SUPPLEMENTARY METHODS

LD-score Regression and Multi-phenotype Analyses

We assessed the genetic correlation between FVIII and VWF (proportion of shared heritability) using LD-score regression [1] using summary statistics for the European-only meta-analysis of FVIII and VWF (including 25,897 and 42,257 individuals respectively). Given the strong genetic correlation, and in order to gain statistical power, genome-wide autosomal data from the multi-ethnic meta-analysis for FVIII and VWF were combined using aSPU package[2]. Briefly, z-scores for each phenotype were calculated and then used to estimate the covariance between both phenotypes. To overcome the computational burden of running high number of permutations (B) in the whole GWAS dataset, we first run the whole dataset at B=1000, then selected those SNPs where the p-value was smaller than 0.0025 (2.5/number of permutations), and re-run the smaller set of SNPs at increasingly higher B values until the limit of $B=10^9$ permutations. Since the p-values obtained with this method is limited by the number of permutations, the final run of $B=10^9$ permutations limit the minimum p-value to $2x10^{-9}$.

Proportion of variance explained

Proportion of variance r^2 explained by the associated variants was calculated using the formula:

$$r^2 = \frac{2f(1-f)\beta^2}{var(y)}$$

Where f is the effect allele frequency and β is the per-allele increment in natural log-transformed phenotype. We calculated beta value using summary-statistics beta values from all participating cohorts and considering the variance, $var(y)$, as square of the weighted mean of standard deviations for the log-transformed phenotype across all cohorts. In addition, we

confirmed these values by calculating the proportion of variance explained by the top SNPs in every individual cohort and then summing them up, which resulted in minor differences.

Putative Variants

Putative functional SNPs for all new loci were investigated *in silico* using publicly available data from Haploreg [3] and SNiPA [4].

Expression quantitative trait loci (eQTL) analysis

The SNP with lowest p-value for each locus was selected and eQTL analysis was performed by checking association with nearby genes (at +/- 1Mb from the selected SNP) in the multi-tissue expression database curated by the Johnson lab at Framingham. A list of all studies included in the database along with their PubMed id is included in below, and results are shown in Supplementary Table 3a and 3b. Results from whole blood and liver studies were prioritized since they have the largest sample size and come from relevant tissues. Since most studies in the database were not performed using 1000G data, proxy SNPs in high LD (r2>0.8) and present in HapMap were selected using SNiPA 4 and also queried.

Database material: A general overview of a subset of >50 eQTL studies has been published (24973796), with specific citations for >100 datasets included in the current query following here.

Blood cell related eQTL studies included fresh lymphocytes (17873875), fresh leukocytes (19966804), leukocyte samples in individuals with Celiac disease (19128478), whole blood samples (18344981, 21829388, 22692066, 23818875, 23359819, 23880221, 24013639, 23157493, 23715323, 24092820, 24314549, 24956270, 24592274, 24728292, 24740359,

25609184, 22563384, 25474530, 25816334, 25578447), lymphoblastoid cell lines (LCL) derived from asthmatic children (17873877, 23345460), HapMap LCL from 3 populations (17873874), a separate study on HapMap CEU LCL (18193047), additional LCL population samples (19644074, 22286170, 22941192, 23755361, 23995691, 25010687, 25951796), neutrophils (26151758, 26259071), CD19+ B cells (22446964), primary PHA-stimulated T cells (19644074, 23755361), CD4+ T cells (20833654), peripheral blood monocytes (19222302,20502693,22446964, 23300628, 25951796, 26019233), long non-coding RNAs in monocytes (25025429) and CD14+ monocytes before and after stimulation with LPS or interferon-gamma (24604202), CD11+ dendritic cells before and after Mycobacterium tuberculosis infection (22233810) and a separate study of dendritic cells before or after stimulation with LPS, influenza or interferon-beta (24604203). Micro-RNA QTLs (21691150, 26020509), DNase-I QTLs (22307276), histone acetylation QTLs (25799442), and ribosomal occupancy QTLs (25657249) were also queried for LCL. Splicing QTLs (25685889) and micro-RNA QTLs (25791433) were queried in whole blood.

