Genome-wide association trans-ethnic meta-analyses identifies novel associations regulating coagulation Factor VIII and von Willebrand Factor plasma levels

Maria Sabater-Lleal, Jennifer E. Huffman, Paul S. de Vries, Jonathan Marten *et al.*

Short title: Genetic regulation of FVIII and VWF

Corresponding authors:

Nicholas L. Smith, PhD
Department of Epidemiology, University of Washington
1730 Minor Ave, Suite 1360, Seattle, WA, USA, 98101
Phone: +1 206 221 7775
Email: nlsmith@u.washington.edu

Maria Sabater-Lleal, PhD
Department of Medicine, Karolinska Institutet
Karolinska Hospital, 17176 Stockholm, Sweden
Email: maria.sabater.lleal@ki.se

Total word count: 9,241
Abstract word count: 249
Total number of references: 36
Total number of tables: 2
Total number of figures 3
Total number of supplemental information: 2 documents

ABSTRACT

Background: Factor VIII (FVIII) and its carrier protein von Willebrand factor (VWF) are associated with risk of arterial and venous thrombosis and with hemorrhagic disorders. We aimed to identify and functionally test novel genetic associations regulating plasma FVIII and VWF.

Methods: We meta-analyzed genome-wide association results from 46,354 individuals of European, African, East Asian, and Hispanic ancestry. All studies performed linear regression analysis using an additive genetic model and associated approximately 35 million imputed variants with natural-log transformed phenotype levels. *In vitro* gene silencing in cultured endothelial cells was performed for candidate genes to provide additional evidence on association and function. Two-sample Mendelian randomization (MR) analyses were applied to test the causal role of FVIII and VWF plasma levels on the risk of arterial and venous thrombotic events.

Results: We identified 13 novel genome-wide significant ($p \leq 2.5 \times 10^{-8}$) associations; 7 with FVIII levels (*FCHO2/TMEM171/TNPO1*, *HLA*, *SOX17/RP1*, *LINC00583/NFIB*, *RAB5C-KAT2A*, *RPL3/TAB1/SYNGR1*, and *ARSA*) and 11 with VWF levels (*PDHB/PXK/KCTD6*, *SLC39A8*, *FCHO2/TMEM171/TNPO1*, *HLA*, *GIMAP7/GIMAP4*, *OR13C5/NIPSNAP*, *DAB2IP*, *C2CD4B*, *RAB5C-KAT2A*, *TAB1/SYNGR1*, and *ARSA*, beyond 10 previously reported associations with these phenotypes. Functional validation provided further evidence of association for all loci on VWF except *ARSA* and *DAB2IP*. MR suggested causal effects of plasma FVIII activity levels on venous thrombosis and coronary artery disease risk and plasma VWF levels on ischemic stroke risk.

Conclusions: The meta-analysis identified 13 novel genetic loci regulating FVIII and VWF plasma levels, 10 of which we validated functionally. We provide some evidence for a causal role of these proteins in thrombotic events.

CLINICAL PERSPECTIVE

What is new?

- Plasma coagulation factor VIII (FVIII) and von Willebrand factor (VWF) concentrations are associated with risk of cardiovascular disease, but the factors that control their levels are not fully understood.
- Using a multi-ethnic meta-analysis of genome wide association studies, we identified 7 genome-wide significant novel associations for FVIII and 11 for VWF.

What are the clinical implications?

- We evaluated the effect of genetic variants with coronary artery disease, ischemic stroke, and venous thrombosis through Mendelian randomization analyses and found evidence of a causal effect of FVIII activity levels on venous thrombosis and coronary artery disease risk, and a causal effect of plasma VWF levels on stroke risk.
- Our findings suggest that FVIII and VWF may be potential therapeutic targets to prevent thrombotic events.

INTRODUCTION

Factor VIII (FVIII) and its carrier protein von Willebrand factor (VWF) regulate hemostasis and thrombosis, and higher plasma levels of these factors have been associated with risk of arterial and venous thrombosis, while lower levels are associated with hemorrhagic disorders[1-4]and with reduced risk of thrombotic events [5]. Previously published genetic association studies have investigated the contribution of nucleotide variation to plasma levels of these factors using genome-wide and exome-wide markers[6-8]. These studies identified and replicated 8 genetic loci associated with plasma VWF levels (*STXBP5, SCARA5, ABO, STAB2, STX2, VWF, TCN2* and *CLEC4M*), 5 of which were also associated with FVIII levels (*STXBP5, SCARA5, ABO, STAB2,* and *VWF*). These discoveries have broadened our understanding of the regulation of hemostasis through follow-up functional investigations[9, 10].

The causal effect of these factors on bleeding is well-established, since severe FVIII and VWF deficiencies lead to bleeding disorders hemophilia A and von Willebrand disease, respectively. While it is currently unclear whether FVIII and VWF levels causally influence the risk of thrombotic diseases, some genetic and observational evidence point towards an effect of these proteins on thrombotic disease. Genetic variants in *F8* gene and in 3 VWF-associated genes (*ABO, STXBP5 and VWF*) are robustly associated with risk of venous thrombosis but no causal association has been established[11-13].

The aim of this investigation was to identify new genetic associations that influence plasma levels of FVIII and VWF by expanding the size and ancestral diversity of the discovery sample from previous genome-wide association studies (GWAS) and by improving coverage of the

genome through the use of *1000 Genomes* imputed data and the inclusion of chromosome X variants[14]. For discoveries that reached genome-wide significance, we conducted first-pass functional characterization of the candidate loci both to provide additional evidence of association and to better understand the biology regulating plasma levels of these coagulation phenotypes. Last, by applying our genetic findings as instrument variables, we characterized the causal effect of plasma FVIII and VWF levels on clinical cardiovascular (CV) events using Mendelian randomization (MR) analyses.

METHODS

Due to patient confidentiality agreements and to comply with the study participants consent, the original data and study materials cannot be made available to other researchers for purposes of reproducing the results or replicating the procedure. Analytic methods will be made available upon request, and summary statistics have been made publicly available through dbGaP.

Study Design and Participating Cohorts

This project was organized within the Cohorts of Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium Hemostasis Working Group[15]. We meta-analyzed phenotype-genotype associations of low-frequency and common (minor allele frequency [MAF] > 0.01) variants in 32,610 individuals from 9 studies with FVIII levels, and in 46,354 individuals from 18 studies with VWF levels. A total of 20 studies contributed to one or both of the analyses; these included participants of European (EUR), African (AFR), East Asian (ASI), and Hispanic (HIS) ancestry. Descriptions and ancestry composition of participating cohorts are found in Supplementary Table S1. All studies were approved by appropriate institutional review

committees and all participants gave written informed consent for themselves and their minor children for the use of their DNA.

Study-Level Methods

Genotype Calling and Quality Control

All participating cohorts performed genotyping using commercial genotyping platforms available from Illumina or Affymetrix. Each study performed genotyping quality control checks and imputed the approximately 35 million polymorphic autosomal and X-chromosome variants described in the *1000 Genomes* population phase 1 version 3 for each participant using available imputation methods[16]. Variant calling and quality control procedures for each cohort are described in Supplementary Table S1.

Association Analyses

Plasma FVIII activity or VWF antigen levels were measured in all participants and reported in % or IU/ml*100 units. Participants with plasma FVIII or VWF levels (or activity levels) 3-standard-deviations above or below the population mean were removed, as were individuals on anticoagulation therapy. Natural-log transformed FVIII activity and VWF antigen levels (% or IU/ml*100 units) were analyzed separately within each study. Study-specific regression analyses using an additive model of inheritance were performed for imputed variant dosages and phenotype levels, adjusting for sex, age, study design variables, and population substructure using principal components. All analyses were stratified by ancestry and then meta-analyzed. X-chromosome variants were additionally stratified by sex, with dosage values for males coded as 0/2.

Meta-Analysis Methods

Quality Control

Study-specific findings were uploaded centrally and a quality control (QC) pipeline of all individual studies prior to meta-analysis was conducted using the EasyQC software package[17]. Variants with beta or standard errors (SE) values > 5 or imputation values < 0.3 were excluded from the analysis. Estimated minor allele counts (eMAC) calculated for all SNPs were a function of allele frequency, total number of samples per cohort, and imputation quality; values <10 were excluded from analysis. Alleles were harmonized according to *1000 Genomes phase1 version3* reference panel and duplicated SNPs or SNPs that had inconsistencies with the reference were excluded.

Meta, Trans-Ethnic, and Multi-Phenotype Discovery Analyses

Meta-analyses were performed in METAL within each ancestry group using a fixed-effects inverse-variance weighted approach then combined in a trans-ethnic analysis using the same method[18]. The trans-ethnic analyses are presented as discovery results and we used the ancestry-specific analyses to inform and interpret these signals. An association was considered genome-wide statistically significant at p-value $<2.5 \times 10^{-8}$ to correct for the low-frequency variants that were not included in the initial generation of GWAS[19] and only variants that passed QC in at least 3 cohorts were reported. Variants with MAF below 1% were filtered out after the meta-analyses. A genomic control coefficient was computed for each discovery cohort and was used to correct for cryptic relatedness. Finally, a locus was defined as +/- 1Mb from the SNP with the lowest p-value, and the SNP with the lowest p-value was selected to represent the locus. Multi-phenotype methods are described in Supplementary Methods.

Functional Characterization of Candidate Loci through Gene Silencing

In the absence of replication cohorts, we conducted first-pass functional characterization of the candidate loci to provide additional evidence of association. For each genome-wide significant locus, we selected candidate genes that could be responsible for the observed associations. Selection was based on proximity to the most associated SNPs in each region, information from public databases on putative effect of the SNPs in terms of regulation of expression and function of nearby genes, and hypothesis for a biological mechanism to regulate VWF/FVIII. This selection process identified 1 to 3 candidate genes for each associated locus. To screen for functionality, human umbilical vein endothelial cells (HUVEC; Life Line Cell Technology) were plated on collagen coated 96-well plates and transfected with siRNA (Silencer Select, ThermoFisher Scientific) directed against the candidate genes using the transfection reagent oligofectamine (ThermoFisher Scientific). Cells were cultured for 4 days after transfection, and media was then replaced with control media or media containing 10 μM of histamine for 30 minutes, to stimulate an inflammatory response. The FVIII and VWF in the media was measured by an ELISA using antibodies from Fitzgerald Industries and had detection ranges of 0.003-0.21 IU/ml for the FVIII assay and 0.5-120 ng/ul for the VWF assay. Every experiment was repeated 4 times and results are expressed as the mean ± standard deviation (SD) of relative expression compared with a negative control (transfected with scramble siRNA).

Follow-up Genetic Analyses

Conditional Analyses

To identify additional independent genetic signals at the associated loci, we used an approximate method implemented in GCTA for conditional and joint analysis using meta-analysis summary

statistics[20]. We used best-guess imputation for variants with imputation quality >0.3 in 8,481 European-ancestry individuals from the Framingham Heart Study (FHS) as the reference panel. A description of the FHS is given in the Supplementary Methods. In order to prevent spurious conditional associations arising from a discrepancy between linkage disequilibrium in our GWAS and the reference panel, we also performed the conditional analysis on the results of the European-ancestry meta-analysis as a sensitivity analysis, since different genetic variants showed the strongest association in the trans-ethnic analysis compared with the European-only analysis.

Mendelian Randomization

For the sentinel variant in each locus in FVIII and VWF analyses, we conducted *in silico* lookups for the associations of each individual variant with 3 major CV events: coronary artery disease (CAD) in the CARDIOGRAMplusC4D Consortium[21, 22], ischemic stroke (IS) in the MEGASTROKE analysis within the International Stroke Genetics Consortium[23], and venous thromboembolism (VTE) in the International Network on Venous Thrombosis (INVENT) Consortium[11]. We conducted 2-sample Mendelian Randomization (MR) analyses to detect any potential causal effects of plasma FVIII and VWF levels on each CV outcome, separately. We used summary statistics to generate 1 causal estimate per significant locus as the ratio of the variant's association with disease to the variant's association with the exposure, and estimates were then meta-analyzed using an inverse-variance weighted approach as our primary MR estimate, known as the inverse-variance weighted (IVW) estimate[24]. Additional methods to avoid bias due to heterogeneity, and the final variants composing the instrumental variables are further described in Supplementary Methods and in Supplementary Tables S2, S3, and S4. Since FVIII plasma levels are largely determined by VWF plasma levels owing to VWF's carrier role for

FVIII in plasma, essentially all genetic predictors of plasma VWF levels are also predictors of FVIII plasma levels. The inverse, however is not true, and a small subset of variants predict FVIII plasma levels without predicting VWF levels. To investigate the independent causal role of FVIII plasma levels from that of VWF plasma levels on CVD events, we applied a multivariable MR (MVMR) approach where we adjusted for VWF variants effects in the estimate of causal association between FVIII and CVD outcomes[25].

