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

Nicholas L. Smith, PhD*; Ming-Huei Chen, PhD*; Abbas Dehghan, MD, DSc*; David P. Strachan, MD*; Saonli Basu, PhD*; Nicole Soranzo, PhD; Caroline Hayward, PhD; Igor Rudan, PhD; Maria Sabater-Lleal, PhD; Joshua C. Bis, PhD; Moniek P.M. de Maat, PhD; Ann Rumley, PhD; Xiaoxiao Kong, MS, ME; Qiong Yang, PhD; Frances M.K. Williams, PhD; Veronique Vitart, PhD; Harry Campbell, MD, PhD; Anders Målarstig, PhD; Kerri L. Wiggins, MS, RD; Comelia M. Van Duijn, PhD; Wendy L. McArdle, PhD; James S. Pankow, PhD, MPH; Andrew D. Johnson, PhD; Angela Silveira, PhD; Barbara McKnight, PhD; Andre G. Uitterlinden, PhD; Wellcome Trust Case Control Consortium; Nena Aleksic, PhD; James B. Meigs, MD, MPH; Annette Peters, PhD; Wolfgang Koenig, MD; Mary Cushman, MD, MSc; Sekar Kathiresan, MD; Jerome I. Rotter, MD; Edwin G. Bovill, MD; Albert Hofman, MD, PhD; Eric Boerwinkle, PhD; Geoffrey H. Tofler, MD; John F. Peden, PhD; Bruce M. Psaty, MD, PhD; Frank Leebeek, MD, PhD†; Aaron R. Folsom, MD, MPH; Martin G. Larson, SD; Timothy D. Spector, MD; Alan F. Wright, MD, PhD; James F. Wilson, DPhil; Anders Hamsten, MD, PhD; Thomas Lumley, PhD†; Jacqueline C.M. Witterman, PhD†; Weihong Tang, MD, PhD†; Christopher J. O’Donnell, MD, MPH†

Background—Plasma levels of coagulation factors VII (FVII), VIII (FVIII), and von Willebrand factor (vWF) influence risk of hemorrhage and thrombosis. We conducted genome-wide association studies to identify new loci associated with plasma levels.

Methods and Results—The setting of the study included 5 community-based studies for discovery comprising 23,608 European-ancestry participants: Atherosclerosis Risk In Communities Study, Cardiovascular Health Study, British 1958 Birth Cohort, Framingham Heart Study, and Rotterdam Study. All subjects had genome-wide single-nucleotide polymorphism (SNP) scans and at least 1 phenotype measured: FVII activity/antigen, FVIII activity, and vWF antigen. Each study used its genotype data to impute to HapMap SNPs and independently conducted association analyses of hemostasis measures using an additive genetic model. Study findings were combined by meta-analysis. Replication was conducted in 7,604 participants not in the discovery cohort. For FVII, 305 SNPs exceeded the genome-wide significance threshold of 5.0 × 10⁻⁸ and comprised 5 loci on 5 chromosomes: 2p23 (smallest P value 6.2 × 10⁻²⁴), 4q25 (3.6 × 10⁻¹²), 11q12 (2.0 × 10⁻¹⁰), 13q34 (9.0 × 10⁻²⁵⁰), and 20q11.2 (5.7 × 10⁻³⁷). Loci were within or near genes, including 4 new candidate genes and F7 (13q34). For vWF, 400 SNPs exceeded the threshold and marked 8 loci on 6 chromosomes: 6q24 (1.2 × 10⁻⁵), 8p21 (1.3 × 10⁻⁶), 9q34 (5.0 × 10⁻³²⁴), 12p11 (1.7 × 10⁻³²), 12q23 (7.3 × 10⁻¹⁰), 12q43 (3.8 × 10⁻¹³), 14q32 (2.3 × 10⁻¹⁰), and 19p13.2 (1.3 × 10⁻³⁷). All loci were within genes, including 6 new candidate genes, as well as ABO (9q34) and VWF (12p13). For FVIII, 5 loci were identified and overlapped vWF findings. Nine of the 10 new findings were replicated.

Conclusions—New genetic associations were discovered outside previously known biological pathways and may point to novel prevention and treatment targets of hemostasis disorders. (Circulation. 2010;121:1382-1392.)

Key Words: genetic variation ■ factor VII ■ factor VIII ■ von Willebrand factor ■ epidemiology ■ meta-analysis ■ thrombosis ■ hemostasis
A complex cascade of coagulation factors underlies hemostasis and prevents life-threatening blood loss from damaged blood vessels. The hemostatic factors VII and VIII, both produced in the liver, play central roles in the initiation and propagation, respectively, of fibrin formation. In the tissue-factor pathway, blood coagulation factor VII (FVII), once activated, serves as a catalyst for factor X (FX) activation, which converts prothrombin to thrombin. During propagation, activated factor VIII (FVIII) activates FX in the presence of activated factor IX. Von Willebrand factor (vWF), produced by endothelial cells and megakaryocytes, has multiple roles in hemostasis. Its primary role is to serve as an adhesion molecule that anchors platelets to exposed collagen after endothelial cell damage. The factor also acts as a carrier protein of FVIII, thereby prolonging the half-life of FVIII.

Clinical Perspective on p 1392

Elevated circulating levels of FVII and vWF are risk factors for venous thrombosis, but the data supporting an association of FVII levels with arterial thrombosis are less consistent.1–5 Hemorrhagic complications are associated with deficiency in FVII and vWF (von Willebrand disease), as well as X-linked deficiency in FVIII (hemophilia A).6–9 Plasma levels of these proteins are affected by environmental factors, but they also are genetically influenced.10–13 Heritability estimates range from 0.53 to 0.63 for FVII, from 0.40 to 0.61 for FVIII, and from 0.31 to 0.75 for vWF.12,13 To date, our understanding of the way genetic variation influences plasma levels has been focused primarily on cis-acting variation in the genes encoding each protein product (F7, F8, and VWF, respectively). A large-scale genome-wide investigation of the genomic correlates of plasma levels has not been published previously. Using data from 23,608 adults, we investigated genome-wide associations between common genetic variation and plasma levels of FVII, FVIII, and vWF.

Methods

Setting

The discovery meta-analysis was conducted in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, which includes data from 5 prospective, population-based cohorts of adults in the United States and Europe.14 Phenotype data were available from 4 of the cohorts: The Atherosclerosis Risk in Communities (ARIC), the Cardiovascular Health Study (CHS), the Framingham Heart Study (FHS), and the Rotterdam Study (RS). In addition, participants from the British 1958 Birth Cohort (B58C) who had genome-wide data and were used as controls for the Wellcome Trust Case-Control Consortium were included in the present analyses.15 The designs of these studies have been described elsewhere.15–22

Subjects

Eligible participants for these analyses had high-quality data from the genome-wide scans (see below), had at least 1 of the 3 phenotypes measured, and were not using a coumarin-based anticoagulant at the time of the phenotype measurement. Participants were of European (n = 23,608) ancestry by self-report. Each study received institutional review board approval, and all participants provided written informed consent for the use of their DNA in research.

Measures

Coagulation Measures

Plasma measures of FVII, FVIII, and vWF were obtained at the time of cohort entry for ARIC, CHS, and RS (except vWF, which was measured at visit 3) and at a follow-up visit for B58C (examination 2002 and 2003) and FHS (examination cycle 5, 1991 to 1995). Only the FVII phenotype in CHS was measured more than once, and this included 2 measures, which were averaged for 2653 participants (81% of 3266 participants). The FVII phenotype was measured in 4 cohorts and included both antigen (FHS23) and activity (ARIC,24 CHS,26 and RS27). FVIII activity was measured in 3 cohorts (ARIC,24,25 CHS,26 and RS), and vWF antigen was measured in 4 cohorts (ARIC,24,25 B58C,28 FHS,29 and RS). Table I in the online-only Data Supplement provides details about assays in each of the 5 cohorts.