Non-blood cell tissue eQTLs searched included omental and subcutaneous adipose (18344981, 21602305, 22941192, 23715323, 25578447), visceral fat (25578447) stomach (21602305), endometrial carcinomas (21226949), ER+ and ER- breast cancer tumor cells (23374354), liver (18462017,21602305,21637794, 22006096, 24665059, 25578447), osteoblasts (19654370), intestine (23474282) and normal and cancerous colon (25079323, 25766683), skeletal muscle (24306210, 25578447), breast tissue (normal and cancer)(24388359, 22522925), lung (23209423, 23715323, 24307700, 23936167, 26102239), skin (21129726, 22941192, 23715323, 25951796), primary fibroblasts (19644074, 23755361, 24555846), sputum (21949713), pancreatic islet cells (25201977), prostate (25983244), rectal mucosa (25569741), arterial wall (25578447) and heart tissue from left ventricles (23715323,

24846176) and left and right atria (24177373). Micro-RNA QTLs were also queried for gluteal and abdominal adipose (22102887) and liver (23758991). Methylation QTLs were queried in pancreatic islet cells (25375650). Further mRNA and micro-RNA QTLs were queried from ER+ invasive breast cancer samples, colon-, kidney renal clear-, lung- and prostate-adenocarcinoma samples (24907074).

Brain eQTL studies included brain cortex (19222302, 19361613, 22685416, 25609184, 25290266), cerebellar cortex (25174004), cerebellum (20485568, 22685416, 22212596, 22832957, 23622250), frontal cortex (20485568, 22832957, 25174004), gliomas (24607568), hippocampus (22832957, 25174004), inferior olivary nucleus (from medulla) (25174004), intralobular white matter (25174004), occiptal cortex (25174004), parietal lobe (22212596), pons (20485568), pre-frontal cortex (22031444, 20351726, 22832957, 23622250), putamen (at the level of anterior commussure) (25174004), substantia nigra (25174004), temporal cortex (20485568, 22685416, 22832957, 25174004), thalamus (22832957) and visual cortex (23622250).

Additional eQTL data was integrated from online sources including ScanDB, the Broad Institute GTEx Portal, and the Pritchard Lab (eqtl.uchicago.edu). Cerebellum, parietal lobe and liver eQTL data was downloaded from ScanDB and cis-eQTLs were limited to those with P<1.0E-6 and trans-eQTLs with P<5.0E-8. Results for GTEx Analysis V4 for 13 tissues were downloaded from the GTEx Portal and then additionally filtered as described below [www.gtexportal.org: thyroid, leg skin (sun exposed), tibial nerve, aortic artery, tibial artery, skeletal muscle, esophagus mucosa, esophagus muscularis, lung, heart (left ventricle), stomach, whole blood, and subcutaneous adipose (23715323)]. Splicing QTL (sQTL) results generated with sQTLseeker with false discovery rate P≤0.05 were retained. For all gene-level

eQTLs, if at least 1 SNP passed the tissue-specific empirical threshold in GTEx, the best SNP for that eQTL was always retained. All gene-level eQTL SNPs with P<1.67E-11 were also retained, reflecting a global threshold correction of P=0.05/(30,000 genes X 1,000,000 tests).

Pathway analyses

We used DEPICT (Data-driven Expression-Prioritized Integration for Complex Traits) [5] to infer genes and pathways that were enriched by variants associated with plasma FVIII and VWF. All independent SNPs (defined with r^2< 0.95 in every 500 Kb region, using 1000G-v3_GIANT reference data downloaded from http://csg.sph.umich.edu/abecasis/mach/download/1000G.2012-03-14.html) with p-value < 10^{-5} from the European-specific analysis were used as input for the pathway analysis, as recommended in the pipe-line.

Mendelian Randomization

The final variants composing instrumental variables (IV) for FVIII and VWF are listed in Supplementary Table 2a. We conducted MR-Egger regressions [6] and weighted median estimates (WME) [7] to diagnose whether the instrumental variables for the main analysis estimates were valid. To avoid bias due to pleiotropic effect in MR, we visually examined the causal effect estimates produced by each of the individual variants using scatter plots, funnel plots, and forest plots (Supplementary Figure 2), and calculated heterogeneity statistics.[8] This allowed us to select IV as follows: first, we removed the *ABO* and *HLA* loci, which showed known effects with multiple CVD related traits; second, we further removed the *DABI2P* locus for IV of VWF, which was suggested to be an outliner in the heterogeneity test for IVW estimate of VWF effect on CAD; third, we also tested the directional pleiotropic

effect using the Egger regression, but no locus was suggested further removed in this step (Supplementary Table 2c).