RESULTS

FVIII, VWF, and Multi-phenotype Meta-Analyses

Agnostic Discovery

Data used for FVIII meta-analysis was available from 25,897 European (EA), 4,500 African (AA), 773 East or Indian Asian (EAA, IAA), and 1,440 Hispanic (HA) participants. Trans-ethnic meta-analysis for FVIII resulted 13,887,196 variants passing all filters, and identified 1,431 variants that reached genome-wide statistical significance at 11 loci. Data used for VWF was available from 42,379 EA, 3,700 AA, and 275 HA participants. Meta-analysis for VWF resulted in 10,537,485 variants passing all filters, and identified 2,453 genome-wide significant variants at 17 loci (Figures 1A-B). European-specific meta-analysis identified one significant variant at one additional locus. Analysis using combined FVIII and VWF phenotypes (see Supplementary Methods) identified 2,828 variants reaching genome-wide significance at 2 additional loci, which were not identified in single-phenotype analyses.

Table 1 shows the genetic discovery results for the FVIII, VWF, and combined FVIII-VWF phenotypes. Overall, 23 unique loci were identified. Among these, 13 were new associations not

previously reported. Among the newly identified loci, 7 were associated with FVIII levels (*FCHO2/TMEM171/TNPO1, HLA, SOX17/RP1, LINC00583/NFIB, RAB5C/KAT2A, RPL3/TAB1/SYNGR1/PDGB,* and *ARSA*) and 11 were associated with VWF levels (*PDHB/PXK/KCTD6, SLC39A8, FCHO2/TMEM171/TNPO1, HLA, GIMAP7/GIMAP4, OR13C5/NIPSNAP, DAB2IP, C2CD4B, RAB5C/KAT2A,* RPL3/*TAB1/SYNGR1/PDGB,* and *ARSA*). Supplementary Figures S1a-n shows regional plots for the novel loci plotted for the 2 phenotypes. The lowest MAF for the index variant was 0.02 while the effect size per allele ranged from 0.015 to 0.032 (in log transformed units) for FVIII levels and from 0.014 to 0.060 for VWF levels.

Among the 23 genome-wide significant findings, 10 loci were previously reported to be associated with plasma levels of FVIII or VWF or both: *STXBP5, SCARA5, ABO, ST3GAL4, STAB2, STX2, VWF, TCN2, CLEC4M,* and *TMLHE-F8 region.*

Conditional Analyses and Variant Characterization

In follow-up analyses, we conditioned on sentinel variants to determine if secondary independent genome-wide significant signals were present. Results and additional independent variants are summarized in Table 2 along with findings from *in silico* investigations of the putative functional variant, and in Supplementary Tables S5 and S6. *SCARA5, ABO, VWF* and *STAB2* were predicted to have more than one independent signal both for FVIII and VWF analyses (details in Supplementary Methods and in Supplementary Tables S5 and S6), some of which are in agreement with previous publications[6]. Among the independently associated variants within the *ABO* locus, SNPs rs10901252 and rs687621 perfectly discriminate B and O blood groups from

A, and rs8176685 can reasonably capture information to tag A1/A2 (r2 0.59/D' 0.99 with the tag SNP), confirming that ABO blood groups are essential determinants of VWF and FVIII plasma levels.

Variance Explained

Overall, the top variants for these loci (including the strongest independent associated variants in each locus that reached genome-wide significance after conditional analyses) explain 17% of the phenotypic variance for FVIII and 21.3% of the variance for VWF. *ABO* locus was by far the strongest determinant, alone explaining 13.6% and 16.2% of these variances, respectively.

Functional Analyses

We silenced 21 genes across 12 loci to assess the *in vitro* impact on FVIII and VWF secretion (Figures 2a-b). These include the main candidate genes identified by proximity (Table 1). Our results suggest that 10 of the 12 candidate loci had one or more genes that changed VWF levels in media under basal and/or histamine-stimulated conditions. Specifically, silencing *PDHB, SLC39A8, TMEM171, TNPO1, HLA-C, GIMAP7, NIPSNAP3A, NIPSNAP3B, C2CD4B,* and *SYNGR1* increased VWF release into media under basal conditions whereas *ST3GAL4* silencing decreased VWF secretion. When cells were stimulated with histamine, silencing *TMEM171, TNPO1, HLA-C, SNIPSNAP3A* (but not *SNIPSNAP3B*), *C2CD4B, KAT2A,* and *TAB1* increased VWF release in the media, and *RAB5C* decreased VWF secretion (Table 1; Figures 2a-b). For the experiments on the 5 genes that were only shown to be associated with FVIII levels (*LINC00583, NFIB, SOX17, RP1* and *TMLHE-F8*), we could not find detectable levels of FVIII in media from treated HUVEC cells, and therefore the experiments were inconclusive.

Mendelian Randomization Analyses and Cardiovascular Events

Figure 3 show forest plots representing the results from MR analyses. We first analyzed FVIII and VWF individually using the IVW estimates that included the sentinel variant in each locus (after exclusion of variants with pleiotropic effects, see Supplementary Tables S2, S3, and S4). Both VWF and FVIII plasma levels showed a significant causal effect on CAD, IS and VTE risk. For CAD, the ORs associated with a per SD change in natural log-transformed FVIII and VWF were (OR(CI_{95}) =1.15 (1.05, 1.27) and 1.14 (1.05, 1.23), respectively. For IS, the ORs(CI_{95}) were 1.28 (1.14, 1.43) and 1.19 (1.10, 1.29), respectively. For VTE, the ORs(CI_{95}) were 2.75 (2.14, 3.55) and 2.31 (1.89, 2.81), respectively. Sensitivity analyses using both MR-Egger regression and weighted median estimates support the IVW estimates and no significant pleiotropic effect was evident after exclusion of the pleiotropic loci (Figure 3, Supplementary Table S3, Supplementary Figures S2a-c).

We then performed MVMR analyses of the FVIII phenotype to identify causal effects of FVIII activity levels independent of VWF levels. For VTE and CAD outcomes, adjustment of FVIII results by the effect of VWF, the ORs were modestly attenuated (20% and 16% respectively) compared with the unadjusted estimates and confidence intervals widened. For IS, however, adjustment of FVIII results by the effect of VWF resulted in an 86% attenuation of the OR(CI_{95}) to 0.88 (0.51, 1.51). We could not demonstrate a causal association of VWF levels with VTE and CAD independent of FVIII levels.

Of note, both the *ABO* and *HLA* loci were excluded from the instrumental variables for the MR analyses due to evidence of pleiotropic effects shown in the heterogeneity tests (Supplementary

Table S3). When we estimated causal effects using *ABO* alone as an instrument, estimates of causal effects were essentially the same across phenotypes and outcomes: OR(CI_{95}) 2.57 (2.47-2.67) for FVIII and VTE; OR(CI_{95}) 2.28 (2.18-2.38) for VWF and VTE; OR(CI_{95}) 1.10 (1.06-1.14) for FVIII and IS; OR(CI_{95}) 1.09 (1.05-1.13) for VWF and IS; OR(CI_{95}) 1.10 (1.06-1.14) for FVIII and CAD; and OR(CI_{95}) 1.08 (1.04-1.12) for VWF and CAD.

DISCUSSION

In the present study, we meta-analyzed data from more than 36,000 individuals with FVIII levels and more than 46,000 with VWF and identified 13 novel loci, 7 of which associated with FVIII plasma levels and 11 associated with VWF levels. Overall, new discoveries yielded an additional 6.2% and 8.1% proportion of variance explained for FVIII and VWF respectively from previous estimations[8], and suggest that a great proportion of the genetic variance is explained by common variation. Further, we presented experimental evidence of biological function on VWF regulation for 14 of these genes from gene silencing *in vitro*: *PDHB, SLC39A8, TMEM171, TNPO1, HLA-C, GIMAP7, NIPSNAP3A and B, ST3GAL4, C2CD4B, RAB5C, KAT2A, TAB1, SYNGR1*. Last, we provide evidence in support of a causal role of FVIII levels on VTE and CAD events and of VWF levels on IS events.

Characterization of the Novel Loci Regulating FVIII and VWF

As expected for traits with strong genetic correlation, most of the newly associated loci regulate both FVIII and VWF levels in blood. Our results show that most of the highest-signal independent variants associated with these traits were located in introns or non-coding regions, although a substantial proportion were in strong LD (R^2>0.8) with mutations

causing missense or frameshift mutations in the nearby genes (Table 2 and Supplementary Table S7). We also explored eQTL associations using publically available data and we conducted pathway analyses for the novel loci. See Supplementary Methods and Supplementary Tables S8-S13 for this information.

For most loci, several genes were identified within the associated region, and we selected 1 or more genes for further characterization using *in vitro* cell models. Based on our initial functional characterization, 1 or more plausible culprit genes regulating VWF secretion could be isolated at most loci. Interestingly, several candidate genes that showed a clear change in VWF levels upon silencing have been shown to participate in vesicle trafficking and exocytosis, as well as intracellular signaling and inflammatory response. The most relevant functional genes are described below.

VWF Biogenesis, Vesicle Trafficking and Secretion

ST3GAL4 is a Golgi transferase that catalyzes transfer of sialic acids in VWF glycan branches that are essential to its biogenesis, adhesive activity and clearance[26]. It also has a role in clearance of desialylated platelets with effects on platelet homeostasis. Genetic variants in *ST3GAL4* locus have been associated with total cholesterol, LDL cholesterol, alkaline phosphatase, increased platelet aggregation, fibrinogen, CRP, and CAD (see further details and references in Supplementary Table S7). Our functional analyses showed a substantial reduction of VWF release upon *ST3GAL4* silencing, which strengthens the evidence of this gene as a novel VWF regulator in basal conditions.

SYNGR1 (Synaptogryn-1) encodes an integral membrane protein associated with presynaptic vesicles in neuronal cells. Several commonalities have been described between the exocytic machinery that drives vesicle trafficking and membrane fusion in endothelial cells and synaptic machinery found in neurons [27, 28], which suggest that SYNGR1 could have a role in vesicle trafficking and exocytosis of VWF from the Weibel-Palade bodies. Genetic variation in this locus has also been associated with IgG glycosylation, rheumatoid arthritis, and inflammatory bowel disease/Crohn's disease, the last 2 consistent with an effect of deregulation of interleukin and inflammatory signaling pathways.

NIPSNAP3A and *NIPSNAP3B* were selected as the main biologically plausible genes for locus on chromosome 9, and results from the functional study show evidence of significant upregulated levels of VWF upon silencing of either gene. Again, a reported role of these genes in vesicular trafficking[29] suggest that these genes could be important in Weibel-Palade formation and exocytosis of VWF, both in basal conditions and under inflammatory stimuli.

Among the 2 new loci found in the trans-ethnic multi-phenotype analysis, *RAB5C* is particularly interesting. It is a member of the Rab protein family, thought to ensure fidelity in the process of docking and fusion of vesicles with their correct acceptor compartment[30], which may be relevant to the process of fusion of Weibel-Palate vesicles to release VWF in endothelial cells. *RAB5C* silencing caused a significant decrease of VWF release in media of endothelial cells upon stimulation with histamine.

Our *in vitro* cell work showed a significantly increased VWF secretion upon *PDHB* silencing. *PDHB* codes for a subunit of the pyruvate dehydrogenase complex, which converts pyruvate to acetyl-CoA in the mitochondrion. We speculate that it is possible that the metabolism of endothelial cells regulates vesicle trafficking and exocytosis of VWF, meaning that the exocytosis process is dependent on the energetic status of the endothelial cell. Genetic variation in this locus has also been associated with total cholesterol, SLE, and RA.

Intracellular Signaling and Inflammatory Response

TAB1 silencing increased VWF released in media in our *in vitro* functional analyses, whereas no effect could be seen for *PDGFB,* a gene that has been implicated in CAD and VTE risk. TAB1 is a regulatory protein that acts as a mediator of several intracellular signaling pathways, especially those mediated by TGF-ß, WNT-1 and interleukin-1 which suggest it might have a role mediating VWF release upon certain cellular stimuli.

Similarly, silencing *C2CD4B* gene in cultured endothelial cells resulted in strong upregulation of VWF release both in basal and under stimulus conditions. Allelic variants in this gene have also been associated with fasting glucose homeostasis and type 2 diabetes. Transcripts of this gene are predominantly found in the nuclear compartment of endothelial cells, and a possible role in regulation of transcription that might increase vascular permeability in acute inflammation has been suggested[31]. Similarly, *TNPO1* codes for a nuclear receptor (Transportin-1[32]) which mediates nuclear import of several proteins, which could also suggest a role in regulation of transcription under certain circumstances.

DAB2IP is involved in several relevant cell-signaling pathways in response to inflammation, innate immune response, and cell growth inhibition, apoptosis, cell survival, angiogenesis, cell migration and maturation in endothelial cells, and genetic variation in this gene has been associated with abdominal aortic aneurysm and heart rate. Despite the strong genetic signal in our data, functional confirmation could not be achieved for *DAB2IP* in our secretion experiment so additional investigative work is needed.

GIMAP7 showed a significant increase of VWF release upon silencing. GTPases of immunity-associated proteins (GIMAPs) are regulators of lymphocyte survival and homeostasis [33] although limited data have been published regarding the function of these proteins.

Finally, although it did not reach genome-wide significance in the trans-ethnic meta-analysis, we found a single locus that close to *SLC39A8* and that was genome-wide significant in our meta-analysis VWF associations in European-ancestry samples. This gene, which encodes a zinc transporter that functions in the cellular import of zinc at the onset of inflammation, has also been associated with blood pressure, high-density lipoprotein (HDL) cholesterol levels and BMI. Our functional work also suggested a strong effect on VWF levels in media from endothelial cells *in vitro* upon *SLC39A8* silencing.