Baseline measures of clinical and demographic characteristics were obtained at the time of cohort entry for ARIC, CHS, and RS and at the time of phenotype measurement for B58C and FHS. Measures, taken by use of standardized methods as specified by each study, included in-person measures of height and weight, as well as self-reported treatment of diabetes and hypertension, current alcohol consumption, and prevalent cardiovascular diseases (history of myocardial infarction, angina, coronary revascularization, stroke, or transient ischemic attack).

Genotyping and Imputation

Genotyping used DNA collected from phlebotomy (ARIC, CHS, FHS, and RS) or cell lines (B58C). Genome-wide assays of single-nucleotide polymorphisms (SNPs) were conducted independently in each cohort with various technologies: Affymetrix 6.0 for ARIC, Affymetrix 500K for B58C, Illumina 370CNV for CHS, Affymetrix 500K and gene-centric 50K for FHS, and Illumina 550K version 3 for RS. Genotype quality control and data cleaning that included assessment of Hardy-Weinberg equilibrium and variant call rates were conducted independently by each study; details have been published elsewhere14 and are also provided in Table II of the online-only Data Supplement.

We investigated genetic variation in the 22 autosomal chromosomes and the X chromosome. Genotypes were coded as 0, 1, and 2 to represent the number of copies of the coded alleles for all chromosomes except the X chromosome in males, for which genotypes were coded as 0 and 2.30 Each study independently imputed its genotype data to the ~2.6 million SNPs identified in the HapMap Caucasian (CEU) sample from the Centre d’Etude du Polymorphisme Humain.31–33 Imputation software that included MACH, BIMBAM, or IMPUTE, with SNPs that passed quality-control criteria (online-only Data Supplement Table II), was used among studies to impute unmeasured genotypes based on phased haplotypes observed in HapMap. Imputed data for the X chromosome were not available in ARIC or RS. Imputation results were summarized as an “allele dosage,” defined as the expected number of copies of the minor allele at that SNP (a continuous value between 0 and 2) for each genotype. Each cohort calculated a ratio of observed to expected variance of the dosage statistic for each SNP. This value, which generally ranges from 0 to 1 (ie, from poor to excellent), reflects imputation quality.

Statistical Analyses

Investigators from all cohorts developed the prespecified analytic plan described below. Each study independently analyzed its genotype-phenotype data. All studies used linear regression to conduct association analyses between measured and imputed SNPs and untransformed phenotype measures except for FHS, which used a linear mixed-effects model to account for family relationships.34 An additive genetic model with 1 degree of freedom was adjusted for age and sex. In addition, CHS and ARIC adjusted for field site and FHS adjusted for generation and ancestry using principal components.35 For each analysis, a genomic control coefficient, which estimated the extent of underlying population structure on the basis of test-statistic inflation, was used to adjust standard errors.36

Downloaded from http://circ.ahajournals.org/ by guest on April 20, 2017
For each phenotype, within-study findings were combined across studies to produce summary results by use of standard meta-analytic approaches. For vWF antigen and FVIII activity levels, fixed-effects, inverse-variance–weighted meta-analysis was performed, and summary $P$ values and $\beta$-coefficients were calculated. The parameter coefficient represents change (% vWF antigen or % FVIII activity) associated with a 1-unit change in allele dosage. This approach was not appropriate for the FVII phenotype, which combined analysis of antigen and activity measures. Instead, effective-sample-size–weighted meta-analysis was performed to estimate summary $P$ values only. All meta-analyses were conducted with METAL software (http://www.sph.umich.edu/csg/abecasis/METAL/index.html). For loci that contained genes already known to be associated with the phenotype, we conducted secondary analyses, adjusting for 1 or more SNPs within the known gene. This allowed us to assess possible novel associations in neighboring genes independently of the strong signal in the known SNPs.

The a priori threshold of genome-wide significance was set at $P=5.0 \times 10^{-8}$. When more than 1 SNP clustered at a locus, we picked the SNP with the smallest $P$ value to represent the locus and to be used for replication. For each phenotype, the amount of variation explained by the top SNPs was the difference in the $R^2$ value when a model with adjustment variables only was compared with a model that also included top genome-wide significant SNPs. Each study calculated the amount of variation explained, and these estimates were combined across cohorts by use of weighted averages. Linkage disequilibrium (LD) between SNPs was estimated with a model that also included top genome-wide significant SNPs.

Cohort-specific $P$ values were subjected to meta-analysis with sample-size (FVII) values and \( \beta \)-coefficients were calculated. The parameter $R^2$ value to represent the locus and $P$ value when a model with adjustment variables only was compared to the model with adjustment variables only. All meta-analyses were conducted with METAL software (http://www.sph.umich.edu/csg/abecasis/METAL/index.html). For loci that contained genes already known to be associated with the phenotype, we conducted secondary analyses, adjusting for 1 or more SNPs within the known gene. This allowed us to assess possible novel associations in neighboring genes independently of the strong signal in the known SNPs.

For each phenotype, within-study findings were combined across studies to produce summary results by use of standard meta-analytic approaches. For vWF antigen and FVIII activity levels, fixed-effects, inverse-variance–weighted meta-analysis was performed, and summary $P$ values and $\beta$-coefficients were calculated. The parameter coefficient represents change (% vWF antigen or % FVIII activity) associated with a 1-unit change in allele dosage. This approach was not appropriate for the FVII phenotype, which combined analysis of antigen and activity measures. Instead, effective-sample-size–weighted meta-analysis was performed to estimate summary $P$ values only. All meta-analyses were conducted with METAL software (http://www.sph.umich.edu/csg/abecasis/METAL/index.html). For loci that contained genes already known to be associated with the phenotype, we conducted secondary analyses, adjusting for 1 or more SNPs within the known gene. This allowed us to assess possible novel associations in neighboring genes independently of the strong signal in the known SNPs.

The a priori threshold of genome-wide significance was set at $P=5.0 \times 10^{-8}$. When more than 1 SNP clustered at a locus, we picked the SNP with the smallest $P$ value to represent the locus and to be used for replication. For each phenotype, the amount of variation explained by the top SNPs was the difference in the $R^2$ value when a model with adjustment variables only was compared to the model with adjustment variables only. All meta-analyses were conducted with METAL software (http://www.sph.umich.edu/csg/abecasis/METAL/index.html). For loci that contained genes already known to be associated with the phenotype, we conducted secondary analyses, adjusting for 1 or more SNPs within the known gene. This allowed us to assess possible novel associations in neighboring genes independently of the strong signal in the known SNPs.

The a priori threshold of genome-wide significance was set at $P=5.0 \times 10^{-8}$. When more than 1 SNP clustered at a locus, we picked the SNP with the smallest $P$ value to represent the locus and to be used for replication. For each phenotype, the amount of variation explained by the top SNPs was the difference in the $R^2$ value when a model with adjustment variables only was compared to the model with adjustment variables only. All meta-analyses were conducted with METAL software (http://www.sph.umich.edu/csg/abecasis/METAL/index.html). For loci that contained genes already known to be associated with the phenotype, we conducted secondary analyses, adjusting for 1 or more SNPs within the known gene. This allowed us to assess possible novel associations in neighboring genes independently of the strong signal in the known SNPs.
Figure. A through C, Genome-wide –log10 P value plots of FVII, FVIII, and vWF. The horizontal line marks the 5.0×10⁻⁸ P value threshold of genome-wide significance.

of variance in the FVII phenotype explained by the 5 loci was 7.7%.

Genome-wide significant signals at chromosomal position 2p23 (online-only Data Supplement Figure I) were within or close to 1 gene, GCKR (glucokinase [hexokinase 4] regulator). Rs1260326 was associated with the smallest P value (6.2×10⁻²⁴; overall MAF=0.422) in this region, and this SNP codes a nonsynonymous substitution (proline to leucine) at amino acid position 446 in GCKR. Genome-wide significant signals at chromosomal position 4q25 (online-only Data Supplement Figure II) were within the region of 2 genes, ADH4 (alcohol dehydrogenase 4 [class II], pi polypeptide) and ADH5 (alcohol dehydrogenase 5 [class III], chi polypeptide). Rs1126670, intronic to ADH4, had the smallest P value (3.6×10⁻¹²; MAF=0.316). This SNP was in high LD ($r^2=0.77$) with the top SNP in ADH5 (rs896992; $P=3.2×10^{-13}$). At chromosomal position 11q12 (online-only Data Supplement Figure III), genome-wide significant signals were within 2 genes, MS4A2 and MS4A6A (membrane-spanning 4 domains, subfamily A, member 2 and member 6A, respectively). Rs11230180 had the smallest P value (2.0×10⁻¹⁰) and was close to the MS4A6A gene and in high LD with top SNPs in MS4A6A ($r^2=0.97$ for rs17602572 [7.3×10⁻¹⁰]) and in MS4A2 ($r^2=0.84$ for rs2847666 [3.8×10⁻⁸]).