For the multivariate MR approach to test the causal association of FVIII independent of VWF, we used all loci associated with at least one of FVIII and VWF after exclusion strategies. If different sentinel SNPs were reported for FVIII and VWF in the same locus (but in LD with each other), then a random SNP was selected from these two to represent the locus.

Study descriptions

The Age, Gene/Environment Susceptibility-Reykjavik Study (**AGES**-Reykjavik) was initiated in 2002. AGES-Reykjavik was designed to examine risk factors, including genetic susceptibility and gene/environment interaction, in relation to disease and disability in old age [9]. The AGES-Reykjavik sample is drawn from an established population-based cohort, the Reykjavik Study. This cohort of men and women born between 1907 and 1935 has been followed in Iceland since 1967 by the Icelandic Heart Association. The concentration of von Willebrand factor was determined by means of a sensitive enzyme immunoassay [10].

The **Atherosclerosis Risk in Communities (ARIC)** study has been described in detail previously [11]. Men and women aged 45-64 years at baseline were recruited from four communities: Forsyth County, North Carolina; Jackson, Mississippi; Minneapolis, Minnesota; and Washington County, Maryland. A total of 15,792 individuals, predominantly White and African American, participated in the baseline examination in 1987-1989, with three additional triennial follow-up examinations and a fifth exam in 2011-2013, and a sixth exam in 2016-2017. FVIII activities was measured using a clotting assay (% activity) while

VWF antigen was measured using ELISA (% antigen) [12, 13] in plasma samples obtained at the baseline examination.

The **British 1958 birth cohort (B58C)** is a national population sample followed periodically from birth. At age 44-45 years, 9377 cohort members were examined by a research nurse in the home as described previously and non-fasting blood samples were collected with permission for DNA extraction and creation of immortalised cell cultures [14]. DNA samples from unrelated subjects of white ethnicity, with nationwide geographic coverage, were genotyped either by the Wellcome Trust Case Control Consortium (WTCCC), the Type 1 Diabetes Genetics Consortium, or the GABRIEL consortium [15-17]. Details of the blood collection, VWF measurement and covariate adjustment have been described elsewhere [18]. In brief, VWF antigen was measured by ELISA assays that used a double-antibody sandwich (DAKO, Copenhagen, Denmark). The standard curve was constructed using the 9th British standard for Blood Coagulation Factors from the National Institute for Biological Standards and Controls (NIBSC), South Mimms, Herefordshire UK, and the results were expressed as International units/decilitre (IU/dl). As a control, the pooled plasma of 20 healthy middle-aged persons was run on each ELISA plate. The intra-assay CV was 6%, the inter-assay CV was 8%, and reference range was 50 to 200 IU/dl. Measurements were adjusted for sex, laboratory batch, time of day, month of examination, and postal delay. Adjustment for age was not required as all subjects were aged 44-45 years. Use of anticoagulant therapy was a contraindication to blood sampling. Valid VWF measurements were available for 6093 (93.9%) of the 6491 subjects with genotype data.

The **Coronary Artery Risk Development in Young Adults (CARDIA)** study is a prospective multicenter study with 5115 Caucasian and African American participants ages 18-30 years at baseline, recruited from four centers. The recruitment was done from the total community in Birmingham, AL, from selected census tracts in Chicago, IL and Minneapolis, MN; and from the Kaiser Permanente health plan membership in Oakland, CA. The details of the study design for the CARDIA study have been published before [19]. Nine examinations have been completed since the baseline examination in 1985–1986, with follow-up examinations 2, 5, 7, 10, 15, 20, 25, and 30 years after baseline. Coagulation FVIII and VWF were measured in plasma citrate [20]. FVIII coagulant activities were assayed by a one-stage system with reagents from Pacific Hemostasis and George King Biomedical, Inc. The standard curve was prepared by using universal reference plasma from Curtin Matheson Scientific and the results calculated as a percentage of standard with the data management system of the MLAElectra 800. von Willebrand antigen was measured by an enzyme-linked immunosorbent assay obtained from American Bioproducts Co [21].