Although further functional characterization of these genes is needed to fully characterize the role of all the investigated genes in VWF regulation, our results demonstrate that these studies

are a valid tool to elucidate functional genes coming from genetic associations, and to shed light into the most relevant biological pathways implicated in the regulation of the phenotype under study.

Mendelian Randomization and Clinical Implications

Our results provide insights into the causal role of FVIII and VWF in 3 CV events, which are the leading causes of deaths globally.

Biological and genetic evidence indicate that circulating FVIII levels are mainly determined by levels of VWF[34]. In the present study, we calculated the genetic correlation between VWF and FVIII based on the genome-wide association results from European-descent individuals (see Supplementary Methods) and found that the proportion of shared heritability of between these 2 phenotypes is 83.5%. This result is strengthened by the overlapping findings found in the individual GWAS, and suggests that, with some exceptions, the genetic pathways that regulate VWF levels indirectly regulate FVIII levels. Given the role of VWF regulating FVIII, we used 3 loci that were uniquely associated with FVIII independent of VWF and pursued conditional analyses that adjusted for the effect of VWF plasma levels to test the causal effect of FVIII on CV events. For IS, we found no evidence of a causal effect of FVIII independent of the VWF effect, which suggests that VWF biology may causally contribute to IS risk. For VTE and CAD, however, we found evidence supporting a causal effect of FVIII independent of the VWF effect. As there were no genetic loci that independently associated with VWF levels and not FVIII levels, we could not adjust the VWF analyses for FVIII. Nonetheless, given the similarities in the magnitude of the VWF-adjusted FVIII causal ORs with the VWF causal ORs for VTE and CAD,

our data suggest that the VWF causal association for VTE and CAD may be driven primarily by the biologic effect of FVIII, although this hypothesis could not be tested.

The results of the MR analyses suggest that both FVIII and VWF may be reasonable targets for the prevention or intervention of CAD and VTE while VWF may be a reasonable target for IS. These molecules are not currently pharmaceutical targets for the prevention of thrombotic events. In this paper, we report on 23 unique genetic loci associated with plasma levels of FVIII and/or VWF, of which 13 are newly reported associations. These discoveries offer new potential targets for the development of pharmaceutical agonistics or antagonists that may modulate thrombotic risk.

Strengths and Limitations

A major strength of the study was the relatively large sample size and the use of a denser imputation panel than was used in past discovery efforts. With this approach, we had hoped to identify uncommon associated variants but the MAF of the variants in the newly associated loci where relatively common, with just 1 variant having an MAF of less than 0.10. Our study design did not identify new associations marked by rare variation. Increasing the number of study participants to increase statistical power or improving the quality of the imputation from genotyping arrays may help to identify uncommon or rare variants associated with the outcomes. Some of the novel findings may be false positives, as we did not have access to independent populations to replicate our discoveries. Replication is required to validate genetic associations, especially for those close to the threshold for statistical significance. To offset this limitation, we conducted functional validation by silencing candidate genes and measuring VWF release; we

view this functional work as a strength of the study. We were able to test only the regulation of VWF expression and not the regulation of VWF clearance by macrophages[35]. Nor were we able to test other mechanisms that regulates synthesis in megakaryocytes but not endothelial cells. Further, the need for a particular cellular stimulus that cannot be mimicked by histamine stimulation for the effect to be produced would be missed by our approach. Finally, it could be that the effect of some genetic associations can only be seen through overexpression rather than silencing of the gene. We attempted to also measure FVIII release but levels were too low so new models are required to validate the impact the candidate genes on FVIII levels; this is a limitation of our approach. All functional work was done *in vitro*, which carries limitations relative to *in vivo* investigations. The strong genetic co-regulation of both FVIII and VWF levels allowed us to conduct multi-phenotypes analyses and increase statistical power for discovery. Our MR approach using improved instrumental variants allows to establish for the first time a causal relationship between VWF and FVIII and several CV events. With only 3 loci associated with FVIII alone, the power of the VWF-adjusted MR analyses for FVIII and CV events was limited and we could not investigate the association of VWF on CV events independent of FVIII. There is a degree of overlap between our sample and the sample from consortia providing CV events GWAS data, which might create some bias in MR analyses[36]; this is a limitation of our work.

CONCLUSIONS

We found 13 novel genetic loci with modest contributions to plasma levels of FVIII and/or VWF. Our discovery approach including first-pass functional validation has provided relevant information on the best candidate gene at the novel loci. Finally, MR analyses provided some

evidence implicating FVIII plasma levels in the risk of CAD and VTE, and VWF plasma levels in the risk of IS. In summary, our work has identified novel loci regulating proteins essential for hemostasis and coagulation. These findings may provide genetic tools for therapeutic and preventive strategies and may be useful to identify new biologic pathways upon which to intervene to reduce the burden of arterial and venous outcomes.

Authors:

Maria Sabater-Lleal*, Jennifer E. Huffman*, Paul S. de Vries*, Jonathan Marten*, Michael A. Mastrangelo, Ci Song, Nathan Pankratz, Cavin Ward-Caviness, Lisa R. Yanek, Stella Trompet, Graciela Delgado, Xiuqing Guo, Traci M. Bartz, Angel Martinez-Perez, Marine Germain, Hugoline G. de Haan, Ayse B. Ozel, Ozren Polasek, Albert V. Smith, John D. Eicher, Alex P Reiner, Weihong Tang, Neil Davies, David J. Stott, Jerome I. Rotter, Geoffrey H. Tofler, Eric Boerwinkle, Moniek PM. de Maat, Paul Welsh, Marcus E. Kleber, Jennifer A. Brody, Ming-Huei Chen, Dhananjay Vaidya, José Manuel Soria, Pierre Suchon, Astrid van Hylckama Vlieg, Karl C. Desch, Ivana Kolcic, Peter K. Joshi, Lenore J. Launer, Tamara B.Harris, Harry Campbell, Igor Rudan, Diane M. Becker, Jun Z. Li, Fernando Rivadeneira, André G. Uitterlinden, Albert Hofman, Oscar H. Franco, Mary Cushman, Bruce M. Psaty, Pierre-Emmanuel Morange, Barbara McKnight, Michael R. Chong, Israel Fernandez-Cadenas, Jonathan Rosand, Arne Lindgren, INVENT Consortium, MEGASTROKE consortium of the International Stroke Genetics Consortium (ISGC), Vilmundur Gudnason, David- James F. Wilson, Caroline Hayward, David Ginsburg, Myriam Fornage, Frits R. Rosendaal, Juan Carlos Souto, Lewis C. Becker, Nancy S. Jenny, Winfried März, J. Wouter Jukema, Abbas Dehghan, David-Alexandre Trégouët, Alanna C. Morrison, Andrew D. Johnson, Christopher O'Donnell, David P. Strachan, Charles J. Lowenstein[#], Nicholas L. Smith[#]

*The authors contributed equally to this work as first authors

[#]The authors contributed equally to this work as last authors

Affiliations:

Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet; Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden (M.S-L)

Unit of Genomics of Complex Diseases. Institut d'Investigació Biomèdica Sant Pau. IIB-Sant Pau. Barcelona, Spain (M.S-L, A.M-P, J.M.S)

Population Sciences Branch, National Heart, Lung, and Blood Institute. The Framingham Heart Study, Framingham, MA, USA (J.E.H, C.S., M-H.C, J.D.E, A.D.J, C.J.O'D)

Human Genetics Center, Department of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA (P.S.dV, A.C.M, E.B, M.F)
Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands (P.S.dV, A.H, O.H.F, A.D)
Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scothland (J.M, J.W,C.H)
Aab Cardiovascular Research Institute, University of Rochester Medical Center, Rochester, NY, USA (M.A.M, C.J.L)
Department of Laboratory Medicine and Pathology, University of Minnesota School of Medicine, Minneapolis, MN, USA (N.P)
Environmental Public Health Division, National Health and Environmental Effects Research Laboratory, US Environmental Protection Agency, Chapel Hill, NC, USA (C.W-C)
GeneSTAR Research Program, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA (L.R.Y, D.V, D.M.B, L.C.B)
Department of Geriatrics and Gerontology, Leiden University Medical Center, Leiden, the Netherlands (S.T)
Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands (S.T, J.W.J)
Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany (G.E.D, M.E.K, W.M)
The Institute for Translational Genomics and Population Sciences, Department of Pediatrics and Medicine, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA (X.G, J.I.R)
Department of Biostatistics, University of Washington, Seattle WA USA (T.M.B, B.McK)
Sorbonne Universités, UPMC Univ. INSERM, UMR_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases; ICAN Institute for Cardiometabolism and Nutrition, Paris, France (M.G, D-A.T)
Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands (H.G.dH, A.vHV, F.R.R)
Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA (A.B.O, J.Z.L, D.G)

Faculty of Medicine, University of Split, Split, Croatia (O.P, I.K)
School of Public Health, Department of Biostatistics, University of Michigan, Ann Arbor, MI, 48109, USA (A.V.S)
Fred Hutchinson Cancer Research Center (A.P.R)
Department of Epidemiology, University of Washington (A.P.R, N.L.S)
Division of Epidemiology and Community Health, University of Minnesota School of Public Health (W.T)
Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom (N.M.D)
Bristol Medical School, University of Bristol, Bristol, United Kingdom (N.M.D)
Academic Section of Geriatrics, Faculty of Medicine, University of Glasgow, Glasgow, United Kingdom (D.J.S)
Royal North Shore Hospital, University of Sydney, Sydney, Australia (G.H.T)
Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA (E.B)
Department of Hematology, Erasmus University Medical Center, Rotterdam, the Netherlands (M.PM.dM)
Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, United Kingdom (P.W)
Institute of Nutrition, Friedrich-Schiller-University Jena, Mannheim, Germany (M.E.K)
Department of Medicine, Epidemiology, and Health Services, University of Washington, Seattle WA USA (J.A.B, B.M.P)
Laboratory of Haematology, La timone Hospital, Marseille, France (P.S, P-E.M)
INSERM, UMR_S 1062, Nutrition Obesity and Risk of Thrombosis, Marseille, France (P.S, P-E.M)
Department of Pediatrics and Communicable Disease, University of Michigan, Ann Arbor, MI, USA (K.C.D)
Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, Scotland (P.K.J, H.C, I.R, J.F.W)
Laboratory of Epidemiology and Population Sciences National Institute on Aging, Bethesda, MD, USA (L.J.L, T.B.H)

Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands (F.R, A.G.U)
Department of Epidemiology, Harvard H.T. Chan School of Public Health, Boston, MA, USA (A.H)
Larner College of Medicine, University of Vermont, Colchester, VT, USA (M.C)
Kaiser Permanente Washington Research Institute, Kaiser Permanente Washington, Seattle, WA, USA (B.M.P, N.L.S)
Cardiovascular Health Research Unit, University of Washington, Seattle, WA (B.McK)
McMaster University, Population Health Research Institute, Population Health Research Institute, Biochemistry and Biomedical Sciences. Hamilton, Ontario, Canada (M.R.C)
Stroke Pharmacogenomics and genetics, Department of Neurology. IIB-Sant Pau, Barcelona, Spain (I.F-C)
Massachusetts General Hospital, Broad Institute, Harvard Medical School (J.R)
Department of Clinical Sciences, Lund University (A.L)
Department of Neurology and Rehabilitation Medicine, Neurology, Skåne University Hospital (A.L)
Icelandic Heart Association, Kopavogur, Iceland (V.G)
Faculty of Medicine, University of Iceland, Reykjavik, Iceland (V.G)
Brown Foundation Institute of Molecular Medicine, The University of Texas Health Science Center, Houston, TX, USA (M.F)
Einthoven Laboratory of Experimental Vascular Medicine, Leiden University Medical Center, Leiden, The Netherlands (F.R.R, J.W.J)
Unit of Hemostasis and Thrombosis. Hospital de la Sant Creu i Sant Pau. Barcelona, Spain (J.C.S)
Department of Pathology and Laboratory Medicine, The University of Vermont College of Medicine, Colchester, VT, USA (N.S.J)
Synlab Academy, Synlab Holding Deutschland GmbH, Mannheim, Germany (W.M)
Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University Graz, Mannheim, Germany (W.M)
Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands (J.W.J)
Department of Epidemiology and Biostatistics, Imperial College London, London, UK (A.D)

Cardiology Section Administration, Boston VA Healthcare System, West Roxbury, MA, USA (C.J.O'D)
Population Health Research Institute, St George's, University of London, London, UK (D.P.S)
Seattle Epidemiologic Research and Information Center, Department of Veterans Affairs Office of Research and Development, Seattle WA, USA (N.L.S)

Acknowledgements of the INVENT Consortium and MEGASTROKE consortium of the International Stroke Genetics Consortium (ISGC) can be found in the Supplementary Materials.

Disclosures

CHS

Psaty serves on the DSMB of a clinical trial funded by Zoll LifeCor and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson.

GABC and TSS

David Ginsburg: Shire: Equity Ownership, Membership on an entity's Board of Directors or advisory committees, Patents & Royalties: recombinant VWF and recombinant ADAMTS13; Portola Pharmaceuticals: Membership on an entity's Board of Directors or advisory committees.