The 63 genome-wide significant signals at chromosomal position 13q34 (online-only Data Supplement Figure IV) were within the FVII structural gene (F7) and several other genes. Rs488703 had the smallest P value ($9.0×10^{-259}$) and is located in an F7 intron. After adjustment for rs6046 (4.4×10⁻²⁵⁹), a missense SNP that leads to the arginine-glutamine (R353Q) functional substitution, 22 SNPs in the 13q34 region retained their genome-wide significance. These SNPs were within 2 genes, rs3211727 (5.3×10⁻²³) in F10 (coagulation factor X) and rs1755693 (3.1×10⁻¹⁹) in MCF2L (MCF.2 cell line–derived transforming sequence-like).

The final chromosomal region that contained genome-wide signals for FVII levels was position 2q11.2 (online-only Data Supplement Figure V). Rs867186 in PROCR had the smallest P value (5.7×10⁻³⁷), and this variant led to the serine-glycine (S219G) substitution in exon 4. The region also covered another 7 genes that contained SNPs with P values that exceeded genome-wide significance: ITCH (itchy E3 ubiquitin protein ligase homolog), PIGU (phosphatidylinositol glycan anchor biosynthesis, class U), ACS52 (acyl-CoA synthetase short-chain family member 2), MYH7B (myosin, heavy chain 7B, cardiac muscle, beta), TRPC4AP (transient receptor potential cation channel, subfamily C, member 4 associated protein), EDEM2 (ER [endoplasmic reticulum] degradation enhancer, mannosidase α-like 2), and MMP24 (matrix metallopeptidase 24). As depicted in online-only Data Supplement Figure V, all top SNPs were in strong LD with rs867186.

Across all studies, significant heterogeneity of effect ($P<0.01$) was detected for rs1126670 and rs488703. Effect estimates were in the same direction, but magnitudes differed between studies (Table 3). When we restricted the discovery cohort to include only those studies that measured FVII activity (ARIC, CHS, and RS), $P$ values weakened for rs1260326 (1.7×10⁻¹⁹) and rs867186 (2.8×10⁻³⁴), were virtually unchanged for rs11230180 and rs488703, and were strengthened for rs1126670 (4.4×10⁻¹³). A sixth locus emerged with 4 genome-wide significant SNPs at 11q13. Among these, rs1149606 had the smallest P value at 1.9×10⁻⁹ (MAF=0.176). This SNP was 4.0 kb from TSKU (tsukushin) and had a P value of 2.2×10⁻⁶ in the full
Replication
The 4 new findings were tested in the ARIC, TUK (1 and 2), FVII Replication discovery cohort. When we restricted the discovery cohort to the single study that measured FVII antigen (FHS), there were no SNPs that reached genome-wide significance.

FVII Replication
The 4 new findings were tested in the ARIC, TUK (1 and 2), and Swedish PROCARDIS cohorts and were all replicated. Replication P values were 6.5×10^{-4} for rs1260326, 3.8×10^{-2} for rs1126670, 1.4×10^{-3} for rs11230180, and 1.6×10^{-7} for rs867186 (online-only Data Supplement Table IIIa). When we separated the FVII activity (ARIC, TUK-1) and FVII antigen (TUK-2 and PROCARDIS) cohorts, replication results were similar, except rs1260326 did not replicate in the activity cohorts (P=9.0×10^{-7}).

vWF and FVIII Discovery
Within each cohort that conducted genome-wide association analyses for vWF (ARIC, B5SC, FHS, and RS) and FVIII (ARIC, CHS, and RS), the genomic control coefficients for analyses were small (<1.03), which suggests negligible test-statistic inflation. Panels B and C in the Figure present all meta-analysis P values (2 729 294 for FVIII and 2 742 821 for vWF) organized by chromosome and genomic position.

For vWF, 400 SNPs exceeded the genome-wide significance threshold and marked 8 regions on 6 chromosomes:

<table>
<thead>
<tr>
<th>Region</th>
<th>SNP</th>
<th>Allele</th>
<th>MAF</th>
<th>P</th>
<th>Parameter Coefficient</th>
<th>Closest Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>6q24</td>
<td>rs9390459</td>
<td>G→A</td>
<td>0.442</td>
<td>1.2×10^{-22}</td>
<td>-4.8 (−7.5, −2.1)</td>
<td>0.0 kb from STXBPs (exon)†</td>
</tr>
<tr>
<td>8q21</td>
<td>rs2726953</td>
<td>C→T</td>
<td>0.309</td>
<td>1.3×10^{-16}</td>
<td>4.5 (1.5, 7.5)</td>
<td>0.0 kb from SCARAS (intron)†</td>
</tr>
<tr>
<td>9q34</td>
<td>rs687621</td>
<td>T→C</td>
<td>0.337</td>
<td>&lt;5.0×10^{-24}</td>
<td>24.1 (21.4, 26.8)</td>
<td>0.0 kb from ABO (intron)</td>
</tr>
<tr>
<td>12p13</td>
<td>rs1603857</td>
<td>C→T</td>
<td>0.357</td>
<td>1.7×10^{-32}</td>
<td>6.0 (3.2, 8.8)</td>
<td>0.0 kb from VWF (exon)</td>
</tr>
<tr>
<td>12q23</td>
<td>rs4981022</td>
<td>C→T</td>
<td>0.315</td>
<td>7.3×10^{-10}</td>
<td>-3.6 (−6.8, −0.4)</td>
<td>0.0 kb from STAB2 (intron)†</td>
</tr>
<tr>
<td>12q24</td>
<td>rs7979987</td>
<td>G→A</td>
<td>0.346</td>
<td>3.8×10^{-11}</td>
<td>3.4 (0.6, 6.2)</td>
<td>0.0 kb from STX2 (intron)†</td>
</tr>
<tr>
<td>14q32</td>
<td>rs10133762</td>
<td>G→T</td>
<td>0.444</td>
<td>2.3×10^{-10}</td>
<td>3.1 (0.4, 5.8)</td>
<td>0.0 kb from TC2N (intron)†</td>
</tr>
<tr>
<td>19p13.2</td>
<td>rs8868875</td>
<td>A→G</td>
<td>0.262</td>
<td>1.3×10^{-9}</td>
<td>−4.0 (−7.6, −0.4)</td>
<td>0.0 kb from CLEC4M (3′ UTR)†</td>
</tr>
</tbody>
</table>

CI indicates confidence interval based on a 2-sided α=0.00000005; NE, not estimated because antigen and antibody measures were combined.