The **Cardiovascular Health Study (CHS)** is a population-based cohort study of risk factors for coronary heart disease and stroke in adults ≥65 years conducted across four field centers [22]. The original predominantly European -ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists; subsequently, an additional predominantly African-American cohort of 687 persons were enrolled for a total sample of 5,888. Blood samples were drawn from all participants at their baseline examination and DNA was subsequently extracted from available samples. Genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai among CHS participants who consented to genetic testing and had DNA available using the Illumina 370CNV BeadChip system (for European ancestry participants, in 2007) or the

Illumina HumanOmni1-Quad_v1 BeadChip system (for African-American participants, in 2010). CHS was approved by institutional review committees at each field center. Participants included in the present analyses had available DNA and gave informed consent including consent to use of genetic information for the study of cardiovascular disease. FVIII activity determined by using factor VIII-deficient plasma (Organon-Teknika) and partial thromboplastin (Organon-Teknika) [23]. Unassayed pooled normal plasma (George King Biomedical, Overland Park, KS, USA) was used as the standard and calibrated with the World Health Organization reference plasma for both assays. Activity measured using Coag-A-Mate X2 (Organon-Teknika). Values were expressed as percentage of the standard. The intra-assay CV was 12.6% and inter-assay CV was 13.8%. No reference range available.

The **CROATIA-Vis** study, Croatia, is a family-based, cross-sectional study in the isolated island of Vis that included 1,056 examinees aged 18-93. Blood samples were collected in 2003 and 2004 along with many clinical and biochemical measures and lifestyle and health questionnaires. The VWF antigen was measured by ELISA assays that used a double-antibody sandwich (DAKO, Copenhagen, Denmark).

The **Framingham Heart Study (FHS)** was started in 1948 with 5,209 randomly ascertained participants from Framingham, Massachusetts, US, who had undergone biannual examinations to investigate cardiovascular disease and its risk factors. In 1971, the Offspring cohort (comprising 5,124 children of the original cohort and the children's spouses) and in 2002, the Third Generation (consisting of 4,095 children of the Offspring cohort) were recruited. FHS participants in this study are of European ancestry. The methods of recruitment and data collection for the Offspring and Third Generation cohorts have been described elsewhere [24]. Von Willebrand factor was assessed using ELISA at

exam 5 (1991-1995) in the Offspring cohort. In our laboratory, the intra-assay coefficient of variation was 8.8%. No reference range is available [25].

The **Genes and Blood-Clotting Study (GABC)** consists of a cohort of 1,150 healthy siblings recruited from the University of Michigan, Ann Arbor, between June 26, 2006 and January 30, 2009. Participants were between the ages of 14 and 35 y, and had at least one eligible healthy sibling. Subjects who indicated that they were pregnant, had a known bleeding or blood-clotting disorder, or any illness requiring regular medical care were excluded. All participants provided informed consent. Subjects completed an online phenotyping survey and donated a blood sample for DNA extraction and plasma biochemical phenotyping. VWF levels were determined at the University of Michigan using a custom AlphaLISA™ assay (Perkin-Elmer), which utilized a polyclonal anti-von Willebrand factor antibody from DAKO Cytomation.

The **GAIT (Genetic Analysis of Idiopathic Thrombophilia**) project is a family based study where 935 subjects in 35 extended pedigrees were collected. To be included in the study, a family was required to have at least 10 living individuals in 3 or more generations. Families were selected through a proband with idiopathic thrombophilia, which was defined as recurrent thrombotic events (at least one of which was spontaneous), a single spontaneous thrombotic episode plus a first-degree relative also affected, or onset of thrombosis before age 45. Thrombosis in these probands was considered idiopathic when biological causes as antithrombin deficiency, protein S and C deficiencies, activated protein C resistance, plasminogen deficiency, heparin cofactor II deficiency, Factor V Leiden, dysfibrogenemia, lupus anticoagulant and antiphospholipid antibodies, were excluded. Subjects were interviewed by a physician to determine their health and reproductive history, current

medications, alcohol consumption, use of sex hormones (oral contraceptives or hormonal replacement therapy) and their smoking history. The study was performed according to the Declaration of Helsinki. All procedures of the study were reviewed by the Institutional Review Board of the Hospital de la Santa Creu i Sant Pau, Barcelona, Spain. Adult subjects gave informed consent for themselves and for their minor children. Coagulation factors were measured in platelet-poor citrated plasma [26]. Coagulation FVIII was assayed with deficient plasma from Diagnostica Stago (Asnieres). von Willebrand factor was measured by an ELISA method with antibodies from Dako. To reduce measurement error, assays were performed in duplicate, and the average value was calculated for each person. Intra-assay and interassay coefficients of variation were generally estimated to be between 2% and 6%.