LURIC

W.M. reports grants and personal fees from AMGEN, BASF, Sanofi, Siemens Diagnostics, Aegerion Pharmaceuticals, Astrazeneca, Danone Research, Numares, Pfizer, Hoffmann LaRoche: personal fees from MSD, Alexion; grants from Abbott Diagnostics, all outside the submitted work. W.M. is employed with synlab Holding Deutschland GmbH.

Disclosures for **AGES**, **ARIC**, **B58C**, **CARDIA, CROATIA-VIS**, **FHS**, **GAIT**, **GeneSTAR**, **MARTHA**, **MEGA**, **MESA**, **ORCADES**, **PROSPER**, and **Rotterdam Study**; there are no conflicts of interest to declare.

Sources of Funding

This study is supported in part by the National Heart, Lung, and Blood Institute (NHLBI) grant HL134894. **The views expressed in this manuscript 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; or the U.S. Department of Health and Human Services.** Maria Sabater-Lleal was partially supported by a EHA-ISTH fellowship to study genetic determinants of FVIII and VWF, by the Swedish Heart-Lung Foundation (20160290) and is a recipient of a Miguel Servet contract from the Spanish Ministry of Health (ISCIII CP17/00142). Paul S. de Vries is supported by American Heart Association Grant #17POST33350042.

The **Age, Gene/Environment Susceptibility** (**AGES**)--Reykjavik Study has been funded by NIH contract N01-AG-12100, the NIA Intramural Research Program (Z01AG007380), Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063.

Atherosclerosis Risk in Communities (ARIC) study is carried out as a collaborative study supported by the National Heart, Lung, and Blood Institute (NHLBI) contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly

supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. The authors thank the staff and participants of the ARIC study for their important contributions. Funding support for “Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium” was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419).

We acknowledge use of phenotype and genotype data from the **British 1958 Birth Cohort (B58C)** DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). The B58C-GABRIEL genotyping was supported by a contract from the European Commission Framework Programme 6 (018996) and grants from the French Ministry of Research.

The **Coronary Artery Risk Development in Young Adults Study (CARDIA)** is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201300025C & HSN268201300026C), Northwestern University (HHSN268201300027C), University of Minnesota (HHSN268201300028C), Kaiser Foundation Research Institute (HHSN268201300029C), and Johns Hopkins University School of Medicine (HHSN268200900041C). CARDIA is also partially supported by the Intramural Research Program of the National Institute on Aging NIA) and an intra-agency agreement between NIA and NHLBI (AG0005). Genotyping was funded as part of the NHLBI Candidate-gene Association Resource (N01-HC-65226) and the NHGRI Gene Environment Association Studies (GENEVA) (U01-HG004729, U01-HG04424, and U01-HG004446). This manuscript has been reviewed and approved by CARDIA for scientific content.

Cardiovascular Health Study (CHS) research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268200960009C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL085251, R01HL087652, R01HL105756, R01HL103612, R01HL120393, and R01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping

data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

The **CROATIA-Vis** study was funded by grants from the Medical Research Council (UK) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, the Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health. The SNP genotyping for the CROATIA-Vis cohort was performed in the core genotyping laboratory of Clinical Research Facility at the Western General Hospital, Edinburgh, Scotland

Framingham Heart Study (FHS) was partially supported by the National Heart, Lung, and Blood Institute's (NHLBI's) Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc. for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. Geoffrey Tofler acknowledges funding from the National Institutes of Health (RO1-HL-48157). The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. The views expressed in this manuscript 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; or the U.S. Department of Health and Human Services.

The **Genes and Blood Clotting Study (GABC)** was funded by the National Institutes of Health grants, R37HL039693 and RO1HL112642. David Ginsburg is a Howard Hughes Investigator.

Genetic Analysis for Idiopathic Thrombophilia (GAIT) project was supported partially by grants PI-11/0184, PI-14/0582 and Red Investigación Cardiovascular RD12/0042/0032 from the Instituto Carlos III (Fondo de Investigación Sanitaria–FIS), and 2014SGR-402 Consolidated Research Group of the Generalitat de Catalunya.

Genetic Study of Atherosclerosis Risk (GeneSTAR) was supported by grants from the National Institutes of Health/National Heart, Lung, and Blood Institute (U01 HL72518, HL087698) and by a grant from the National Institutes of Health/National Center for Research Resources (M01-RR000052) to the Johns Hopkins General Clinical Research Center.

We thank the **LURIC study** team who were either temporarily or permanently involved in patient recruitment as well as sample and data handling, in addition to the laboratory staff at the Ludwigshafen General Hospital and the Universities of Freiburg and Ulm, Germany. LURIC was supported by the 7th Framework Program RiskyCAD (grant agreement number

305739) of the European Union. The work of W.M. and M.E.K. is supported as part of the Competence Cluster of Nutrition and Cardiovascular Health (nutriCARD) which is funded by the German Federal Ministry of Education and Research.

The **MARTHA** project was supported by grants from the Program Hospitalier de Recherche Clinique. MARTHA genetics research programs are supported and funded by the GenMed LABEX (ANR-10-LBX-0013), the ICAN Institute for Cardiometabolism and Nutrition (ANR-10-IHU-05) and the French INvestigation Network on Venous Thrombo-Embolism (INNOVTE). Statistical analysis of the MARTHA data were performed on the C2BIG computing cluster funded by the Région Ile de France, Pierre and Marie Curie University and the ICAN Institute for Cardiometabolism and Nutrition.

The **MEGA** study was supported by Netherlands Heart Foundation (NHS 98.113), the Dutch Cancer Foundation (RUL 99/1992), the Netherlands Organisation for Scientific Research (912–03–033| 2003), and partially by the GenMed LABEX (ANR-10-LABX-0013). We would like to thank all colleagues from the French Centre National de Génotypage for the genotyping contribution.

MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-

001079, UL1-TR-001420, UL1-TR-001881, and DK063491. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, California, USA) and the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

The **Orkney Complex Disease Study (ORCADES)** was supported by the Chief Scientist Office of the Scottish Government (CZB/4/276, CZB/4/710), the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

The **Prospective Study of Pravastatin in the Elderly at Risk (PROSPER)** trial was supported by an investigator-initiated grant from Bristol-Myers Squibb, USA. The study was conducted, analysed and reported independently of the company. The GWAS project PHASE has received funding from the European Union's Seventh Framework Programme (FP7/2007–2013) under grant agreement HEALTH-F2-2009-223004. A part of the genotyping was funded by The Netherlands Consortium for Healthy Ageing (NGI: 05060810). Professor Dr J.W.J. is an

established clinical investigator of The Netherlands Heart Foundation (2001 D 032). This work was performed as part of an ongoing collaboration of the PROSPER study group in the universities of Leiden, Glasgow and Cork.

The generation and management of GWAS genotype data for the **Rotterdam Study (RS)** is supported by the Netherlands Organisation of Scientific Research NWO Investments (no. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project no. 050-060-810. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera, Marjolein Peters and Carolina Medina-Gomez for their help in creating the Rotterdam Study GWAS database, and Karol Estrada and Carolina Medina-Gomez for the creation and analysis of imputed data.

The **Trinity Student Study (TSS)** was supported by the Intramural Research Programs of the National Institutes of Health, the Eunice Kennedy Shriver National Institute of Child Health and Human Development and the National Human Genome Research Institute. The principle investigators for the TSS study were Lawrence C. Brody (Genome Technology Branch,

National Human Genome Research Institute, National Institutes of Health**),** James Mills (Eunice Kennedy Shriver National Institute of Child Health and Human Development) and Anne M. Molloy (School of Medicine, Trinity College, Dublin, Ireland) who provided Karl Desch and David Ginsburg plasma samples for the determination of VWF concentrations and de-identified individual level genotyping data which was used in this study.

The Medical Research Council (MRC) and the University of Bristol support the MRC Integrative Epidemiology Unit [MC_UU_12013/1, MC_UU_12013/9].

Acknowledgments

We thank Martin Dichgans, Stéphanie Debette, and Rainer Malik for their assistance in coordinating access to data from the MEGASTROKE consortium.

REFERENCES:

1. Folsom AR. Hemostatic risk factors for atherothrombotic disease: An epidemiologic view. *Thromb Haemost*. 2001;86:366-373
2. Koster T, Blann AD, Briet E, Vandenbroucke JP, Rosendaal FR. Role of clotting factor viii in effect of von willebrand factor on occurrence of deep-vein thrombosis. *Lancet*. 1995;345:152-155
3. Smith A, Patterson C, Yarnell J, Rumley A, Ben-Shlomo Y, Lowe G. Which hemostatic markers add to the predictive value of conventional risk factors for coronary heart disease and ischemic stroke? The caerphilly study. *Circulation*. 2005;112:3080-3087
4. Spiel AO, Gilbert JC, Jilma B. Von willebrand factor in cardiovascular disease: Focus on acute coronary syndromes. *Circulation*. 2008;117:1449-1459
5. Seaman CD, Yabes J, Comer DM, Ragni MV. Does deficiency of von willebrand factor protect against cardiovascular disease? Analysis of a national discharge register. *J Thromb Haemost*. 2015;13:1999-2003
6. Huffman JE, de Vries PS, Morrison AC, Sabater-Lleal M, Kacprowski T, Auer PL, Brody JA, Chasman DI, Chen MH, Guo X, Lin LA, Marioni RE, Muller-Nurasyid M, Yanek LR, Pankratz N, Grove ML, de Maat MP, Cushman M, Wiggins KL, Qi L, Sennblad B, Harris SE, Polasek O, Riess H, Rivadeneira F, Rose LM, Goel A, Taylor KD, Teumer A, Uitterlinden AG, Vaidya D, Yao J, Tang W, Levy D, Waldenberger M, Becker DM, Folsom AR, Giulianini F, Greinacher A, Hofman A, Huang CC, Kooperberg C, Silveira A, Starr JM, Strauch K, Strawbridge RJ, Wright AF, McKnight B, Franco OH, Zakai N, Mathias RA, Psaty BM, Ridker PM, Tofler GH, Volker U, Watkins H, Fornage M, Hamsten A, Deary IJ, Boerwinkle E, Koenig W, Rotter JI, Hayward C, Dehghan A, Reiner AP, O'Donnell CJ, Smith NL. Rare and low-frequency variants and their association with plasma levels of fibrinogen, fvii, fviii, and vwf. *Blood*. 2015;126:e19-29
7. Johnsen JM, Auer PL, Morrison AC, Jiao S, Wei P, Haessler J, Fox K, McGee SR, Smith JD, Carlson CS, Smith N, Boerwinkle E, Kooperberg C, Nickerson DA, Rich SS, Green D, Peters U, Cushman M, Reiner AP, Project NES. Common and rare von willebrand factor (vwf) coding variants, vwf levels, and factor viii levels in african americans: The nhlbi exome sequencing project. *Blood*. 2013;122:590-597
8. Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, Hayward C, Rudan I, Sabater-Lleal M, Bis JC, de Maat MP, Rumley A, Kong X, Yang Q, Williams FM, Vitart V, Campbell H, Malarstig A, Wiggins KL, Van Duijn CM, McArdle WL, Pankow JS, Johnson AD, Silveira A, McKnight B, Uitterlinden AG, Wellcome Trust Case Control C, Aleksic N, Meigs JB, Peters A, Koenig W, Cushman M, Kathiresan S, Rotter JI, Bovill EG, Hofman A, Boerwinkle E, Tofler GH, Peden JF, Psaty BM, Leebeek F, Folsom AR, Larson MG, Spector TD, Wright AF, Wilson JF, Hamsten A, Lumley T, Witteman JC, Tang W, O'Donnell CJ. Novel associations of multiple genetic loci with plasma levels of factor vii, factor viii, and von willebrand factor: The charge (cohorts for heart and aging research in genome epidemiology) consortium. *Circulation*. 2010;121:1382-1392
9. Zhu Q, Yamakuchi M, Ture S, de la Luz Garcia-Hernandez M, Ko KA, Modjeski KL, LoMonaco MB, Johnson AD, O'Donnell CJ, Takai Y, Morrell CN, Lowenstein CJ.