*Parameter coefficient represents change (% of activity or antigen) associated with 1-unit change in allele dosage.
†Novel associations.
Table 4. Cohort-Specific Results for the Top Loci Associated With vWF

<table>
<thead>
<tr>
<th>CHR</th>
<th>SNP</th>
<th>Variant</th>
<th>ARIC β</th>
<th>ARIC P</th>
<th>BSBC β</th>
<th>BSBC P</th>
<th>FHS β</th>
<th>FHS P</th>
<th>RS I β</th>
<th>RS I P</th>
<th>RS II β</th>
<th>RS II P</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>rs9390459</td>
<td>G→A</td>
<td>-4.1</td>
<td>4.1×10^{-10}</td>
<td>-6.9</td>
<td>4.0×10^{-9}</td>
<td>-4.9</td>
<td>4.2×10^{-5}</td>
<td>-5.7</td>
<td>1.5×10^{-4}</td>
<td>-5.2</td>
<td>2.4×10^{-3}</td>
</tr>
<tr>
<td>8</td>
<td>rs2726953</td>
<td>C→T</td>
<td>5.0</td>
<td>3.8×10^{-12}</td>
<td>1.1</td>
<td>5.2×10^{-7}</td>
<td>5.1</td>
<td>3.6×10^{-5}</td>
<td>3.4</td>
<td>3.6×10^{-3}</td>
<td>5.8</td>
<td>2.3×10^{-3}</td>
</tr>
<tr>
<td>9</td>
<td>rs687621</td>
<td>T→C</td>
<td>23.5</td>
<td>6.0×10^{-27}</td>
<td>33.6</td>
<td>4.0×10^{-24}</td>
<td>17.4</td>
<td>7.0×10^{-48}</td>
<td>29.4</td>
<td>1.8×10^{-83}</td>
<td>27.5</td>
<td>7.9×10^{-56}</td>
</tr>
<tr>
<td>12</td>
<td>rs1063857</td>
<td>T→C</td>
<td>6.1</td>
<td>4.1×10^{-19}</td>
<td>7.2</td>
<td>3.5×10^{-9}</td>
<td>3.8</td>
<td>2.2×10^{-3}</td>
<td>5.5</td>
<td>4.0×10^{-9}</td>
<td>8.7</td>
<td>6.6×10^{-7}</td>
</tr>
<tr>
<td>12</td>
<td>rs4981022</td>
<td>T→C</td>
<td>-3.4</td>
<td>1.1×10^{-6}</td>
<td>-4.3</td>
<td>2.7×10^{-2}</td>
<td>-1.1</td>
<td>5.0×10^{-1}</td>
<td>-3.6</td>
<td>2.3×10^{-2}</td>
<td>-6.9</td>
<td>1.4×10^{-4}</td>
</tr>
<tr>
<td>12</td>
<td>rs7978987</td>
<td>G→A</td>
<td>3.9</td>
<td>1.2×10^{-8}</td>
<td>2.1</td>
<td>1.6×10^{-1}</td>
<td>2.9</td>
<td>2.2×10^{-2}</td>
<td>3.2</td>
<td>4.3×10^{-2}</td>
<td>2.7</td>
<td>1.3×10^{-1}</td>
</tr>
<tr>
<td>14</td>
<td>rs10133762</td>
<td>G→T</td>
<td>3.2</td>
<td>7.7×10^{-7}</td>
<td>5.2</td>
<td>4.4×10^{-4}</td>
<td>2.5</td>
<td>3.5×10^{-2}</td>
<td>0.9</td>
<td>5.5×10^{-1}</td>
<td>3.5</td>
<td>4.1×10^{-2}</td>
</tr>
<tr>
<td>19</td>
<td>rs868875</td>
<td>A→G</td>
<td>-3.8</td>
<td>8.7×10^{-9}</td>
<td>-3.1</td>
<td>3.9×10^{-1}</td>
<td>-2.9</td>
<td>2.8×10^{-2}</td>
<td>-5.1</td>
<td>2.2×10^{-7}</td>
<td>-8.6</td>
<td>9.5×10^{-4}</td>
</tr>
</tbody>
</table>

CHR indicates chromosome; β-coefficient, change (% of activity) associated with 1-unit change in allele dosage; RS I, original Rotterdam Study cohort; and RS II, extended Rotterdam Study cohort.

Table 5. Cohort-Specific Results for the Top Loci Associated With FVIII

<table>
<thead>
<tr>
<th>CHR</th>
<th>SNP</th>
<th>Variant</th>
<th>ARIC β</th>
<th>ARIC P</th>
<th>BSBC β</th>
<th>BSBC P</th>
<th>FHS β</th>
<th>FHS P</th>
<th>RS I β</th>
<th>RS I P</th>
<th>RS II β</th>
<th>RS II P</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>rs9390459</td>
<td>G→A</td>
<td>-2.4</td>
<td>9.1×10^{-6}</td>
<td>-3.1</td>
<td>5.5×10^{-4}</td>
<td>-3.2</td>
<td>4.9×10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>rs11780263</td>
<td>G→A</td>
<td>-4.2</td>
<td>2.0×10^{-10}</td>
<td>-4.9</td>
<td>7.1×10^{-6}</td>
<td>-2.6</td>
<td>6.8×10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>rs6872989</td>
<td>C→T</td>
<td>17.2</td>
<td>4.1×10^{-22}</td>
<td>19.0</td>
<td>3.0×10^{-6}</td>
<td>17.6</td>
<td>1.4×10^{-50}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>rs1063856</td>
<td>A→G</td>
<td>2.9</td>
<td>1.6×10^{-7}</td>
<td>3.4</td>
<td>2.4×10^{-4}</td>
<td>0.0</td>
<td>9.8×10^{-1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>rs12229292</td>
<td>G→T</td>
<td>3.5</td>
<td>2.1×10^{-7}</td>
<td>3.3</td>
<td>2.5×10^{-3}</td>
<td>1.1</td>
<td>4.2×10^{-1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CHR indicates chromosome; β-coefficient, change (% of activity) associated with 1-unit change in allele dosage; RS I, original Rotterdam Study cohort.

All observed-expected variances (OEV) measuring imputation quality exceeded 0.95 except those marked with *(OEV range from 0.80–0.94), †(OEV range from 0.50–0.79), and ‡(OEV range from 0.309).
wide significant SNPs at chromosomal position 19p13.2 (online-only Data Supplement Figure XIII) were located within 1 gene, CLEC4M (C-type lectin domain family 4, member M). Rs868875 had the smallest P value (1.3 x 10^-9; MAF=0.262) and was intronic to CLEC4M.

The change in vWF associated with each additional copy of the minor allele was large for the ABO locus (increase of 24.1 in vWF antigen percent) and was smaller for the other SNPs of genome-wide significance, ranging from 3.1 to 6.0 (Table 2). These change values were generally smaller for the FVIII phenotype than for vWF. Only the ABO locus (rs687621) demonstrated heterogeneity of effect size across studies (Table 4).

**vWF Replication**

The 6 new findings were tested in the ARIC, B58C, TUK (1 and 2), VIS, and ORCADES cohorts, and 5 were replicated. Replicated P values were 1.7 x 10^-6 for rs9390459, 3.1 x 10^-5 for rs2726953, 8.7 x 10^-8 for rs4981022, 1.9 x 10^-1 for rs7978987, 3.3 x 10^-5 for rs10133762, and 1.4 x 10^-2 for rs868875 (online-only Data Supplement Table IIIb).

**Discussion**

In addition to confirming previously known associations of FVII levels with F7 variation and associations of vWF and FVIII levels with ABO and VWF variation, respectively, the present meta-analysis of data from 23 608 subjects of European ancestry identified novel, genome-wide significant associations of 4 loci with circulating FVII levels, 6 loci with vWF levels, and 3 loci with FVIII levels. The FVIII loci were a subset of the vWF loci. Several of these discoveries were associated with P values many magnitudes smaller than the threshold of genome-wide significance and included new candidate genes for levels of FVII (GCKR, ADH4, MS4A6A, and PROC1), vWF (STXBP5, SCARA5, STAB2, STX2, TC2N, and CLEC4M), and FVIII (STXBP5, SCARA5, and STAB2). There was independent evidence for replication of 9 of 10 new findings in new populations.