GeneSTAR (Genetic Study of Atherosclerosis Risk) is an ongoing prospective family study begun in 1983 to explore the causes of early-onset cardiovascular disease. Probands with an early-onset coronary disease event prior to 60 years of age were identified at the time of hospitalization in any of 10 Baltimore area hospitals. Their apparently healthy 30-59 year old siblings without known CAD were initially recruited and screened between 1983 and 2006; offspring of the siblings and probands, as well as the co-parent of these offspring, were recruited and assessed between 2003 and 2006. In this study, European and African American participants with both genotyping with the Illumina Human1Mv1_C chip and measured VWF were included. Levels of VWF were measured using ELISA at the University of Maryland Cytokine Laboratory.

The Ludwigshafen Risk and Cardiovascular Health (LURIC) study is a monocentric hospital based prospective study including 3316 individuals referred for coronary angiography recruited in the Ludwigshafen Cardiac Center, southwestern Germany from

1997 – 2000. Clinical indications for angiography were chest pain or a positive non-invasive stress test suggestive of myocardial ischemia. To limit clinical heterogeneity, individuals suffering from acute illnesses other than acute coronary syndrome, chronic non-cardiac diseases and a history of malignancy within the five past years were excluded. All participants were of European ancestry and completed a detailed questionnaire which gathered information on medical history, clinical, and lifestyle factors. Study protocols were approved by the ethics committee of the "Landesärztekammer Rheinland-Pfalz" and the study was conducted in accordance with the "Declaration of Helsinki". Informed written consent was obtained from all participants. Coronary heart disease at baseline was defined as the presence of a visible luminal narrowing (>50% stenosis) in at least one of 15 coronary segments according to the classification of the American Heart Association. Fasting blood samples were obtained by venipuncture in the early morning and stored for later analyses. Coagulation FVIII and VWF were measured in plasma using methods previously described [27]. For this study a subset of 3061 Samples were used that had been genotyped on an Affymetrix 6.0 array.

The **MARseille THrombosis Association (MARTHA)** project has already been extensively [28]. It is composed of unrelated subjects of European origin, with the majority being of French ancestry, consecutively recruited at the Thrombophilia center of La Timone hospital (Marseille, France) between January 1994 and October 2012. All patients had a documented history of VT and free of well characterized genetic risk factors including AT, PC, or PS deficiency, homozygosity for FV Leiden or FII 20210A, and lupus anticoagulant. They were interviewed by a physician on their medical history, which emphasized manifestations of deep vein thrombosis and pulmonary embolism using a standardized questionnaire. The thrombotic events were confirmed by venography, Doppler ultrasound, spiral computed

tomographic scanning angiography, and/or ventilation/perfusion lung scan. Hemostasis-related parameters were centrally performed using the Star automate and commercially available kits and reagents from Diagnostica Stago (Asniéres, France) including the corresponding normal and pathological control plasmas and standard plasmas. FVIII and VWF levels are respectively measured by VIII:C using human FVIII-deficient plasma in a 1-stage factor assay and by STA LIATEST VWF (Diagnostica Stago) on Star automate. They were measured at least 3 months after the most recent VT event to minimize the effect of the acute phase [29]. From the 1542 MARTHA participants with GWAS data, 727 and 877 were included for the present analyses on FVIII and VWF levels.

The **Multiple Environmental and Genetic Assessment of risk factors for venous thrombosis (MEGA)** study is a large population-based case-control study [30]. Data collection and ascertainment of venous thrombotic events have been previously described in detail. In short, patients with a first deep vein thrombosis or pulmonary embolism were recruited at six anticoagulation clinics in the Netherlands between 1999 and 2004. The diagnosis of a deep vein thrombosis was based on compression ultrasonography, whereas a pulmonary embolism was confirmed by perfusion and ventilation scintigraphy, helical computed tomography or pulmonary angiography. Blood samples were taken at least 3 months after discontinuation of vitamin K antagonist treatment, unless patients were still receiving anticoagulant therapy one year after their VT event. For the present analyses, patients who were still receiving anticoagulant treatment at the time of blood collection were excluded. Factor VIII activity levels were measured with a mechanical clot detection method on an STA-R coagulation analyzer (Diagnostica Stago, Asnieres, France), whereas von Willebrand factor antigen levels were measured with the immunoturbidimetric method using the STA liatest kit (rabbit anti-human von Willebrand factor antibodies).