Syntaxin-binding protein stxbp5 inhibits endothelial exocytosis and promotes platelet secretion. *J Clin Invest*. 2014;124:4503-4516
10. Rydz N, Swystun LL, Notley C, Paterson AD, Riches JJ, Sponagle K, Boonyawat B, Montgomery RR, James PD, Lillicrap D. The c-type lectin receptor clec4m binds, internalizes, and clears von willebrand factor and contributes to the variation in plasma von willebrand factor levels. *Blood*. 2013;121:5228-5237
11. Germain M, Chasman DI, de Haan H, Tang W, Lindstrom S, Weng LC, de Andrade M, de Visser MC, Wiggins KL, Suchon P, Saut N, Smadja DM, Le Gal G, van Hylckama Vlieg A, Di Narzo A, Hao K, Nelson CP, Rocanin-Arjo A, Folkersen L, Monajemi R, Rose LM, Brody JA, Slagboom E, Aissi D, Gagnon F, Deleuze JF, Deloukas P, Tzourio C, Dartigues JF, Berr C, Taylor KD, Civelek M, Eriksson P, Cardiogenics C, Psaty BM, Houwing-Duitermaat J, Goodall AH, Cambien F, Kraft P, Amouyel P, Samani NJ, Basu S, Ridker PM, Rosendaal FR, Kabrhel C, Folsom AR, Heit J, Reitsma PH, Tregouet DA, Smith NL, Morange PE. 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
12. Hinds DA, Buil A, Ziemek D, Martinez-Perez A, Malik R, Folkersen L, Germain M, Malarstig A, Brown A, Soria JM, Dichgans M, Bing N, Franco-Cereceda A, Souto JC, Dermitzakis ET, Hamsten A, Worrall BB, Tung JY, Metastroke Consortium, Invent Consortium, Sabater-Lleal M. 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
13. Smith NL, Rice KM, Bovill EG, Cushman M, Bis JC, McKnight B, Lumley T, Glazer NL, van Hylckama Vlieg A, Tang W, Dehghan A, Strachan DP, O'Donnell CJ, Rotter JI, Heckbert SR, Psaty BM, Rosendaal FR. Genetic variation associated with plasma von willebrand factor levels and the risk of incident venous thrombosis. *Blood*. 2011;117:6007-6011
14. de Vries PS, Sabater-Lleal M, Chasman DI, Trompet S, Ahluwalia TS, Teumer A, Kleber ME, Chen MH, Wang JJ, Attia JR, Marioni RE, Steri M, Weng LC, Pool R, Grossmann V, Brody JA, Venturini C, Tanaka T, Rose LM, Oldmeadow C, Mazur J, Basu S, Franberg M, Yang Q, Ligthart S, Hottenga JJ, Rumley A, Mulas A, de Craen AJ, Grotevendt A, Taylor KD, Delgado GE, Kifley A, Lopez LM, Berentzen TL, Mangino M, Bandinelli S, Morrison AC, Hamsten A, Tofler G, de Maat MP, Draisma HH, Lowe GD, Zoledziewska M, Sattar N, Lackner KJ, Volker U, McKnight B, Huang J, Holliday EG, McEvoy MA, Starr JM, Hysi PG, Hernandez DG, Guan W, Rivadeneira F, McArdle WL, Slagboom PE, Zeller T, Psaty BM, Uitterlinden AG, de Geus EJ, Stott DJ, Binder H, Hofman A, Franco OH, Rotter JI, Ferrucci L, Spector TD, Deary IJ, Marz W, Greinacher A, Wild PS, Cucca F, Boomsma DI, Watkins H, Tang W, Ridker PM, Jukema JW, Scott RJ, Mitchell P, Hansen T, O'Donnell CJ, Smith NL, Strachan DP, Dehghan A. Comparison of hapmap and 1000 genomes reference panels in a large-scale genome-wide association study. *PLoS One*. 2017;12:e0167742
15. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, Uitterlinden AG, Harris TB, Witteman JC, Boerwinkle E, Consortium C. Cohorts for heart and aging research in genomic epidemiology (charge) consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet*. 2009;2:73-80

16. 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56-65
17. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Magi R, Ferreira T, Fall T, Graff M, Justice AE, Luan J, Gustafsson S, Randall JC, Vedantam S, Workalemahu T, Kilpelainen TO, Scherag A, Esko T, Kutalik Z, Heid IM, Loos RJ, Genetic Investigation of Anthropometric Traits C. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc*. 2014;9:1192-1212
18. Willer CJ, Li Y, Abecasis GR. Metal: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190-2191
19. Fadista J, Manning AK, Florez JC, Groop L. The (in)famous gwas p-value threshold revisited and updated for low-frequency variants. *Eur J Hum Genet*. 2016;24:1202-1205
20. Yang J, Lee SH, Goddard ME, Visscher PM. Gcta: A tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76-82
21. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB, Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimaki T, Loos RJF, Melander O, Metspalu A, Marz W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M. A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47:1121-1130
22. Loley C, Alver M, Assimes TL, Bjonnes A, Goel A, Gustafsson S, Hernesniemi J, Hopewell JC, Kanoni S, Kleber ME, Lau KW, Lu Y, Lyytikainen LP, Nelson CP, Nikpay M, Qu L, Salfati E, Scholz M, Tukiainen T, Willenborg C, Won HH, Zeng L, Zhang W, Anand SS, Beutner F, Bottinger EP, Clarke R, Dedoussis G, Do R, Esko T, Eskola M, Farrall M, Gauguier D, Giedraitis V, Granger CB, Hall AS, Hamsten A, Hazen SL, Huang J, Kahonen M, Kyriakou T, Laaksonen R, Lind L, Lindgren C, Magnusson PK, Marouli E, Mihailov E, Morris AP, Nikus K, Pedersen N, Rallidis L, Salomaa V, Shah SH, Stewart AF, Thompson JR, Zalloua PA, Chambers JC, Collins R, Ingelsson E, Iribarren C, Karhunen PJ, Kooner JS, Lehtimaki T, Loos RJ, Marz W, McPherson R,

Metspalu A, Reilly MP, Ripatti S, Sanghera DK, Thiery J, Watkins H, Deloukas P, Kathiresan S, Samani NJ, Schunkert H, Erdmann J, Konig IR. No association of coronary artery disease with x-chromosomal variants in comprehensive international meta-analysis. *Scientific reports*. 2016;6:35278

23. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese AK, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD, Jr., Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen WM, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, den Hoed M, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu FC, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jimenez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee JM, Lemmens R, Leys D, Lewis CM, Lin WY, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmae K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rost NS, Rothwell PM, Rotter JI, Rundek T, Sacco RL, Sakaue S, Sale MM, Salomaa V, Sapkota BR, Schmidt R, Schmidt CO, Schminke U, Sharma P, Slowik A, Sudlow CLM, Tanislav C, Tatlisumak T, Taylor KD, Thijs VNS, Thorleifsson G, Thorsteinsdottir U, Tiedt S, Trompet S, Tzourio C, van Duijn CM, Walters M, Wareham NJ, Wassertheil-Smoller S, Wilson JG, Wiggins KL, Yang Q, Yusuf S, Bis JC, Pastinen T, Ruusalepp A, Schadt EE, Koplev S, Bjorkegren JLM, Codoni V, Civelek M, Smith NL, Tregouet DA, Christophersen IE, Roselli C, Lubitz SA, Ellinor PT, Tai ES, Kooner JS, Kato N, He J, van der Harst P, Elliott P, Chambers JC, Takeuchi F, Johnson AD, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT, Jr., Rolfs A, Hata J, Woo D, Rosand J, Pare G, Hopewell JC, Saleheen D, Stefansson K, Worrall BB, Kittner SJ, Seshadri S, Fornage M, Markus HS, Howson JMM, Kamatani Y, Debette S, Dichgans M, Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese AK, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD, Jr., Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen WM, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, Hoed MD, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu FC, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jimenez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee JM, Lemmens R, Leys D, Lewis CM, Lin WY, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmae K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rost NS, Rothwell PM, Rotter JI, Rundek T, Sacco RL, Sakaue S, Sale MM, Salomaa V, Sapkota BR, Schmidt R, Schmidt CO, Schminke U, Sharma P, Slowik A, Sudlow CLM, Tanislav C, Tatlisumak T, Taylor KD, Thijs VNS, Thorleifsson G, Thorsteinsdottir U, Tiedt S, Trompet S, Tzourio C, van Duijn CM,

Walters M, Wareham NJ, Wassertheil-Smoller S, Wilson JG, Wiggins KL, Yang Q, Yusuf S, Amin N, Aparicio HS, Arnett DK, Attia J, Beiser AS, Berr C, Buring JE, Bustamante M, Caso V, Cheng YC, Choi SH, Chowhan A, Cullell N, Dartigues JF, Delavaran H, Delgado P, Dorr M, Engstrom G, Ford I, Gurpreet WS, Hamsten A, Heitsch L, Hozawa A, Ibanez L, Ilinca A, Ingelsson M, Iwasaki M, Jackson RD, Jood K, Jousilahti P, Kaffashian S, Kalra L, Kamouchi M, Kitazono T, Kjartansson O, Kloss M, Koudstaal PJ, Krupinski J, Labovitz DL, Laurie CC, Levi CR, Li L, Lind L, Lindgren CM, Lioutas V, Liu YM, Lopez OL, Makoto H, Martinez-Majander N, Matsuda K, Minegishi N, Montaner J, Morris AP, Muino E, Muller-Nurasyid M, Norrving B, Ogishima S, Parati EA, Peddareddygari LR, Pedersen NL, Pera J, Perola M, Pezzini A, Pileggi S, Rabionet R, Riba-Llena I, Ribases M, Romero JR, Roquer J, Rudd AG, Sarin AP, Sarju R, Sarnowski C, Sasaki M, Satizabal CL, Satoh M, Sattar N, Sawada N, Sibolt G, Sigurdsson A, Smith A, Sobue K, Soriano-Tarraga C, Stanne T, Stine OC, Stott DJ, Strauch K, Takai T, Tanaka H, Tanno K, Teumer A, Tomppo L, Torres-Aguila NP, Touze E, Tsugane S, Uitterlinden AG, Valdimarsson EM, van der Lee SJ, Volzke H, Wakai K, Weir D, Williams SR, Wolfe CDA, Wong Q, Xu H, Yamaji T, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT, Jr., Rolfs A, Hata J, Woo D, Rosand J, Pare G, Hopewell JC, Saleheen D, Stefansson K, Worrall BB, Kittner SJ, Seshadri S, Fornage M, Markus HS, Howson JMM, Kamatani Y, Debette S, Dichgans M, Consortium AF, Cohorts for H, Aging Research in Genomic Epidemiology C, International Genomics of Blood Pressure C, Consortium I, Starnet, BioBank Japan Cooperative Hospital G, Consortium C, Consortium E-C, Consortium EP-I, International Stroke Genetics C, Consortium M, Neurology Working Group of the CC, Network NSG, Study UKYLD, Consortium M, Consortium M. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet*. 2018; 50:524-537.

24. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. *Epidemiology*. 2017;28:30-42

25. Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkila K, Hypponen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikainen LP, Magnusson PK, Mangino M, Mihailov E, Montasser ME, Muller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney AS, Doring A, Elliott P, Epstein SE, Eyjolfsson GI, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJ, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimaki T, Lin SY, Lindstrom J, Loos RJ, Mach F, McArdle WL, Meisinger C, Mitchell BD, Muller G, Nagaraja R, Narisu N, Nieminen TV, Nsubuga RN, Olafsson I,

Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stancakova A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YD, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrieres J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllensten U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Jarvelin MR, Jula A, Kahonen M, Kaprio J, Kesaniemi A, Kivimaki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, Marz W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njolstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PE, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BH, Altshuler D, Ordovas JM, Boerwinkle E, Palmer CN, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Mohlke KL, Ingelsson E, Abecasis GR, Daly MJ, Neale BM, Kathiresan S. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet*. 2013;45:1345-1352

26. Song J, Xue C, Preisser JS, Cramer DW, Houck KL, Liu G, Folsom AR, Couper D, Yu F, Dong JF. Association of single nucleotide polymorphisms in the st3gal4 gene with vwf antigen and factor viii activity. *PLoS One*. 2016;11:e0160757

27. Sudhof TC. The synaptic vesicle cycle: A cascade of protein-protein interactions. *Nature*. 1995;375:645-653

28. Lowenstein CJ, Morrell CN, Yamakuchi M. Regulation of weibel-palade body exocytosis. *Trends Cardiovasc Med*. 2005;15:302-308

29. Buechler C, Bodzioch M, Bared SM, Sigruener A, Boettcher A, Lapicka-Bodzioch K, Aslanidis C, Duong CQ, Grandl M, Langmann T, Dembinska-Kiec A, Schmitz G. Expression pattern and raft association of nipsnap3 and nipsnap4, highly homologous proteins encoded by genes in close proximity to the atp-binding cassette transporter a1. *Genomics*. 2004;83:1116-1124

30. Han HJ, Sudo K, Inazawa J, Nakamura Y. Isolation and mapping of a human gene (rabl) encoding a small gtp-binding protein homologous to the ras-related rab gene. *Cytogenetics and cell genetics*. 1996;73:137-139

31. Warton K, Foster NC, Gold WA, Stanley KK. A novel gene family induced by acute inflammation in endothelial cells. *Gene*. 2004;342:85-95

32. Twyffels L, Gueydan C, Kruys V. Transportin-1 and transportin-2: Protein nuclear import and beyond. *FEBS Lett*. 2014;588:1857-1868

33. Krucken J, Schroetel RM, Muller IU, Saidani N, Marinovski P, Benten WP, Stamm O, Wunderlich F. Comparative analysis of the human gimap gene cluster encoding a novel gtpase family. *Gene*. 2004;341:291-304

34. Lacroix-Desmazes S, Repesse Y, Kaveri SV, Dasgupta S. The role of vwf in the immunogenicity of fviii. *Thromb Res*. 2008;122 Suppl 2:S3-6
35. O'Sullivan JM, Aguila S, McRae E, Ward SE, Rawley O, Fallon PG, Brophy TM, Preston RJ, Brady L, Sheils O, Chion A, O'Donnell JS. N-linked glycan truncation causes enhanced clearance of plasma-derived von willebrand factor. *J Thromb Haemost*. 2016;14:2446-2457
36. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample mendelian randomization. *Genet Epidemiol*. 2016;40:597-608

Table 1: Main association results for FVIII and VWF trans-ethnic GWAS meta-analysis.