**Genes Associated With Genome-Wide Significant Findings: FVII**

The strongest genetic associations with FVII levels resided in F7 and confirm previously reported associations of SNPs in this locus.42,43 After adjustment for rs6046 (R353Q), a residual signal remained in this region with SNPs in MCF2L, F7, and F10 that reached genome-wide significance. On further exploration, the residual signal was only detected in 1 of the 4 cohorts (ARIC), in which the rs6046 SNP was imputed with only modest precision (observed-to-expected variance of 0.607). Although the residual signal is likely attributable to the incomplete adjustments, we cannot rule out the possibility that other sequence variations in the 13q34 region contribute to FVII levels.44

In addition to the manuscript by Ireland,45 this is the first genome-wide association study to attempt to discover novel genetic associations with the 3 hemostasis phenotypes: FVII, FVIII, and vWF. Recent reports have described activated FVII (FVIIa) binding to the endothelial protein C receptor via the Gla domain, which has a homologous counterpart on the protein C molecule.45,46 This binding appears to inhibit procoagulant activity of the FVIIa-tissue factor complex and may impact the clearance of FVIIa and receptor signaling through competition with the binding site. Variation in GCKR, whose gene product inhibits glucokinase in liver and pancreatic islet cells, has not been associated previously with FVII levels but has been associated with C-reactive protein and triglyceride levels.47,48 FVII protein concentrations are associated with triglyceride levels in the fasting state, but there is a direct effect of plasma triglycerides on FVII activity.49 The postprandial increase in plasma triglycerides is directly correlated with an increase in FVII activity, but the protein concentration remains constant. Genetic variation in alcohol dehydrogenase 3 has been associated with risk of myocardial infarction and coronary heart disease.50,51 Plasma levels of FVII are negatively correlated with alcohol consumption and ADH4, a gene that regulates alcohol metabolism and may have an effect on FVII levels. There are no previous reports of associations of genetic variation in ADH4, ADH5, or other alcohol dehydrogenase genes with levels of FVII. Little has been published about the MS4A6A gene.52

When comparing findings by the different measurements of FVII, either antigen or activity level, we found little evidence that the genetic predictors we identified differed by subphenotype. Nonetheless, a more extensive comparison of the 2 subphenotypes is merited and would require a larger representation of FVII antigen measures in the discovery cohort.

**Genes Associated With Genome-Wide Significant Findings: vWF and FVIII**

Because the vWF molecule transports FVIII and the 2 factors are highly correlated, we expected overlap in some of the vWF and FVIII findings.11 No unique genetic predictors of FVIII levels on the 22 autosomal chromosomes or the X chromosome, where F8 is located, were identified. This may indicate that in healthy individuals, genetic determinants of FVIII levels are primarily dependent on vWF levels.

The strong associations between SNPs located in ABO and VWF and FVIII levels were expected, because it is known that individuals with blood group O have 25% to 30% lower vWF levels than individuals with non-O blood groups.53 The 2 SNPs in VWF with the smallest P values for FVIII (rs1063856) and vWF (rs1063857) levels are in strong LD and mark nonsynonymous and synonymous substitutions at amino acids 789 and 795, respectively. These variations are located in exon 18, which encodes for the D′ domain, and are involved in binding of FVIII. Rare genetic variation in VWF has been associated with low levels of vWF, which characterize von Willebrand disease, either through lowering protein expression or through increased clearance. Rare missense mutations in the D′ part of the VWF gene have been associated with normal or reduced levels of vWF and low levels of FVIII, which is characteristic of von Willebrand disease type 2N.54,55

We detected a novel and relatively strong association for the STXB5 gene and also for STX2, the binding substrate for STXB5, with vWF and FVIII levels. We were not able to replicate the STX2 finding. Syntaxin 4, 1 of the soluble NSF
attachment protein receptor (SNARE) proteins, has been linked to exocytosis of Weibel-Palade bodies in endothelial cells by targeting Weibel-Palade bodies to the plasma membrane.56 Knockdown of syntaxin 4 resulted in inhibition of exocytosis of Weibel-Palade bodies.57 vWF is the main constituent of Weibel-Palade bodies; and on stimulation (for instance, due to stress or after endothelial damage), Weibel-Palade bodies are excreted by endothelial cells and give rise to increased plasma vWF levels. Syntaxin 2 and v- and vertebrate STXB5 have not been shown to be involved in vWF secretion; however, our observations may indicate a functional role for these proteins in determining circulating vWF levels.

SCARA5 belongs to the group of scavenger molecules that have as their primary function the initiation of immune responses. SCARA5 recently has been characterized in more detail and is solely expressed in epithelial cells. It has not yet been linked to vWF, and the nature of the genetic associations found in the present study is yet unknown. STAB2 is a transmembrane receptor protein primarily expressed in liver and spleen sinusoidal endothelial cells. The receptor binds various ligands, including heparin, LDL, bacteria, and advanced glycosylation products, and subjects them to endocytosis.58 Taken together, these results generate the hypothesis that several loci may be involved in vWF and FVIII clearance or uptake.

Relevance for Cardiovascular Outcomes

The relevance of the reported associations of genome-wide variation with levels of hemostasis proteins to cardiovascular end points will require expanded research efforts. Elevated circulating levels of FVIII and vWF are independent risk factors for venous thrombosis, and the genetic variants identified here explained roughly 10% of the variance in these measures.63 Variation in ABO is a known risk factor for venous thrombosis, and the risk associated with the newly identified candidate genes presented here can only be hypothesized.69,70 The association between FVII levels and thrombotic risk is less clear, and epidemiological findings have not been consistent.1–3,5 Similarly, published evidence supporting the role of $F7$ variation in atherothrombotic risk has been heterogeneous, and an exploration of newly identified candidate genes in the present report has not been undertaken.61,62

Strengths and Limitations

This study included 23 608 individuals of European ancestry and ~2.6 million markers spread throughout the genome. Data came from 5 large, population-based cohort studies in which cardiovascular outcomes were primary study endpoints. Phenotypes were measured in a standardized fashion within each cohort, but measures between cohorts differed and likely introduced between-group variability. This variability may decrease statistical power to find associations of smaller magnitudes. Online-only Data Supplement Table IV lists the minimum percent change in the hemostasis measures for variants with a range of MAFs and with an assumption of power of 80% or greater and an $\alpha$-level of $5.0 \times 10^{-8}$. For an SNP with a MAF of 0.05, we had sufficient power to detect a 5% change in FVII, a 6% change in FVIII, and a 7% change in vWF. Not all SNPs tested were directly genotyped, and imputation quality varied across SNPs. For poorly imputed SNPs, there was reduced statistical power to detect an association. For each identified locus, we chose the SNP with the smallest $P$ value, but the causal variant, if one exists, need not be the one with the smallest $P$ value, and effect sizes associated with the SNP may be an overestimate of any true effect. In situations in which genes clustered near a gene that contained 1 or more variants that were strongly associated with the phenotype, we used statistical adjustment to determine whether neighboring genes had independent associations with the phenotype. This approach has limitations and cannot indentify novel causal variants that are in LD with known causal variants.

Summary

Using data from 5 community-based cohorts that included 23 608 participants, we identified 4 novel genetic associations with FVII activity and antigen levels and 6 novel genetic associations with vWF antigen levels; 9 of these 10 findings were replicated. New candidate genes of interest were discovered outside known biological pathways that influence these essential hemostasis factors, and our findings may point to new pathways to target for the prevention and treatment of hemorrhagic and thrombotic disorders.

Appendix

The following is a list of institutional and study affiliations:

Atherosclerosis Risk In Communities Study: Divisions of Biostatistics (S.B., X.K.), Epidemiology and Community Health (J.S.P., A.R.F., W.T.), University of Minnesota, Minneapolis, Minn; DNA Laboratory (N.A.), Human Genetics Center and Institute of Molecular Medicine (E.B.), University of Texas Health Science Center at Houston, Houston, Tex.

British 1958 Birth Cohort: Division of Community Health Sciences (D.P.S.), St George’s, University of London, United Kingdom; Division of Cardiovascular and Medical Sciences (A.R.), University of Glasgow, Royal Infirmary, Glasgow, United Kingdom; ALSPAC Laboratory (W.L.M.), University of Bristol, United Kingdom.

Cardiovascular Health Study: Departments of Epidemiology (N.L., B.M.P.), Medicine (J.C.B., K.L.W., B.M.P.), Biostatistics (T.L., B.M.), and Health Services (B.M.P.), University of Washington, and Group Health Research Institute (B.M.P.), Group Health Cooperative, Seattle, Wash; Departments of Pathology (M.C., E.G.B.) and Medicine (M.C.), University of Vermont, Burlington, Vt; Medical Genetics Institute, Cedars-Sinai Medical Center (J.I.R.), Los Angeles, Calif.