For genome-wide genotyping with the Illumina Human660-Quad Beadchip, we sampled 1,499 patients with a first episode of VT. Patients with a cancer diagnosis were excluded. Patients with genotyping success lower than 95% were excluded from the analyses as were patients showing discrepancies between their reported and genotypic sex. Individuals demonstrating a too high or low level of heterozygosity, close relatedness or non-European ancestry were also excluded. Variants showing significant (P-value $<1x10^{-6}$) deviation from HWE, with MAF less than 1%, and with genotyping call rate <98% were filtered out. After quality-control, Imputation to the 1000 Genomes population reference (Phase1 March 2012 release) was performed using IMPUTE2 software. Association analyses were conducted with the score test as implemented in SNPTEST, adjusting for age, sex, and, the first three principal components.

The Multi-Ethnic Study of Atherosclerosis (MESA) is a study of the characteristics of subclinical cardiovascular disease and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease[31]. MESA consisted of a diverse, population-based sample of an initial 6,814 asymptomatic men and women aged 45-84. 38 percent of the recruited participants were white, 28 percent African American, 22 percent Hispanic, and 12 percent Asian, predominantly of Chinese descent. Participants were recruited from six field centers across the United States: Wake Forest University, Columbia University, Johns Hopkins University, University of Minnesota, Northwestern University and University of California - Los Angeles. The first examination took place over two years, from July 2000 - July 2002. Only individuals without clinical CVD were eligible. Each participant received an extensive physical exam and determination of coronary calcification, ventricular mass and function, flow-mediated endothelial vasodilation, carotid intimal-medial wall thickness and presence of echogenic lucencies in the carotid artery, lower extremity vascular

insufficiency, arterial wave forms, electrocardiographic (ECG) measures, standard coronary risk factors, sociodemographic factors, lifestyle factors, and psychosocial factors. It was followed by four examination periods that were 17-20 months in length, and a sixth exam is currently taking place. Factor VIII levels were determined by measuring the clot time of a sample in factor VIII deficient plasma in the presence of activators utilizing the Sta-R analyzer (STA-Deficient VIII; Diagnostica Stago, Parsippany, NJ). The assay was performed at the Laboratory for Clinical Biochemistry Research (University of Vermont, Burlington, VT). Results are given as percent factor VIII. The normal plasma range of factor VIII in healthy adults ranges from 60-150%. The analytical coefficient of variation for the factor VIII assay was 10%. von Willebrand factor (VWF) was measured by immunoturbidimetric assay on the Sta-R analyzer (liatest VWF; Diagnostica Stago, Parsippany, NJ). The assay utilizes latex particles to which specific antibodies have been attached. In the presence of antigen (VWF), the particles agglutinate to form aggregates, which absorb more light. This increase in absorbance is proportional to the VWF present in the test sample. This assay was performed at the Laboratory for Clinical Biochemistry Research (University of Vermont, Burlington, VT). The results are presented as percent VWF. The expected range in healthy adults is 50-160%. Intra- and inter-assay analytical coefficients of variation were 3.7% and 4.5%, respectively.

The **Orkney Complex Disease Study (ORCADES)** is a family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with the high levels of endogamy historically. Data for participants aged 18-100 years, from a subgroup of ten islands, were used for this analysis. Fasting blood samples were collected and over 300 health-related phenotypes and environmental exposures were measured in each individual. All

participants gave informed consent and the study was approved by Research Ethics Committees in Orkney and Aberdeen. Subjects with self-reported non-European ancestry were excluded. The VWF antigen was measured by ELISA assays that used a double-antibody sandwich (DAKO, Copenhagen, Denmark).