Footnotes: "*Freq"* refers to the allele frequency of the effect allele. "B" = baseline; "S" = stimulated; "ns" = not significant; *p-value<0.05; **p-value<0.01; ***p-value<0.001.
Beta and Frequency refer always to the Effect_allele, and they are expressed as natural-log transformed values from the original units (reported in % or IU/ml*100 units).
†SLC39A8 was found in vWF meta-analysis of EA only (N=42,145).
‡ Although not in LD with this variant, a low-frequency variant 665Kb upstream rs9271597 was found significantly associated to vWF levels (rs80082277; p=1x10-8) and we consider it within the HLA region; thus, we pursued this gene for further functional validation.
§Olfactory receptor family was not considered for further functional validation for its low expression in the relevant tissues (mainly artery and whole blood).
|| The *ST3GAL4* locus was new at the time of analyses, although reported in a recent candidate gene study lacking replication (PMID: 27584569).
The highest associated SNP in this locus for FVIII is rs137631 (p=9.5x10-9), located close to **RPL3 gene**, 112Kb downstream TAB1/SYNGR1 locus and in low LD with rs5750823 (R^2=0.14)
¶ Chromosome X variant for VWF only available for EU samples (N=28.685).

Table 2: Characteristics of all associated independent variants after conditional analyses.

Footnotes: †LD with top variant in the region, calculated using FHS data. ‡Primary SNP was not well imputed in FHS and no other SNPs in the region achieved genome-wide significance in conditional analyses. "slct pJ" = joint p-value from GCTA "slct". "Original p-value" = p-value from discovery meta-analysis. The putative functional column indicates the best candidate variant in high linkage disequilibrium with the associated variant (R^2>0.8) that has been identified *in silico* as the best candidate variant to have an impact on the adjacent gene/s. No functional work was performed in known genes, and these are symbolized by "-" in the last column.

Table 1													
rsID	**Effect Allele**	**Other Allele**	**Freq FVIII**	**N FVIII**	**Beta FVIII**	**P FVIII**	**Freq VWF**	**N VWF**	**Beta VWF**	**P VWF**	**Closest Gene(s)**	**Association**	**vWF Release at Baseline (B) and upon Stimulation (S) for Candidate Loci**
rs55954186	a	g	0.36	35,513	0.011	2.2E-05	0.37	46,229	0.015	**5.2E-09**	*PDHB, PXK, KCTD6*	New	***PDHB*** **(B=↑**; S=ns),** *PXK* (B=ns; S=ns), *KCTD6* (B=ns; S=ns)
rs6855246	a	g	0.93	35,513	-0.003	0.672	0.92	46,068	-0.034	**8.68E-10**	*SLC39A8*†	New	***SLC39A8 (B=↑**)***
rs548630	a	c	0.49	36,286	-0.016	**2.1E-10**	0.47	46,137	-0.018	**1.2E-12**	*FCHO2, TMEM171, TNPO1*	New	*FCHO* (B=ns; S=ns)**,** ***TMEM171*** **(B=↑***; S=↑*),** ***TNPO1*** **(B=↑***; S=↑**)**
rs9390460	t	c	0.47	36,286	-0.019	**2.2E-15**	0.46	46,212	-0.033	**5.6E-42**	*STXBP5*	Known	-
rs9271597 ‡	a	t	0.41	28,203	-0.015	**1.4E-08**	0.41	31,364	-0.012	2.1E-04	*HLA region*	New	***HLA-C (B=↑**;s=↑***)***
rs7788962	a	g	0.62	33,773	-0.010	2.3E-04	0.61	46,231	-0.014	**7.3E-09**	*GIMAP7, GIMAP4*	New	***GIMAP7*** **(B=↑***; S=ns),** *GIMAP4* (B=ns; S=ns)
rs4276643	t	c	0.66	36,286	-0.023	**1.3E-19**	0.67	44,168	-0.029	**8.8E-28**	*SCARA5*	Known	-
rs10102164	a	g	0.19	36,286	0.019	**2.4E-09**	0.20	46,230	0.009	2.9E-03	*SOX17, RP1*	New	Too little FVIII released by endothelial cells to be detected *in vitro*
rs6479259	t	c	0.73	28,535	-0.021	1.1E-03	0.73	24,987	-0.056	**1.5E-08**	*OR13C5, NIPSNAP*§	New	***NIPSNAP3A*** **(B=↑*; S=↑**),** ***NIPSNAP3B*** **(B=↑*; S=ns)**
rs10985344	a	g	0.25	36,286	0.011	7.5E-05	0.25	46,178	0.017	**3.5E-09**	*DAB2IP*	New	*DAB2IP (B=ns; S=ns)*
rs687289	a	g	0.36	36,286	0.145	**1.9E-770**	0.33	46,231	0.197	**5.0E-1443**	*ABO*	Known	-
9:13930481:ID	d	i	0.85	22,480	-0.032	**2.7E-10**	0.85	29,409	-0.003	4.4E-01	*LINC00583, NFIB*	New	Too little FVIII released by endothelial cells to be detected *in vitro*
rs35458154	a	g	0.03	33,871	0.048	3.1E-08	0.03	44,020	0.060	**3.0E-12**	*ST3GAL4*	Known ‖	***ST3GAL4*** **(B=↓**; S=ns)**
rs4981022	a	g	0.69	36,286	0.025	**3.0E-20**	0.69	46,232	0.035	**6.6E-41**	*STAB2*	Known	-
rs4759787	a	c	0.40	36,286	0.011	1.1E-05	0.37	46,180	0.023	**7.7E-20**	*STX2*	Known	-
rs2238109	a	t	0.39	36,286	0.026	**3.5E-24**	0.38	46,232	0.050	**1.8E-89**	*VWF*	Known	-
rs4904820	a	g	0.49	36,286	0.014	**1.8E-08**	0.47	46,232	0.022	**6.0E-19**	*TCN2*	Known	-
rs6494314	t	c	0.18	36,286	-0.007	2.4E-02	0.17	46,232	-0.018	**1.1E-08**	*C2CD4B*	New	***C2CD4B*** **(B=↑*; S=↑**)**
rs1869365	Significant in combined multiphenotype analysis									**<5.0E-8**	*RAB5C, KAT2A*	New	***RAB5C (B=ns; S=↓*), KAT2A (B=ns; S=↑**)***
rs2277998	a	g	0.30	33,097	-0.010	4.4E-04	0.30	45,566	-0.022	**6.5E-16**	*CLEC4M*	Known	-
rs5750823 #	t	c	0.70	36,286	-0.013	1.5E-06	0.72	46,230	-0.020	**6.0E-14**	*RPL3, TAB1, SYNGR1, PDGB*	New	***TAB1*** **(B=ns; S=↑**),** ***SYNGR1*** **(B=↑*; S=ns),** *PDGFB* (B=ns; S=ns)
rs9616897	Significant in combined multiphenotype analysis									**<5.0E-8**	*ARSA*	New	*ARSA (B=ns; S=ns)*
rs150926226 ¶	c	g	0.62	20,537	0.017	**3.3E-09**	0.65	28,685	-0.005	7.0E-02	*TMLHE, F8*	Known	Too little FVIII released by endothelial cells to be detected *in vitro*

Table 2					EUR-only	TRANS-ethnic	Top Meta variant	LD† (r2)			
MarkerName	rsID	Closest Gene(s)	variant position	putative functional	Original P-value	Conditional P-value	Original P-value	Conditional P-value	LD* (D')		
FVIII (n=29,573 EUR; n=36,286 TRANS)											
5:72406659	**rs548630**	**TNPO1; FCHO2; TMEM171**	**9.5 kb 5' TMEM171**		**5.14E-10**	**3.96E-10**	**2.10E-10**	**2.06E-10**	**5:72406659**	**Top variant**	**Top variant**
6:147701217	rs9390461	STXBP5	intronic	rs1039084, missense	2.74E-13	4.06E-13	1.73E-14	-	6:147703299	0.98	1.00
6:147703299	**rs9399599**	**STXBP5**	**intronic**	**rs1039084, missense**	**3.73E-13**	-	**1.41E-15**	**4.70E-16**	**6:147703299**	**Top variant**	**Top variant**
8:27778148	rs55829013	SCARA5	intronic		3.39E-10	8.54E-10	5.81E-10	-	8:27805815	0.51	0.77
8:27823832	rs11780263	SCARA5	intronic		3.41E-17	6.96E-17	1.39E-19	-	8:27805815	0.10	0.98
8:27805815	**rs7816579**	**SCARA5**	**intronic**	**SiPhy conserved**	**5.32E-16**	-	**6.52E-21**	**4.07E-21**	**8:27805815**	**Top variant**	**Top variant**
8:55421614	**rs10102164**	**SOX17; RP1**	**intergenic**		**1.66E-07**	-	**2.38E-09**	**2.44E-09**	**8:55421614**	**Top variant**	**Top variant**
9:13930481:ID	**rs35468074**	**LINC00583**	**intronic**		**1.74E-10**	**1.24E-10**	**2.66E-10**	**3.10E-10**	**9:13930481:ID**	**Top variant**	**Top variant**
9:136116662	rs11793768	ABO	14 Kb 3' ABO		3.74E-13	1.29E-62	2.16E-13	-	9:136132908:ID	0.08	0.72
9:136137065	rs687621	ABO	intronic		1.18E-647	p<1E-320	6.85E-778	-	9:136132908:ID	0.89	0.98
9:135976698	rs35108384	ABO	intronic		4.58E-12	-	2.24E-12	6.53E-12	9:136132908:ID	0.00	0.14
9:136114000	rs78490142	ABO	17 kb 3' ABO		1.12E-03	-	6.53E-05	1.48E-13	9:136132908:ID	0.01	1.00
9:136132908:ID	**rs8176719**	**ABO**	**intronic**	**SiPhy conserved**	**1.62E-307**	-	**3.52E-403**	**p<1E-320**	**9:136132908:ID**	**Top variant**	**Top variant**
9:136178821	rs574237	ABO	intergenic		2.89E-29	-	6.94E-38	5.38E-26	9:136132908:ID	0.01	0.18
9:136181539	rs551924	ABO	intergenic		2.81E-48	-	9.21E-75	2.20E-63	9:136132908:ID	0.00	0.14
9:136207218	rs607900	ABO	intergenic	rs116779216, missense SURF1	6.86E-03	-	3.21E-15	1.85E-13	9:136132908:ID	0.00	0.19
9:136255149	rs62575992	ABO	intergenic		3.46E-42	-	8.45E-45	1.83E-26	9:136132908:ID	0.02	0.80
9:136344853	rs3124758	ABO	intergenic		2.71E-21	-	1.74E-18	5.88E-15	9:136132908:ID	0.00	0.14
12:6062894	rs57040304	VWF	intronic	rs7962217, missense, SiPhy cons	1.51E-15	1.22E-14	7.71E-12	-	12:6160614	0.00	0.09
12:6160614	**rs7135039**	**VWF**	**intronic**	**rs1063856, missense**	**3.00E-19**	**7.02E-18**	**2.32E-24**	**7.21E-20**	**12:6160614**	**Top variant**	**Top variant**
12:6070845	rs12423482	VWF	intronic	rs7962217, missense, SiPhy cons	2.49E-15	-	2.92E-18	1.57E-17	12:6160614	0.00	0.15
12:6160146	rs11064010	VWF	intronic		8.36E-07	-	2.41E-11	2.51E-09	12:6160614	0.01	0.33
12:104000319	rs1070073	STAB2	intronic		1.07E-13	5.04E-13	2.25E-14	-	12:104149874	0.00	0.16
12:104147207	rs3751198	STAB2	intronic		2.16E-17	5.91E-17	5.88E-18	-	12:104149874	0.70	0.98
12:104000470	rs2723889	STAB2	intronic		1.08E-13	-	2.14E-14	1.21E-14	12:104149874	0.00	0.16
12:104149874	**rs4981022**	**STAB2**	**intronic**		**1.16E-17**	-	**2.95E-20**	**4.47E-21**	**12:104149874**	**Top variant**	**Top variant**
14:92268531	rs10498631	TC2N	intronic		6.69E-09	8.63E-09	2.30E-08	-	14:92302972	0.54	0.82
14:92302972	**rs58204830**	**TC2N**	**intronic**		**2.57E-08**	-	**6.28E-09**	**7.03E-09**	**14:92302972**	**Top variant**	**Top variant**
22:39717706	**rs137631**	**RPL3**	**1.3Kb 5´of RPL3**		**2.35E-07**	-	**9.48E-09**	**1.34E-08**	**22:39717706**	**Top variant**	**Top variant**