Framingham Heart Study: National Heart, Lung, and Blood Institute’s (NHLBI’s) Framingham Heart Study (M.H.C., Q.Y., A.D.J., S.K., M.G.L., C.J.O.), Framingham, Mass; Departments of Biostatistics (Q.Y.), Neurology (M.H.C.), and Mathematics (M.G.L.), Boston University, Boston, Mass; Division of Intramural Research (A.D.J., C.J.O.), NHLBI, Bethesda, Md; General Medicine Division (J.B.M.), Massachusetts General Hospital, Harvard Medical School, Boston, Mass; Cardiology Division (S.K., C.J.O.), Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Mass; Royal North Shore Hospital (G.H.T.), Sydney, Australia.

Rotterdam Study: Departments of Epidemiology (A.D., C.V.D., A.H., A.U., J.C.M.), Hematology (M.D.M., F.L.), and Internal Medicine (A.U.), Erasmus Medical Center, Rotterdam, Netherlands; Member of the Netherlands Consortium on Healthy Aging (A.D., J.C.M.).
Twins UK: Wellcome Trust Sanger Institute (N.S.), Hinxton, United Kingdom; Department of Twin Research & Genetic Epidemiology (N.S., F.M.K.W., T.D.S.), King’s College London, London, United Kingdom.

VIS: MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine (C.H., V.V., A.F.W.), Edinburgh, Scotland, United Kingdom; Croatian Centre for Global Health (I.R.), University of Split Medical School, Split, Croatia; Centre for Population Health Sciences (I.R., H.C., J.F.W.), University of Edinburgh Medical School, Edinburgh, Scotland, United Kingdom.

ORCADES: Centre for Population Health Sciences (I.R., H.C., J.F.W.), University of Edinburgh Medical School, Edinburgh, Scotland, United Kingdom; MRC Human Genetics Unit (C.H., V.V., A.F.W.), Institute of Genetics and Molecular Medicine, Edinburgh, Scotland, United Kingdom; Croatian Centre for Global Health (I.R.), University of Split Medical School, Split, Croatia.

PROCARDIS: Atherosclerosis Research Unit (M.S.L., A.M., A.S., A.H.), Department of Medicine, Karolinska Institutet, Stockholm, Sweden; Department of Cardiovascular Medicine (J.F.P.), University of Oxford, Oxford, United Kingdom.

KORA: Institute of Epidemiology (A.P.), Helmholtz Zentrum München–German Research Center for Environment and Health, Neuherberg; Internal Medicine II–Cardiology (W.K.), University of Munich, Munich, Germany.

Acknowledgments

The authors acknowledge the essential role of the CHARGE (Co-horts for Heart and Aging Research in Genome Epidemiology) Consortium in development and support of this report. CHARGE members include the NHLBI’s Atherosclerosis Research Consortium; National Institute on Aging’s Iceland Age, Gene/Environment Susceptibility (AGES) Study; NHLBI’s CHS and FHS; and the Netherlands’ RS. The authors also acknowledge the thousands of study participants who volunteered their time to help advance science and the scores of research staff and scientists who have made this research possible. Furthermore, ARIC would like to thank the University of Minnesota Supercomputing Institute for use of the blade supercomputers; TUK acknowledges the essential contribution of Peter Grant and Angela Carter from the Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, United Kingdom, for measurements of clotting factors phenotypes; and VIS acknowledges the invaluable contributions of the recruitment team (including those from the Institute of Anthropological Research in Zagreb) in Vis, the administrative teams in Croatia and Edinburgh, and the people of Vis. Genotyping was performed at the Wellcome Trust Clinical Research Facility in Edinburgh, United Kingdom.

Sources of Funding

ARIC is supported by NHLBI contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022 and grants R01-HL-087641, R01-HL-59367, and R01-HL-086694; National Human Genome Research Institute contract U01-HG-004402; and National Institutes of Health (NIH) contract HHSN268200625226C. The infrastructure was supported in part by grant No. UL1-RR-025005, a component of the NIH Roadmap for Medical Research. We acknowledge use of the National Institute on Aging’s Iceland Age, Gene/Environment Susceptibility (AGES) Study; NHLBI’s CHS and FHS; and the Netherlands’ RS. The authors also acknowledge the thousands of study participants who volunteered their time to help advance science and the scores of research staff and scientists who have made this research possible. Furthermore, ARIC would like to thank the University of Minnesota Supercomputing Institute for use of the blade supercomputers; TUK acknowledges the essential contribution of Peter Grant and Angela Carter from the Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, United Kingdom, for measurements of clotting factors phenotypes; and VIS acknowledges the invaluable contributions of the recruitment team (including those from the Institute of Anthropological Research in Zagreb) in Vis, the administrative teams in Croatia and Edinburgh, and the people of Vis. Genotyping was performed at the Wellcome Trust Clinical Research Facility in Edinburgh, United Kingdom.

The authors acknowledge the essential role of the CHARGE (Co-horts for Heart and Aging Research in Genome Epidemiology) Consortium in development and support of this report. CHARGE members include the NHLBI’s Atherosclerosis Research Consortium; National Institute on Aging’s Iceland Age, Gene/Environment Susceptibility (AGES) Study; NHLBI’s CHS and FHS; and the Netherlands’ RS. The authors also acknowledge the thousands of study participants who volunteered their time to help advance science and the scores of research staff and scientists who have made this research possible. Furthermore, ARIC would like to thank the University of Minnesota Supercomputing Institute for use of the blade supercomputers; TUK acknowledges the essential contribution of Peter Grant and Angela Carter from the Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, United Kingdom, for measurements of clotting factors phenotypes; and VIS acknowledges the invaluable contributions of the recruitment team (including those from the Institute of Anthropological Research in Zagreb) in Vis, the administrative teams in Croatia and Edinburgh, and the people of Vis. Genotyping was performed at the Wellcome Trust Clinical Research Facility in Edinburgh, United Kingdom.

Disclosure

Dr Meigs reports receiving grant support from Sanofi-Aventis and GlaxoSmithKline and serving as a consultant/advisor to Interleukin Genetics and Eli Lilly. The remaining authors report no conflicts.
References


16. Smith et al. Genome-Wide Associations With FVII, FVIII, and vWF 1391


---

**CLINICAL PERSPECTIVE**

Elevated circulating levels of coagulation factor VIII (FVIII) and von Willebrand factor (vWF) are risk factors for venous thrombosis, but the data supporting an association of coagulation factor VII (FVII) levels with arterial thrombosis are less consistent. Hemorrhagic complications are associated with deficiency of FVII and vWF (von Willebrand disease), as well as X-linked deficiency in FVIII (hemophilia A). To date, our understanding of genetic variation influencing plasma levels has been focused primarily on variation in the genes encoding each protein product (F7, F8, and VWF, respectively). Using data from 23 608 adults drawn from community populations, we investigated genome-wide associations between common genetic variation and plasma levels of FVII, FVIII, and vWF. For FVII, we identified 5 loci on 5 chromosomes that exceeded genome-wide significance. All loci were within or near genes, including 4 new candidate genes and F7. For vWF, we identified 8 loci on 6 chromosomes. All loci were within genes, including 6 new candidate genes, as well as ABO and VWF. For FVIII, 5 loci were identified and overlapped vWF findings. We replicated 9 of the 10 new findings in independent samples. In summary, new genetic associations were discovered in biological pathways not previously associated with circulating levels of these factors, including proteins implicated in uptake and intracellular transport of the factors. These findings may point to novel prevention and treatment targets of hemostasis disorders.
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; originally published online March 15, 2010; doi: 10.1161/CIRCULATIONAHA.109.869156

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2010 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/121/12/1382

