ARTICLE

Meta-analysis of 65,734 Individuals Identifies *TSPAN15* and *SLC44A2* as Two Susceptibility Loci for Venous Thromboembolism

Marine Germain,[1,2,3] Daniel I. Chasman,[4] Hugoline de Haan,[5] Weihong Tang,[6] Sara Lindström,[7] Lu-Chen Weng,[6] Mariza de Andrade,[8] Marieke C.H. de Visser,[9] Kerri L. Wiggins,[10] Pierre Suchon,[11,12,13] Noémie Saut,[11,12,13] David M. Smadja,[14,15,16] Grégoire Le Gal,[17,18] Astrid van Hylckama Vlieg,[5] Antonio Di Narzo,[19] Ke Hao,[19] Christopher P. Nelson,[20,21] Ares Rocanin-Arjo,[1,2,3] Lasse Folkersen,[22] Ramin Monajemi,[23] Lynda M. Rose,[24] Jennifer A. Brody,[25] Eline Slagboom,[26] Dylan Aïssi,[1,2,3] France Gagnon,[27] Jean-Francois Deleuze,[28] Panos Deloukas,[29,30] Christophe Tzourio,[31] Jean-Francois Dartigues,[31] Claudine Berr,[32] Kent D. Taylor,[33] Mete Civelek,[34] Per Eriksson,[35] Cardiogenics Consortium, Bruce M. Psaty,[25,36] Jeanine Houwing-Duitermaat,[23] Alison H. Goodall,[20,21] François Cambien,[1,2,3] Peter Kraft,[7] Philippe Amouyel,[37,38] Nilesh J. Samani,[20,21] Saonli Basu,[39] Paul M. Ridker,[4] Frits R. Rosendaal,[5] Christopher Kabrhel,[40] Aaron R. Folsom,[6] John Heit,[41] Pieter H. Reitsma,[9] David-Alexandre Trégouët,[1,2,3] Nicholas L. Smith,[10,36,42,*] and Pierre-Emmanuel Morange[11,12,13,*]

Venous thromboembolism (VTE), the third leading cause of cardiovascular mortality, is a complex thrombotic disorder with environmental and genetic determinants. Although several genetic variants have been found associated with VTE, they explain a minor proportion of VTE risk in cases. We undertook a meta-analysis of genome-wide association studies (GWASs) to identify additional VTE susceptibility genes. Twelve GWASs totaling 7,507 VTE case subjects and 52,632 control subjects formed our discovery stage where 6,751,884 SNPs were tested for association with VTE. Nine loci reached the genome-wide significance level of 5×10^{-8} including six already known to associate with VTE (*ABO*, *F2*, *F5*, *F11*, *FGG*, and *PROCR*) and three unsuspected loci. SNPs mapping to these latter were selected for replication in three independent case-control studies totaling 3,009 VTE-affected individuals and 2,586 control subjects. This strategy led to the identification and replication of two VTE-associated loci, *TSPAN15* and *SLC44A2*, with lead risk alleles associated with odds ratio for disease of 1.31 ($p = 1.67 \times 10^{-16}$) and 1.21 ($p = 2.75 \times 10^{-15}$), respectively. The lead SNP at the *TSPAN15* locus is the intronic rs78707713 and the lead *SLC44A2* SNP is the non-synonymous rs2288904 previously shown to associate with

[1]Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR_S) 1166, 75013 Paris, France; [2]Sorbonne Universités, Université Pierre et Marie Curie (UPMC Univ Paris 06), UMR_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, 75013 Paris, France; [3]Institute for Cardiometabolism and Nutrition (ICAN), 75013 Paris, France; [4]Division of Preventive Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02215, USA; [5]Department of Thrombosis and Hemostasis, Department of Clinical Epidemiology, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands; [6]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN 55454, USA; [7]Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA; [8]Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN 55905, USA; [9]Einthoven Laboratory for Experimental Vascular Medicine, Department of Thrombosis and Hemostasis, Leiden University Medical Center, 2300 RC Leiden, the Netherlands; [10]Department of Epidemiology, University of Washington, Seattle, WA 98195, USA; [11]Laboratory of Haematology, La Timone Hospital, 13385 Marseille, France; [12]INSERM, UMR_S 1062, Nutrition Obesity and Risk of Thrombosis, 13385 Marseille, France; [13]Nutrition Obesity and Risk of Thrombosis, Aix-Marseille University, UMR_S 1062, 13385 Marseille, France; [14]Université Paris Descartes, Sorbonne Paris Cité, 75006 Paris, France; [15]AP-HP, Hopital Européen Georges Pompidou, Service d'Hématologie Biologique, 75015 Paris, France; [16]Faculté de Pharmacie, INSERM, UMR_S 1140, 75006 Paris, France; [17]Université de Brest, EA3878 and CIC1412, 29238 Brest, France; [18]Ottawa Hospital Research Institute at the University of Ottawa, Ottawa, ON K1Y 4E9, Canada; [19]Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY 10029, USA; [20]Department of Cardiovascular Sciences, University of Leicester, LE1 7RH Leicester, UK; [21]National Institute for Health Research (NIHR) Leicester Cardiovascular Biomedical Research Unit, Leicester LE3 9QP, UK; [22]Department of PharmacoGenetics, Novo Nordisk Park 9.1.21, 2400 Copenhagen, Denmark; [23]Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, 2300 RC Leiden, the Netherlands; [24]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02215, USA; [25]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, WA 98195-5852, USA; [26]Department of Molecular Epidemiology, Leiden University Medical Center, 2300 RC Leiden, the Netherlands; [27]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, ON M5T 3M7, Canada; [28]Commissariat à l'Energie Atomique/Direction des Sciences du Vivant/Institut de Génomique, Centre National de Génotypage, 91057 Evry, France; [29]William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 4NS, UK; [30]Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah 21589, Saudi Arabia; [31]Inserm Research Center U897, University of Bordeaux, 33000 Bordeaux, France; [32]Inserm Research Unit U1061, University of Montpellier I, 34000 Montpellier, France; [33]Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor-UCLA Medical Center, Torrence, CA 90502, USA; [34]Department of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA; [35]Atherosclerosis Research Unit, Center for Molecular Medicine, Department of Medicine, Karolinska Institutet, 171 77 Stockholm, Sweden; [36]Group Health Research Institute, Group Health Cooperative, Seattle, WA 98101, USA; [37]Institut Pasteur de Lille, Université de Lille Nord de France, INSERM UMR_S 744, 59000 Lille, France; [38]Centre Hospitalier Régional Universitaire de Lille, 59000 Lille, France; [39]Division of Biostatistics, University of Minnesota, Minneapolis, MN 55455, USA; [40]Department of Emergency Medicine, Massachusetts General Hospital, Channing Network Medicine, Harvard Medical School, Boston, MA 2114, USA; [41]Division of Cardiovascular Diseases, Mayo Clinic, Rochester, MN 55905, USA; [42]Seattle Epidemiologic Research and Information Center, VA Office of Research and Development, Seattle, WA 98108, USA