X:154721357	**rs150926226**	**TMLHE,F8**				‡	**3.25E-09**	‡	**X:154721357**	-	-
VWF (n=42,256 EUR; n=46,232 TRANS)											
3:58383174	rs55692656	PXK; KCTD6; PDHB	intronic PXK	rs56384862, missense PXK; rs34579268 missense PXK; rs200687616 frameshift PDHB	6.48E-09	1.01E-08	3.33E-08	-	3:58436476	0.94	1.00
3:58436476	**rs55954186**	**PXK; KCTD6; PDHB**	**17 kb 3´of PXK**	**rs56384862, missense PXK; rs34579268 missense PXK; rs200687616 frameshift PDHB**	**8.72E-09**	-	**5.20E-09**	**8.06E-09**	**3:58436476**	**Top variant**	**Top variant**
5:72403453	rs7733340	TNPO1; FCHO2; TMEM171	intergenic		5.80E-13	1.52E-12	3.40E-12	-	5:72406659	0.93	0.99
5:72406659	**rs548630**	**TNPO1; FCHO2; TMEM171**	**9.5 kb 5' TMEM171**		**1.33E-12**	-	**1.22E-12**	**6.19E-13**	**5:72406659**	**Top variant**	**Top variant**
6:31192766	rs9263993	HLA region	intergenic		8.54E-09	3.50E-09	1.69E-07	-	6:31925848	0.00	0.63
6:31941629	rs116420479	HLA region	intronic STK19		3.10E-08	1.49E-08	1.02E-08	-	6:31925848	0.94	0.98
6:31158633	rs9263861	HLA region	intergenic		3.08E-08	-	3.25E-08	8.64E-09	6:31925848	0.00	1.00
6:31906828	rs78593564	HLA region	intronic C2		3.15E-08	-	1.03E-08	2.76E-09	6:31925848	1.00	1.00
6:147694334	**rs9390460**	**STXBP5**	**intronic**	**rs1039084, missense**	**6.32E-38**	**1.67E-38**	**5.58E-42**	**1.27E-42**	**6:147694334**	**Top variant**	**Top variant**
7:150296496	rs13230842	GIMAP7; GIMAP4	intergenic		2.06E-08	1.96E-08	2.95E-08	-	7:150227227	0.18	0.97
7:150227227	**rs7788962**	**GIMAP7; GIMAP4**	**9.1Kb 3´of GIMAP7**		**6.29E-08**	-	**7.30E-09**	**1.36E-08**	**7:150227227**	**Top variant**	**Top variant**
8:27803599	**rs4276643**	**SCARA5**	**intronic**		**5.77E-27**	**1.19E-29**	**8.77E-28**	**7.69E-31**	**8:27803599**	**Top variant**	**Top variant**
8:27815481	rs62496810	SCARA5	intronic		3.74E-08	4.38E-11	4.98E-08	6.04E-11	8:27803599	0.02	1
9:124416940	**rs10985344**	**DAB2IP**	**intronic**		**4.11E-09**	**3.76E-09**	**3.47E-09**	-	**9:124416940**	**Top variant**	**Top variant**
9:124421965	rs4837886	DAB2IP	intronic		3.84E-09	-	4.43E-09	3.94E-09	9:124416940	0.79	0.97
9:135919501	rs138796740	ABO	intergenic		2.61E-08	7.06E-11	3.59E-08	-	9:136132908:ID	0.00	0.15
9:136128546	rs7857390	ABO	2.5Kb 3' of ABO		p<1E-320	7.81E-294	2.15E-393	-	9:136132908:ID	0.36	1
9:136145414:ID	rs202001822	ABO	intronic		5.90E-78	1.46E-39	2.26E-80	-	9:136132908:ID	0.10	1
9:136147553	rs660340	ABO	intronic		p<1E-320	p<1E-320	1.11E-492	-	9:136132908:ID	0.33	0.87
9:136177394	rs656105	ABO	intergenic		1.20E-66	2.62E-129	4.23E-73	-	9:136132908:ID	0.02	0.19
9:136128000	rs10901252	ABO	3.1 Kb 3' of ABO	rs8176747, missense; rs8176746, missense; rs8176743 missense	1.78E-291	-	2.61E-351	p<1E-320	9:136132908:ID	0.16	0.99
9:136130677	rs62641786	ABO	3'UTR		2.79E-96	-	2.59E-106	8.85E-30	9:136132908:ID	0.03	0.91
9:136138765:ID	rs8176685	ABO	intronic		p<1E-320	-	5.95E-507	p<1E-320	9:136132908:ID	0.31	1
9:136316367	rs28680325	ABO	intronic ADAMTS13		1.04E-09	-	2.88E-10	1.12E-09	9:136132908:ID	0.01	0.25
11:126296825	**rs35458154**	**ST3GAL4**	**intronic**		**6.35E-12**	**8.31E-12**	**2.96E-12**	**2.42E-12**	**11:126296825**	**Top variant**	**Top variant**
12:6157394	rs2283335	VWF	intronic	rs1063856, missense	6.48E-83	5.55E-79	2.08E-89	-	12:6153967	1	1
12:6160146	rs11064010	VWF	intronic		1.60E-27	8.75E-22	7.90E-27	2.40E-21	12:6153967	0.01	0.30

12:6225931	rs112814955	VWF	intronic		1.43E-42	3.14E-49	9.02E-43	4.00E-50	12:6153967	0.01	0.34
12:6153967	**rs2238109**	**VWF**	**intronic**	**rs1063856, missense**	**1.37E-82**	**-**	**1.77E-89**	**3.41E-87**	**12:6153967**	**Top variant**	**Top variant**
12:104000319	rs1070073	STAB2	intronic		2.28E-36	8.23E-31	2.03E-36	6.72E-32	12:104149874	0.00	0.16
12:104007418	rs11111679	STAB2	intronic		2.31E-14	4.79E-10	1.22E-14	1.60E-10	12:104149874	0.00	0.02
12:104127353	rs73192004	STAB2	intronic	rs17034433, missense	2.29E-08	2.33E-08	5.09E-09	6.57E-09	12:104149874	0.01	0.53
12:104149874	**rs4981022**	**STAB2**	**intronic**		**5.44E-37**	**4.69E-41**	**6.57E-41**	**8.65E-45**	**12:104149874**	**Top variant**	**Top variant**
12:131287011	rs6486599	STX2	intronic	rs17564, missense	4.23E-18	1.92E-18	1.14E-18	-	12:131290180	1.00	1
12:131290180	**rs4759787**	**STX2**	**intronic**	**rs17564, missense**	**4.45E-18**	**-**	**7.73E-20**	**3.84E-20**	**12:131290180**	**Top variant**	**Top variant**
14:92290744	rs10498632	TCN2	intronic		2.58E-18	4.74E-18	5.00E-18	-	14:92318935	0.53	0.99
14:92318935	**rs4904820**	**TCN2**	**intronic**		**8.30E-18**	**-**	**6.04E-19**	**2.40E-19**	**14:92318935**	**Top variant**	**Top variant**
15:62455019	**rs6494314**	**C2CD4B**	**700bp 3´of gene**	**rs8040712, missense**	**9.63E-08**	**-**	**1.14E-08**	**1.30E-08**	**15:62455019**	**Top variant**	**Top variant**
19:7831628	**rs2277998**	**CLEC4M**	**missense**	**rs2277998, missense**	**2.02E-15**	**7.30E-16**	**6.47E-16**	**1.74E-15**	**19:7831628**	**Top variant**	**Top variant**
22:39790191	rs2413590	PFGFB; SYNGR1; TAB1	5.6Kb of TAB1		3.96E-12	4.44E-12	2.02E-12	-	22:39829973	0.75	0.94
22:39829973	**rs5750823**	**PFGFB; SYNGR1; TAB1**	**intronic TAB1**		**2.98E-11**	**-**	**5.95E-14**	**7.58E-14**	**22:39829973**	**Top variant**	**Top variant**

Figure Legends

Figure 1A: Manhattan plot for the trans-ethnic analyses FVIII (named by closest gene)

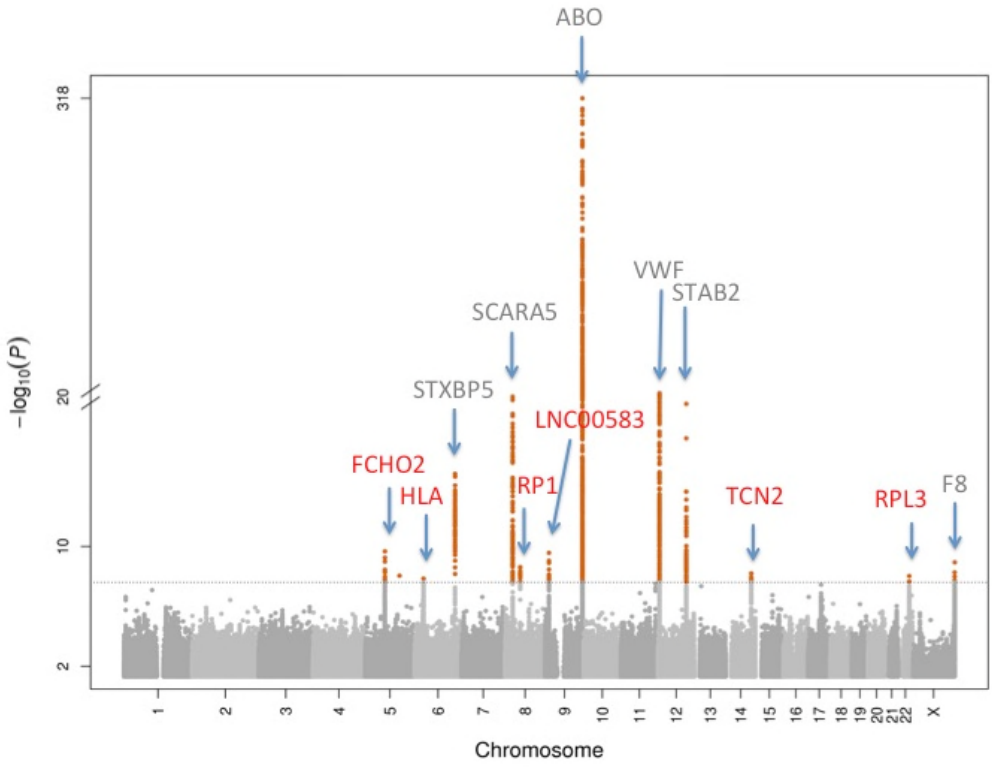


Figure 1B: Manhattan plot for the trans-ethnic analyses VWF (named by closest gene).

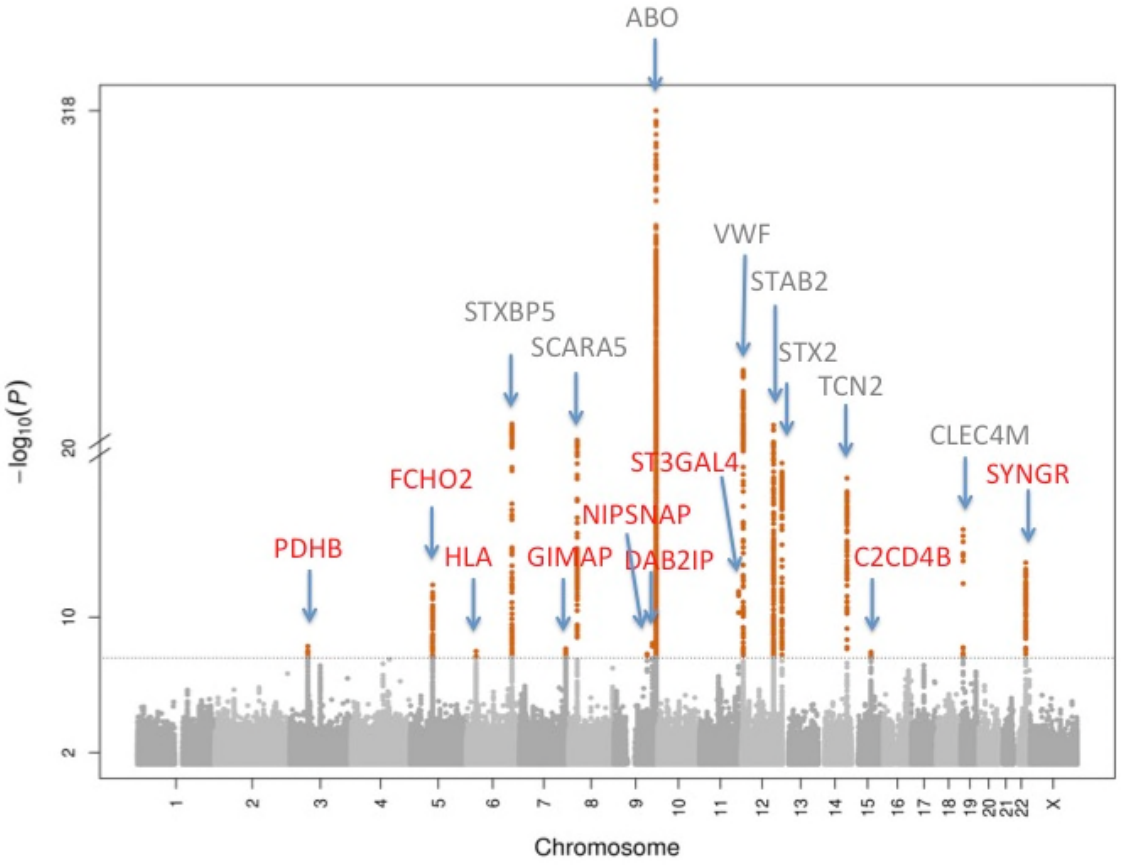


Figure 2A. Silencing candidate genes changes basal release of VWF.

