens and conjugated equine estrogens and the risk of venous thrombosis. *JAMA.* 2004;292:1581–1587. doi: 10.1001/jama.292.13.1581
100. Holmen J, Midthjell K, Kruger O, Langhammer A, Holmen TL, Bratberg GH, Vatten L, Lund-Larsen PG. The Nord-Trøndelag Health Study 1995-97 (HUNT 2). *Norsk Epidemiol.* 2003;13:19–32. doi: 10.5324/nje.v13i1.305
101. Glynn RJ, Danielson E, Fonseca FAH, Genest J, Gotto AM, Kastelein JJP, Koenig W, Libby P, Lorenzatti AJ, MacFadyen JG, et al. A randomized trial of rosuvastatin in the prevention of venous thromboembolism. *N Engl J Med.* 2009;360:1851–1861. doi: 10.1056/NEJMoa0900241
102. Ridker PM, Danielson E, Fonseca FAH, Genest J, Gotto AM, Kastelein JJP, Koenig W, Libby P, Lorenzatti AJ, MacFadyen JG, et al. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med.* 2008;359:2195–2207. doi: 10.1056/NEJMoa0807646
103. Chasman DI, Giulianini F, MacFadyen J, Barratt BJ, Nyberg F, Ridker PM. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction: the Justification for the Use of Statins in Prevention: An Intervention Trial Evaluating Rosuvastatin (JUPITER) Trial. *Circ Cardiovasc Genet.* 2012;5:257–264. doi: 10.1161/CIRCGENETICS.111.961144
104. Oudot-Mellakh T, Cohen W, Germain M, Saut N, Kallel C, Zelenika D, Lathrop M, Trégouët D-A, Morange P-E. Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol.* 2012;157:230–239. doi: 10.1111/j.1365-2141.2011.09025.x
105. 3C Study Group. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology.* 2003;22:316–325. doi: 10.1159/000072920
106. Blom JW, Doggen CJM, Osanto S, Rosendaal FR. Malignancies, prothrombotic mutations, and the risk of venous thrombosis. *JAMA.* 2005;293:715–722. doi: 10.1001/jama.293.6.715
107. Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J, Whitbourne S, Deen J, Shannon C, Humphries D, et al. Million Veteran Program: a mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol.* 2016;70:214–223. doi: 10.1016/j.jclinepi.2015.09.016
108. Hankinson SE, Colditz GA, Hunter DJ, Manson JE, Willett WC, Stampfer MJ, Longcope C, Speizer FE. Reproductive factors and family history of breast cancer in relation to plasma estrogen and prolactin levels in postmenopausal women in the Nurses' Health Study (United States). *Cancer Causes Control.* 1995;6:217–224. doi: 10.1007/BF00051793
109. Tworoger SS, Sluss P, Hankinson SE. Association between plasma prolactin concentrations and risk of breast cancer among predominately premenopausal women. *Cancer Res.* 2006;66:2476–2482. doi: 10.1158/0008-5472.CAN-05-3369
110. Jacobsen BK, Eggen AE, Mathiesen EB, Wilsgaard T, Njølstad I. Cohort profile: the Tromso Study. *Int J Epidemiol.* 2012;41:961–967. doi: 10.1093/ije/dyr049
111. Braekkan SK, Mathiesen EB, Njølstad I, Wilsgaard T, Størmer J, Hansen JB. Family history of myocardial infarction is an independent risk factor for venous thromboembolism: the Tromsø study. *J Thromb Haemost.* 2008;6:1851–1857. doi: 10.1111/j.1538-7836.2008.03102.x
112. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials.* 1998;19:61–109. doi: 10.1016/s0197-2456(97)00078-0
113. Anderson GL, Manson J, Wallace R, Lund B, Hall D, Davis S, Shumaker S, Wang C-Y, Stein E, Prentice RL. Implementation of the Women's Health Initiative study design. *Ann Epidemiol.* 2003;13:S5–17. doi: 10.1016/s1047-2797(03)00043-7
114. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Mägi R, Ferreira T, Fall T, Graff M, Justice AE, et al; Genetic Investigation of Anthropometric Traits (GIANT) Consortium. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc.* 2014;9:1192–1212. doi: 10.1038/nprot.2014.071
115. Wolfe D, Dudek S, Ritchie MD, Pendergrass SA. Visualizing genomic information across chromosomes with PhenoGram. *BioData Min.* 2013;6:18. doi: 10.1186/1756-0381-6-18
116. Lee SH, Goddard ME, Wray NR, Visscher PM. A better coefficient of determination for genetic profile analysis. *Genet Epidemiol.* 2012;36:214–224. doi: 10.1002/gepi.21614
117. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature.* 2015;526:68–74. doi: 10.1038/nature15393
118. Zhang J, Dutta D, Köttgen A, Tin A, Schlosser P, Grams ME, Harvey B, CKDGen C, Yu B, Boerwinkle E, et al. Plasma proteome analyses in individuals of European and African ancestry identify cis-pQTLs and models for proteome-wide association studies. *Nat Genet.* 2022;54:593–602. doi: 10.1038/s41588-022-01051-w
119. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, et al. Genomic atlas of the human plasma proteome. *Nature.* 2018;558:73–79. doi: 10.1038/s41586-018-0175-2
120. Folkersen L, Fauman E, Sabater-Lleal M, Strawbridge RJ, Frånberg M, Sennblad B, Baldassarre D, Veglia F, Humphries SE, Rauramaa R, et al. Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. *PLoS Genet.* 2017;13:e1006706. doi: 10.1371/journal.pgen.1006706
121. Desch KC, Ozel AB, Halvorsen M, Jacobi PM, Golden K, Underwood M, Germain M, Tregouet D-A, Reitsma PH, Kearon C, et al. Whole-exome sequencing identifies rare variants in STAB2 associated with venous thromboembolic disease. *Blood.* 2020;136:533–541. doi: 10.1182/blood.2019004161
122. Backman JD, Li AH, Marcketta A, Sun D, Mbatchou J, Kessler MD, Benner C, Liu D, Locke AE, Balasubramanian S, et al. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature.* 2021;599:628–634. doi: 10.1038/s41586-021-04103-z
123. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, et al. Genomic atlas of the human plasma proteome. *Nature.* 2018;558:73–79. doi: 10.1038/s41586-018-0175-2