The **PROspective Study of Pravastatin in the Elderly at Risk (PROSPER)** was a prospective multicenter randomized placebo-controlled trial to assess whether treatment with pravastatin diminishes the risk of major vascular events in elderly. Between December 1997 and May 1999, subjects were enrolled in Scotland (Glasgow), Ireland (Cork), and the Netherlands (Leiden). Men and women aged 70-82 years were recruited if they had pre-existing vascular disease or increased risk of such disease because of smoking, hypertension, or diabetes. A total number of 5,804 subjects were randomly assigned to pravastatin or placebo. A large number of prospective tests were performed including Biobank tests and cognitive function measurements. A whole genome wide screening has been performed in the sequential PHASE project with the use of the Illumina 660K beadchip. Of 5,763 subjects DNA was available for genotyping. Genotyping was performed with the Illumina 660K beadchip, after QC (call rate <95%) 5,244 subjects and 557,192 SNPs were left for analysis. Imputation was based on the March 2012 1000G release using the IMPUTE2 imputation software. For this study, 4927 PROSPER participants with data on VWF levels in citrate (ELISA) were analyzed.

The **Rotterdam Study (RS-I and RS-II)** is a prospective, population-based cohort study of determinants of several chronic diseases in older adults [32]. RS-I comprised 7,983 inhabitants of Ommoord, a district of Rotterdam in the Netherlands, who were 55 years or over. The baseline examination took place between 1990 and 1993. In 1999, the cohort was extended to

include 3011 inhabitants who reached the age of 55 years after the baseline examination and persons aged 55 years or older who migrated into the research area (RS-II). Subjects are of European ancestry based on their self-report. Factor VIII activity was measured with a one-stage clotting assay by using a mixture of micronized silica and phospholipids (Platelin LS, Biomerieux) and factor VIII-deficient plasma (Biopool). The plasma concentrations were expressed as percentage activity by relating the clotting time to a calibration curve constructed of a standardized control plasma. As a control, the pooled plasma of 50 healthy middle-aged persons was used and three control samples were run with each batch of study samples. The intra-assay CV was 2.9%, the inter-assay CV was 5.2%, and the reference range was 0.70-1.40 U/ml. Von Willebrand factor antigen was measured with an in-house ELISA with polyclonal rabbit anti-human VWF antibodies (DAKO). The intra-assay CV was 1.9%, inter-assay CV was 6.3%, and the reference range was 0.60-1.40 U/ml.

The **Trinity Student Study (TSS)** consists of a cohort of 2,524 healthy, ethnically Irish individuals, attending the University of Dublin, Trinity College, with ages between 18 and 28 years, recruited over one academic year in 2003–2004. Ethical approval was obtained from the Dublin Federated Hospitals Research Ethics Committee, which is affiliated with the Trinity College, and reviewed by the Office of Human Subjects Research at the National Institutes of Health. Written informed consent was obtained from participants before recruitment. VWF levels were determined at the University of Michigan using a custom AlphaLISA™ assay (Perkin-Elmer), which utilized a polyclonal anti-von Willebrand factor antibody from DAKO Cytomation.

REFERENCES:

1. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Patterson N, Daly MJ, Price AL and Neale

BM. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47:291-5.
2. Kwak IY and Pan W. Adaptive gene- and pathway-trait association testing with GWAS summary statistics. *Bioinformatics*. 2016;32:1178-84.
3. Ward LD and Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res*. 2016;44:D877-81.
4. Arnold M, Raffler J, Pfeufer A, Suhre K and Kastenmuller G. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics*. 2015;31:1334-6.
5. Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J, Lui JC, Vedantam S, Gustafsson S, Esko T, Frayling T, Speliotes EK, Genetic Investigation of ATC, Boehnke M, Raychaudhuri S, Fehrmann RS, Hirschhorn JN and Franke L. Biological interpretation of genome-wide association studies using predicted gene functions. *Nature communications*. 2015;6:5890.
6. Bowden J, Davey Smith G and Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512-25.
7. Bowden J, Davey Smith G, Haycock PC and Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016;40:304-14.
8. Burgess S, Bowden J, Fall T, Ingelsson E and Thompson SG. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology*. 2017;28:30-42.
9. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF, Hoffman HJ and Gudnason V. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol*. 2007;165:1076-87.
10. Danesh J, Wheeler JG, Hirschfield GM, Eda S, Eiriksdottir G, Rumley A, Lowe GD, Pepys MB and Gudnason V. C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. *N Engl J Med*. 2004;350:1387-97.
11. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *American journal of epidemiology*. 1989;129:687-702.
12. Folsom AR, Wu KK, Rosamond WD, Sharrett AR and Chambless LE. Prospective study of hemostatic factors and incidence of coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. *Circulation*. 1997;96:1102-8.
13. Chambless LE, McMahon R, Wu K, Folsom A, Finch A and Shen YL. Short-term intraindividual variability in hemostasis factors. The ARIC Study. Atherosclerosis Risk in Communities Intraindividual Variability Study. *Ann Epidemiol*. 1992;2:723-33.
14. Strachan DP, Rudnicka AR, Power C, Shepherd P, Fuller E, Davis A, Gibb I, Kumari M, Rumley A, Macfarlane GJ, Rahi J, Rodgers B and Stansfeld S. Lifecourse influences on health among British adults: effects of region of residence in childhood and adulthood. *Int J Epidemiol*. 2007;36:522-31.
15. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, Plagnol V, Pociot F, Schuilenburg H, Smyth DJ, Stevens H, Todd JA, Walker NM, Rich SS and Type 1 Diabetes Genetics C. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*. 2009;41:703-7.

16. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M, Cookson W and Consortium G. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med*. 2010;363:1211-1221.
17. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661-78.
18. Rudnicka AR, Rumley A, Lowe GD and Strachan DP. Diurnal, seasonal, and blood-processing patterns in levels of circulating fibrinogen, fibrin D-dimer, C-reactive protein, tissue plasminogen activator, and von Willebrand factor in a 45-year-old population. *Circulation*. 2007;115:996-1003.
19. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hulley SB, Jacobs DR, Jr., Liu K and Savage PJ. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *Journal of clinical epidemiology*. 1988;41:1105-16.
20. Green D, Ruth KJ, Folsom AR and Liu K. Hemostatic factors in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Arterioscler Thromb*. 1994;14:686-93.
21. Batlle J L-FM. Laboratory assays for von Willebrand factor. In: R. Z. Zimmerman TS, ed. *Coagulation and Bleeding Disorders* New York; 1989: 330-331.
22. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A and et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol*. 1991;1:263-76.
23. Cushman M, Cornell ES, Howard PR, Bovill EG and Tracy RP. Laboratory methods and quality assurance in the Cardiovascular Health Study. *Clin Chem*. 1995;41:264-70.
24. Feinleib M, Kannel WB, Garrison RJ, McNamara PM and Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med*. 1975;4:518-25.
25. Tofler GH, Massaro J, Levy D, Mittleman M, Sutherland P, Lipinska I, Muller JE and D'Agostino RB. Relation of the prothrombotic state to increasing age (from the Framingham Offspring Study). *Am J Cardiol*. 2005;96:1280-3.
26. Souto JC, Almasy L, Borrell M, Gari M, Martinez E, Mateo J, Stone WH, Blangero J and Fontcuberta J. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation*. 2000;101:1546-51.
27. Winkelmann BR, Marz W, Boehm BO, Zotz R, Hager J, Hellstern P, Senges J and Group LS. Rationale and design of the LURIC study--a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. *Pharmacogenomics*. 2001;2:S1-73.
28. Oudot-Mellakh T, Cohen W, Germain M, Saut N, Kallel C, Zelenika D, Lathrop M, Tregouet DA and Morange PE. Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol*. 2012;157:230-9.
29. Cohen W, Castelli C, Alessi MC, Aillaud MF, Bouvet S, Saut N, Brunet D, Barthet MC, Tregouet DA, Lavigne G and Morange PE. ABO blood group and von Willebrand factor levels partially explained the incomplete penetrance of congenital thrombophilia. *Arterioscler Thromb Vasc Biol*. 2012;32:2021-8.
30. Blom JW, Doggen CJ, Osanto S and Rosendaal FR. Malignancies, prothrombotic mutations, and the risk of venous thrombosis. *JAMA*. 2005;293:715-22.
31. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Jr., Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M and Tracy RP. Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am J Epidemiol*. 2002;156:871-81.

32. Hofman A, Brusselle GG, Darwish Murad S, van Duijn CM, Franco OH, Goedegebure A, Ikram MA, Klaver CC, Nijsten TE, Peeters RP, Stricker BH, Tiemeier HW, Uitterlinden AG and Vernooij MW. The Rotterdam Study: 2016 objectives and design update. *European journal of epidemiology*. 2015;30:661-708.