Data Supplement (unedited) at:
http://circ.ahajournals.org/content/suppl/2010/03/11/CIRCULATIONAHA.109.869156.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/
SUPPLEMENTAL MATERIAL
<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Cohort</th>
<th>Assay Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FVII level</td>
<td>ARIC</td>
<td>Factor VII activity was measured by determining the ability of the tested sample to correct the clotting time of human factor VII–deficient plasma obtained from George King Biomedical Inc. (Overland Park, KS). Plasma activity level was determined by relating the clotting time to a calibration curve constructed with each assay. Values were expressed as percentage of the standard. The reliability coefficient obtained from repeated testing of 39 individuals over several weeks was 0.78, and the method CV was 7.3%. No reference range available.</td>
</tr>
<tr>
<td></td>
<td>CHS</td>
<td>Factor VII activity was determined by using factor VII-deficient plasma (Baxter-Dade, Bedford MA, USA) and Thromborel S (Behring Diagnostics, Marburg, Germany) human-placenta derived thromboplastin. Activity measured using Coag-A-Mate X2 (Organon-Teknika, Durham, NC). Values were expressed as percentage of the standard. The intra-assay CV was 5.6% and inter-assay CV was 6.2%. No reference range available.</td>
</tr>
<tr>
<td></td>
<td>FHS</td>
<td>Factor VII antigen levels were determined by ELISA (Diagnostica Stago). Values were expressed as percentage of the standard. The intra-assay CV was 3.0%. No reference range available.</td>
</tr>
<tr>
<td></td>
<td>RS</td>
<td>Factor VII activity was measured with a one-stage clotting assay by using human thromboplastin (Tromborel S, Siemens) and factor VII-deficient plasma (Ortho Diagnostic System). 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 1.7%, the inter-assay CV was 3.7%, and reference range was 0.60-1.40 U/ml.</td>
</tr>
<tr>
<td></td>
<td>Twins UK</td>
<td>For Twins UK-1, factor VII activity was measured using a functional clotting assay with factor VII–deficient plasma on an ACL300 research coagulometer (Instrumentation Laboratory, Milan, Italy). For Twins UK-2, factor VII antigen level was determined by ELISA.</td>
</tr>
<tr>
<td></td>
<td>PROCARDIS</td>
<td>Factor VII antigen was measured with a sandwich ELISA using a matched-pair antibody set from Affinity Biologicals Inc (Ancaster, ON, Canada). The plasma concentration was expressed in U/mL by relating the absorbance of the sample to that of a calibration curve constructed with a standardized coagulation reference plasma (Technoclone GmbH, Vienna, Austria). Normal control plasma for coagulation analyses (Global Hemostasis Institute, Linköping, Sweden) was used for assessment of the variability of the assay, rendering an intra-assay CV of 8% (n=10) and an inter-assay CV of 10% (n=41).</td>
</tr>
<tr>
<td>FVIII level</td>
<td>ARIC</td>
<td>FVIII activity was measured by assessing clotting time after adding factor VIII deficient plasma obtained from George King Biomedical Inc. (Overland Park, KS). Plasma activity level was determined by relating the clotting time to a calibration curve constructed with each assay. Values were expressed as percentage of the standard. The reliability coefficient obtained from repeated testing of 39 individuals over several weeks was 0.86, and the method CV was 5.9%. No reference range available.</td>
</tr>
<tr>
<td></td>
<td>CHS</td>
<td>FVIII activity determined by using factor VIII-deficient plasma (Organon-Teknika) and partial thromboplastin (Organon-Teknika). 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.</td>
</tr>
</tbody>
</table>
|           | RS      | Factor VIII activity was measured with a one-stage clotting assay by using a mixture of micronized silica and
<table>
<thead>
<tr>
<th>vWF level</th>
<th>Experiment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twins UK</td>
<td>Factor VIII antigen was measured using ELISA (Immuno, Vienna, Austria) methods.</td>
<td></td>
</tr>
<tr>
<td>ARIC</td>
<td>Von Willebrand factor antigen was determined by an enzyme-linked immunosorbent assay (ELISA) kit from American Bioproducts Co (Parsippany, NJ). The reliability coefficient obtained from repeated 39 testing of individuals over several weeks was 0.68, and the method CV was 18.5%. No reference range available.</td>
<td></td>
</tr>
<tr>
<td>B58C</td>
<td>Plasma von Willebrand factor antigen levels were measured using an in-house enzyme linked immunosorbent assay (ELISA), employing rabbit anti-human polyclonal antibodies obtained from DAKO plc, Copenhagen, Denmark. The standard curve was constructed using the 9th British standard for Blood Cogulation Factors from 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.</td>
<td></td>
</tr>
<tr>
<td>VIS</td>
<td>Same as B58C</td>
<td></td>
</tr>
<tr>
<td>ORCADES</td>
<td>Same as B58C</td>
<td></td>
</tr>
<tr>
<td>FHS</td>
<td>Von Willebrand factor was assessed using enzyme-linked immunosorbent assays. In our laboratory, the intra-assay coefficient of variation was 8.8%. No reference range available.</td>
<td></td>
</tr>
<tr>
<td>RS</td>
<td>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.</td>
<td></td>
</tr>
<tr>
<td>Twins UK</td>
<td>Von Willebrand factor was measured using an in-house sandwich ELISA.</td>
<td></td>
</tr>
</tbody>
</table>
### Supplemental Table S2. Genotyping and imputation methods for autosomal chromosomes by study

<table>
<thead>
<tr>
<th></th>
<th>ARIC</th>
<th>B58C</th>
<th>CHS</th>
<th>FHS</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Methods</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platform</td>
<td>Affymetrix</td>
<td>Affymetrix</td>
<td>Illumina</td>
<td>Affymetrix</td>
<td>Illumina</td>
</tr>
<tr>
<td>Chip</td>
<td>6.0</td>
<td>500K</td>
<td>370 CNV</td>
<td>500K + 50K</td>
<td>550 v3</td>
</tr>
<tr>
<td>SNPs investigated</td>
<td>906,600</td>
<td>490,032</td>
<td>306,655</td>
<td>490,700 (500K)</td>
<td>561,466</td>
</tr>
<tr>
<td>SNP exclusion criteria*</td>
<td>≤0.90</td>
<td>1.00</td>
<td>≤0.97</td>
<td>≤0.97</td>
<td>≤0.98</td>
</tr>
<tr>
<td>HWE p-value</td>
<td>&lt;1.0x10^-6</td>
<td>≤5.0x10^-7</td>
<td>&lt;1.0x10^-5</td>
<td>&lt;1 x 10E-6</td>
<td>&lt;1.0x10^-6</td>
</tr>
<tr>
<td>Variants included for imputation</td>
<td>602,642</td>
<td>490,032</td>
<td>291,322</td>
<td>343,361 (500K)</td>
<td>512,349</td>
</tr>
<tr>
<td>Percent of variants included</td>
<td>66%</td>
<td>100%</td>
<td>95%</td>
<td>70% (500k)</td>
<td>91.3%</td>
</tr>
<tr>
<td>Imputation software</td>
<td>MACH</td>
<td>IMPUTE</td>
<td>BIMBAM</td>
<td>MACH</td>
<td>MACH</td>
</tr>
<tr>
<td>Imputation software version</td>
<td>1.0.16</td>
<td>0.1.3</td>
<td>0.99</td>
<td>1.0.15</td>
<td>1.0.15</td>
</tr>
<tr>
<td>Genome build</td>
<td>35</td>
<td>35</td>
<td>36</td>
<td>36.2</td>
<td>36</td>
</tr>
<tr>
<td>Total number of SNPs</td>
<td>2,516,203</td>
<td>2,236,936</td>
<td>2,543,887</td>
<td>2,543,887</td>
<td>2,543,887</td>
</tr>
</tbody>
</table>

**Population**

<table>
<thead>
<tr>
<th></th>
<th>ARIC</th>
<th>B58C</th>
<th>CHS</th>
<th>FHS</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of subjects in cohort(s)</td>
<td>15,792</td>
<td>1,461</td>
<td>5,888</td>
<td>16,419</td>
<td>7,983</td>
</tr>
<tr>
<td>Participants eligible for scans**</td>
<td>11,433†</td>
<td>1,461</td>
<td>3,980†</td>
<td>9,274</td>
<td>6,449</td>
</tr>
<tr>
<td>Percent eligible</td>
<td>72%</td>
<td>100%</td>
<td>68%</td>
<td>56%</td>
<td>81%</td>
</tr>
<tr>
<td>Participant with successful scans***</td>
<td>10,651‡</td>
<td>1,461</td>
<td>3,868‡</td>
<td>8,756</td>
<td>5,974</td>
</tr>
<tr>
<td>Percent successful</td>
<td>93%</td>
<td>100%</td>
<td>97%</td>
<td>94%</td>
<td>93%</td>
</tr>
</tbody>
</table>

HWE = Hardy-Weinberg equilibrium

*For CHS, >1 duplicate error or Mendelian inconsistency (for reference CEPH trios), heterozygote frequency = 0, SNP not found in dbSNP.