HUVEC cells were transfected with siRNA against selected genes for 4 days, the media was changed, cells were cultured for 30 min, and VWF was measured in the supernatant via ELISA. n = 4 ± S.D. *p<0.05, **p<0.01, ***p<0.001. All results are relative to VWF release after transfection with a scrambled control siRNA, which is set as reference (100%).

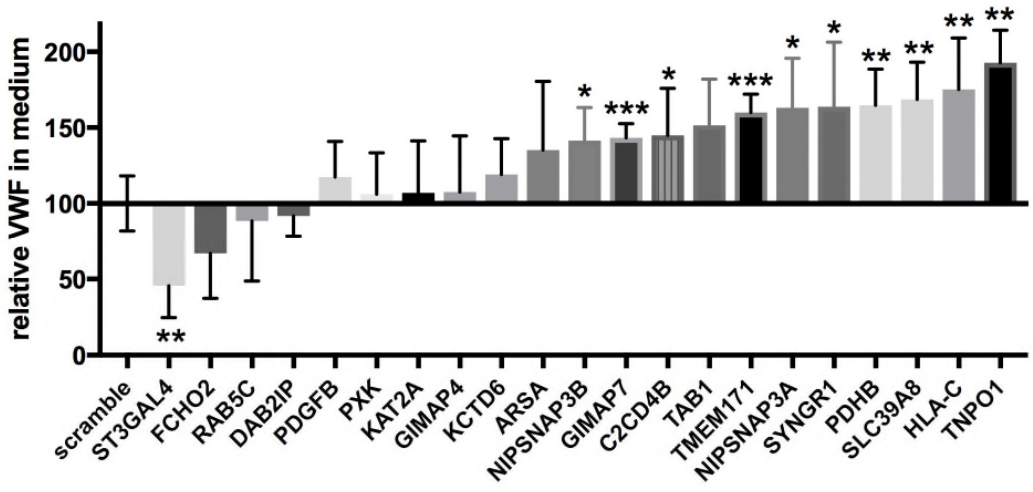


Figure 2B. Silencing candidate genes changes stimulated release of VWF.

HUVEC cells were transfected with siRNA against selected genes for 4 days, the media was changed, cells were stimulated with histamine 10 µM for 30 min, and VWF was measured in the supernatant by an ELISA. n = 4 ± S.D. *p<0.05, **p<0.01, ***p<0.001. All results are relative to VWF release after transfection with a scrambled control siRNA, which is set as reference (100%).

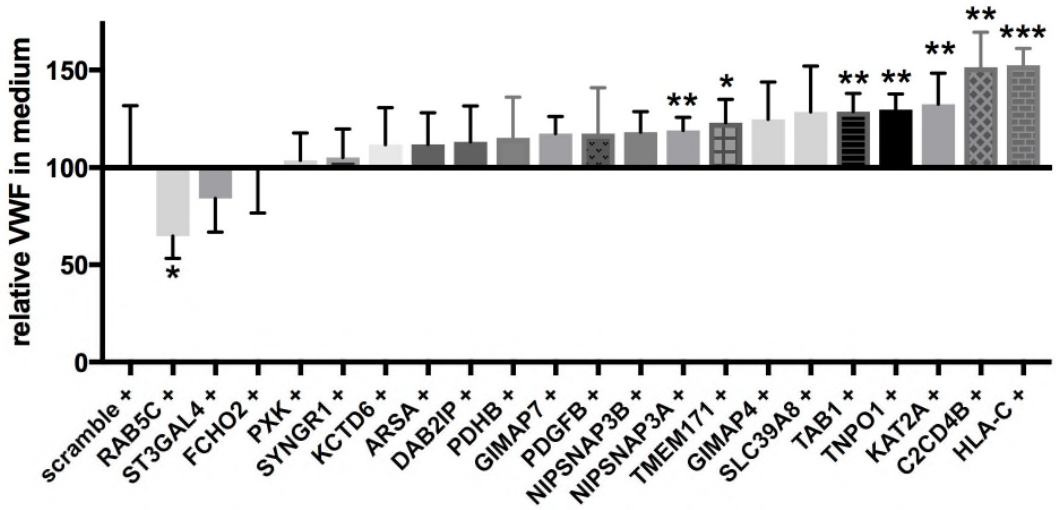
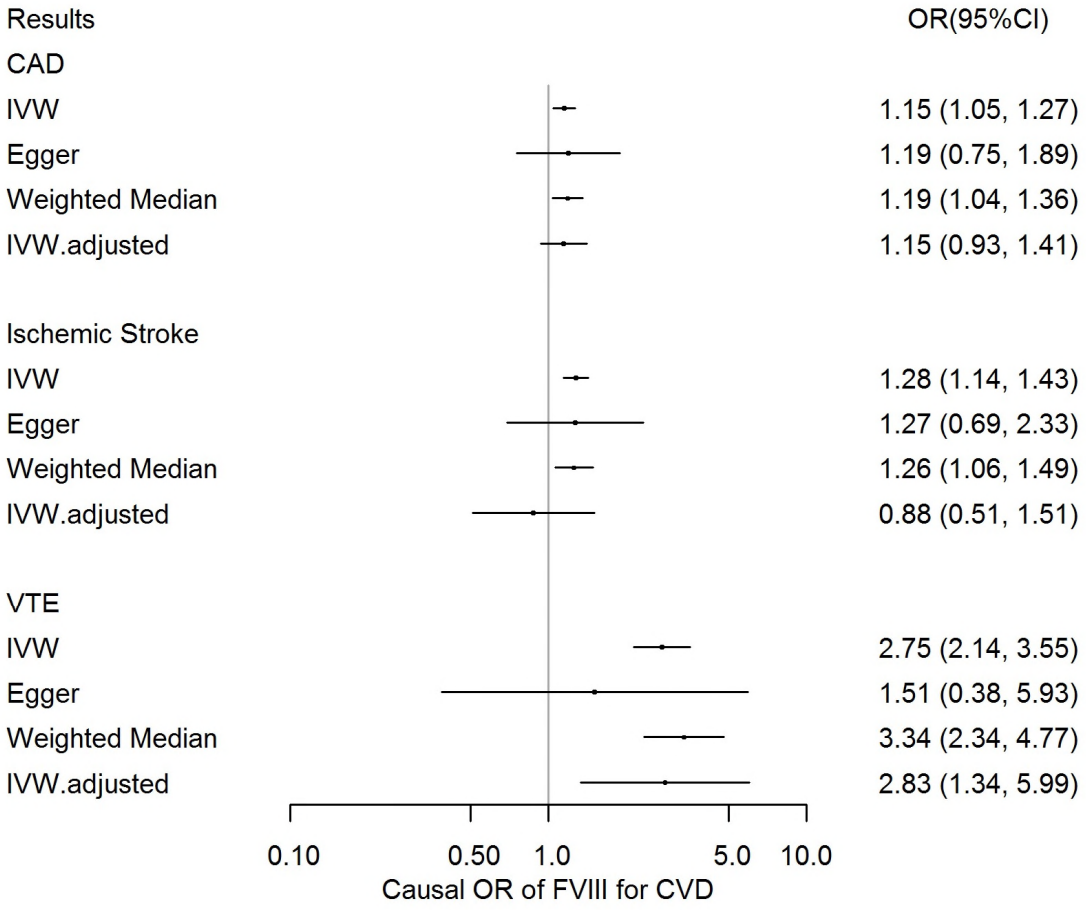

relative VWF in medium
150
100
50
0
*
**
*
**
**
**
***
scramble +
RAB5C +
ST3GAL4 +
FCHO2 +
PXK +
SYNGR1 +
KCTD6 +
ARSA +
DAB2IP +
PDHB +
GIMAP7 +
PDGFB +
NIPSNAP3B +
NIPSNAP3A +
TMEM171 +
GIMAP4 +
SLC39A8 +
TAB1 +
TNPO1 +
KAT2A +
C2CD4B +
HLA-C +

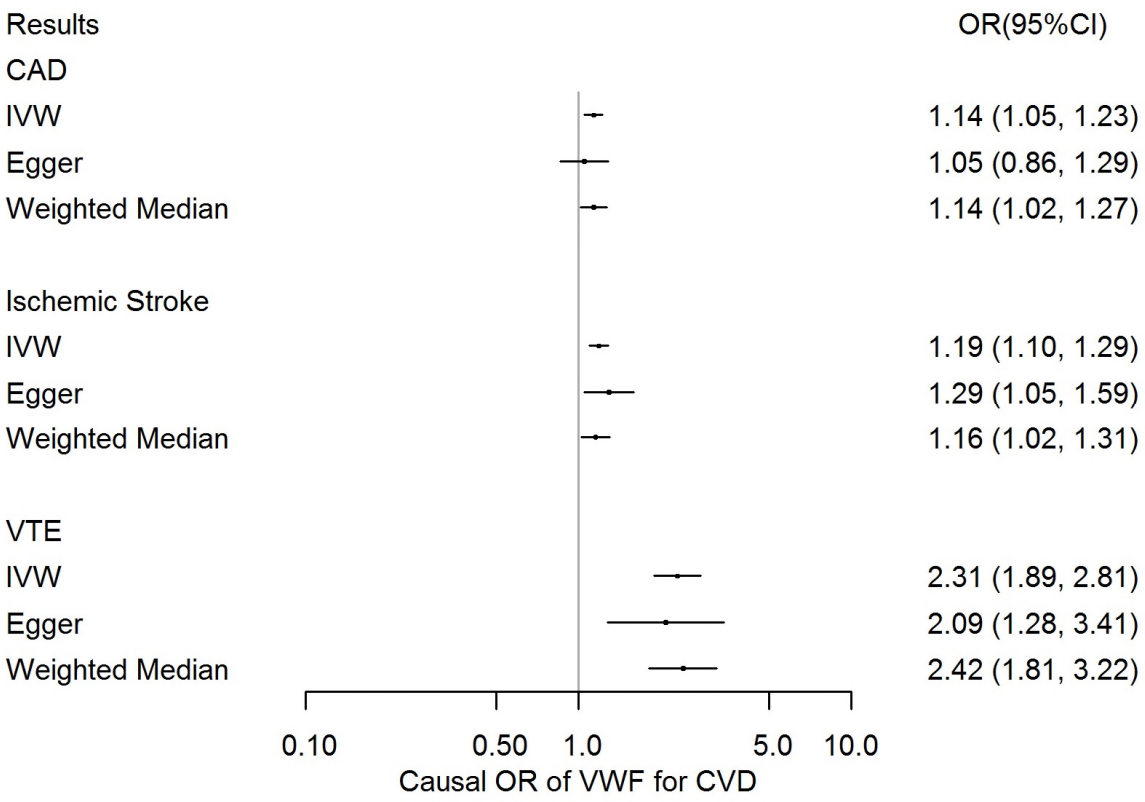
Figure 3

Mendelian Randomization results. Results show OR (95% confidential interval) per every higher standard deviation change in FVIII and VWF. CAD (Coronary Artery Disease), IVW(inverse-variance weighted method), IVW.adjusted (IVW FVIII adjusted for the effects of VWF).

a) FVIII



b) VWF


Results
OR(95%CI)
CAD
IVW
1.14 (1.05, 1.23)
Egger
1.05 (0.86, 1.29)
Weighted Median
1.14 (1.02, 1.27)
Ischemic Stroke
IVW
1.19 (1.10, 1.29)
Egger
1.29 (1.05, 1.59)
Weighted Median
1.16 (1.02, 1.31)
VTE
IVW
2.31 (1.89, 2.81)
Egger
2.09 (1.28, 3.41)
Weighted Median
2.42 (1.81, 3.22)
0.10
0.50
1.0
5.0
10.0
Causal OR of VWF for CVD

Supplementary material

Supplementary Table S1: Descriptions of participating cohorts and resources

Supplementary Table S2: Final SNPs composing instrumental variables

Supplementary Table S3: MR results for the instrumental variables individually

Supplementary Table S4: global results for the Egger Regression to evaluate heterogeneity, before and after exclusions

Supplementary Table S5: Results from conditional analyses from European-only data

Supplementary Table S6: Results from conditional analyses from transethnic data

Supplementary Table S7: Haploreg/SNIPA/GWAS catalogue for all transethnic top-SNPs

Supplementary Table S8: eQTL in other tissues

Supplementary Table S9: eQTL in whole blood

Supplementary Table S10: Factor VIII enrichment networks

Supplementary Table S11: Factor VIII prioritized genes in networks

Supplementary Table S12: von Willebrand factor enrichment networks

Supplementary Table S13: von Willebrand factor genes in networks

Supplementary Figures S1a-n: Regional plots for the novel associations

Supplementary Figures S2a-c: Scatter plots for the genetic associations between FVIII and cardiovascular events

Supplementary Figure S3: Figure shows the hypothesized effect of the genes found in the present study in relation to the possible regulatory points in VWF synthesis and secretion from endothelial cells. The specific regulatory point suggestion is based on previous literature evidence. Of note, VWF clearance or regulation in platelets were not studied in our in vitro first-pass analyses.