**ARIC number indicates the number of participants with available scans at the time of the analysis.

***Sample successfully genotyped on platform with within-participant variant call rate >0.95 and after implementation of all quality controls.

†European ancestry participants in ARIC (8861) and in CHS (3397)

‡European ancestry in ARIC (8127) and in CHS (3295)
Supplemental Table 3a. Replication cohort specific results for the top loci associated with factor VII activity and antigen

<table>
<thead>
<tr>
<th>CHR</th>
<th>SNP</th>
<th>Variant</th>
<th>ARIC (FVII:c)</th>
<th>Twins UK-1 (FVII:ag)</th>
<th>Twins UK-2 (FVII:c)</th>
<th>PROCARDIS (FVII:ag)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>β</td>
<td>P-value</td>
<td>β</td>
<td>P-value</td>
</tr>
<tr>
<td>2</td>
<td>rs1260326</td>
<td>C→T</td>
<td>1.1</td>
<td>2.6x10^-1</td>
<td>5.3</td>
<td>6.2x10^-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>rs1126670</td>
<td>T→G</td>
<td>-2.5</td>
<td>1.8x10^-2</td>
<td>1.2</td>
<td>4.6x10^-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>rs11230180</td>
<td>G→T</td>
<td>2.3</td>
<td>2.1x10^-2</td>
<td>1.5</td>
<td>3.4x10^-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>rs867186</td>
<td>A→G</td>
<td>7.1</td>
<td>2.0x10^-5</td>
<td>7.0</td>
<td>7.1x10^-3</td>
</tr>
</tbody>
</table>

CHR = chromosome; P = p-value; β coefficient represents change (% of activity or antigen) associated with 1-unit change in allele dosage; NA = parameter estimate not available; SNP = single nucleotide polymorphism.

Supplemental Table 3b. Replication cohort specific results for the top loci associated with von Willebrand factor antigen.

<table>
<thead>
<tr>
<th>CHR</th>
<th>SNP</th>
<th>Variant</th>
<th>ARIC (1 and 2)</th>
<th>B58C</th>
<th>Twins UK (1 and 2)</th>
<th>VIS</th>
<th>Orcade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>β</td>
<td>P-value</td>
<td>β</td>
<td>P-value</td>
<td>β</td>
</tr>
<tr>
<td>6</td>
<td>rs9390459</td>
<td>G→A</td>
<td>-3.3</td>
<td>3.3x10^-5</td>
<td>-3.4</td>
<td>3.5x10^-4</td>
<td>-3.3</td>
</tr>
<tr>
<td>8</td>
<td>rs2726953</td>
<td>C→T</td>
<td>5.5</td>
<td>1.4x10^-3</td>
<td>2.7</td>
<td>2.9x10^-2</td>
<td>2.1</td>
</tr>
<tr>
<td>12</td>
<td>rs4981022</td>
<td>T→C</td>
<td>-5.0</td>
<td>6.0x10^-3</td>
<td>-3.6</td>
<td>2.1x10^-2</td>
<td>0.1</td>
</tr>
<tr>
<td>12</td>
<td>rs7978987</td>
<td>G→A</td>
<td>3.4</td>
<td>4.4x10^-2</td>
<td>0.0</td>
<td>9.9x10^-1</td>
<td>NA</td>
</tr>
<tr>
<td>14</td>
<td>rs10133762</td>
<td>G→T</td>
<td>4.0</td>
<td>1.1x10^-2</td>
<td>1.8</td>
<td>1.2x10^-1</td>
<td>0.2</td>
</tr>
<tr>
<td>19</td>
<td>rs868875</td>
<td>A→G</td>
<td>-0.5</td>
<td>7.9x10^-1</td>
<td>-2.8</td>
<td>2.3x10^-2</td>
<td>-25.6</td>
</tr>
</tbody>
</table>

CHR = chromosome; P = p-value; β coefficient represents change (% of activity or antigen) associated with 1-unit change in allele dosage; NA = parameter estimate not available; SNP = single nucleotide polymorphism.
Supplemental Table S4. Minimum detectable percent changes in hemostasis factor with at least 80% power according to minor allele frequency

<table>
<thead>
<tr>
<th>MAF</th>
<th>FVII mean = 100%, SD = 30% (n=15400)</th>
<th>FVIII mean = 100%, SD = 35% (n=15200)</th>
<th>vWF mean = 100%, SD = 45% (n=17600)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>5.0%</td>
<td>6.0%</td>
<td>7.0%</td>
</tr>
<tr>
<td>15%</td>
<td>3.0%</td>
<td>4.0%</td>
<td>4.5%</td>
</tr>
<tr>
<td>25%</td>
<td>2.5%</td>
<td>3.0%</td>
<td>3.5%</td>
</tr>
<tr>
<td>35%</td>
<td>2.5%</td>
<td>3.0%</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

MAF = minor allele frequency;
Supplemental Figures S1-S13. Regional plots for top marker loci. The top SNP is presented as a large diamond in red font and neighboring variants are presented in different colors based on linkage disequilibrium based on HapMap Caucasian data: red: $1 \geq r^2 > 0.8$; orange: $0.8 \geq r^2 > 0.6$; yellow: $0.6 \geq r^2 > 0.3$; green: $0.3 \geq r^2 > 0.1$; blue: $0.1 \geq r^2 > 0.05$; light blue: $0.05 \geq r^2 > 0.0$. The left y-axis is the p-value on the $-\log_{10} p$-value scale and the gray line marks the threshold of genome-wide significance, $5 \times 10^{-8}$. The right y-axis is the recombination rate in centimorgans per mega base and the light blue line maps the rate across the region. Regional genes and their direction of transcription are depicted with green arrows.

S1. rs1260326
Regional plot for rs1126670

rs1126670

P = 3.64 \times 10^{-12}

Chromosome 4 location (Kb)

Recombination rate (cM/Mb)
Regional plot for rs11230180

rs11230180

\( P = 1.95 \times 10^{-10} \)

Chromosome 11 location (Kb)
S4. rs488703

Regional plot for rs488703

rs488703

P = 9 \times 10^{-259}

- \log_{10}(P)

Recombination rate (cM/Mb)

Chromosome 13 location (Kb)
Regional plot for rs867186

rs867186
P = 5.66 \times 10^{-37}

Chromosome 20 location (Kb)

- log_{10}(P)

Recombination rate (cMMb)
Regional plot for rs9390459

rs9390459

$P = 1.16 \times 10^{-22}$

Chromosome 6 location (Kb)
Regional plot for rs2726953

rs2726953

\[ P = 1.27 \times 10^{-15} \]

Chromosome 8 location (Kb)

Recombination rate (cMM/kb)
Regional plot for rs687621

rs687621

$P < 5 \times 10^{-324}$
Regional plot for rs1063857

- log_{10}(P) vs. Recombination rate (cM/Mb)

**rs1063857**

- **P = 1.66 \times 10^{-32}**

Chromosome 12 location (Kb):

- ANO2
- WWF
- CD8
Regional plot for rs7978987

rs7978987

- log₁₀(P) vs Chromosome 12 location (Kb)

Recombination rate (cM/Mb)
Regional plot for rs10133762

rs10133762  \quad P = 2.27 \times 10^{-10}

- \log_{10}(P)

Recombination rate (cM/Mb)

Chromosome 14 location (Kb)

CATSPERB

TC2N

FBLN5

TRIP11
Regional plot for rs868875

rs868875

\[ P = 1.3 \times 10^{-9} \]