*Correspondence: nlsmith@u.washington.edu (N.L.S.), pierre.morange@ap-hm.fr (P.-E.M.)

http://dx.doi.org/10.1016/j.ajhg.2015.01.019.

transfusion-related acute lung injury. We further showed that these two variants did not associate with known hemostatic plasma markers. *TSPAN15* and *SLC44A2* do not belong to conventional pathways for thrombosis and have not been associated to other cardiovascular diseases nor related quantitative biomarkers. Our findings uncovered unexpected actors of VTE etiology and pave the way for novel mechanistic concepts of VTE pathophysiology.

Introduction

Venous thromboembolism (VTE [MIM 188050]) is a common multicausal thrombotic disease with an annual incidence of 1 per 1,000. It includes two main clinical manifestations: deep vein thrombosis (DVT) and pulmonary embolism (PE). The latter is associated with a 1-year mortality of 20%, making VTE the third leading cause of cardiovascular death in industrialized countries.[1] Moreover, among survivors, 25%–50% will have lasting debilitating health problems such as post-thrombotic syndrome, severely hampering mobility and quality of life.

Factors contributing to VTE include endothelial injury or activation, reduced blood flow, and hypercoagulability of the blood, the so-called Virchow triad.[2] Venous thromboembolism has a strong genetic basis that is characterized by an underlying heritability estimate of 50% and a risk of developing the disease in an individual with an affected sib 2.5 higher than for the general population.[3,4] But, like other complex phenotypes, most genetic contributors have not been elucidated because the proportion of heritability explained by replicated variants has been small.[5,6]

There are seven well-established genetic risk factors for VTE, all responsible for inherited hypercoagulable states. The first three are heterozygous deficiencies of the natural coagulation inhibitors (antithrombin, protein C, and protein S). These deficiencies are relatively rare, affecting <1% of the general population, and they increase VTE risk by approximately ten. The other four, Factor V (FV [MIM 612309]) Leiden, prothrombin (MIM 176930) G20210A, fibrinogen γ′ (FGG) (MIM 134850) rs2066865, and blood group non-O, are more frequent with prevalence in European-descent individuals around 5% for the former two and ~25% for the latter two. The increase in VTE risk is about 3-fold for the FV Leiden (RefSeq accession number NP_000121.2, p.Arg534Gln [c.1601G>A]) and prothrombin G20210A (RefSeq NM_000506.3, c.*97G>A) mutations, 2-fold for non-O blood group, and 1.5-fold for *FGG* rs2066865.[7]

The genome-wide association strategy is a powerful method to identify common SNPs associated with a complex disorder without a pre-specified hypothesis. Although previous genome-wide association studies (GWASs) have been reported for VTE, none has included more than 1,961 case subjects[5,6] and none has yielded new genetic loci. In this article, we report the largest investigation to date of the influence of common genetic variations on VTE risk by meta-analyzing GWAS findings from 12 studies.

Subjects and Methods

Study Design

We report on a three-stage investigation of common genetic predictors of VTE. A discovery phase included 7,507 VTE case subjects and 52,632 control subjects from 12 studies and a replication phase included 3,009 case subjects and 2,586 control subjects from three independent studies. In addition, confirmed discoveries were then examined for association with quantitative biomarkers of VTE risk, gene expression in various tissues and cell types, as well as with whole blood DNA methylation levels.

Participants in Discovery and Replication

For the discovery phase, participants were European-ancestry adults in two French case-control studies, two Dutch case-control studies, and four cohort and four case-control studies from the United States. Details of each study have been previously published.[8–19] Three other French case-control studies for VTE were used for the replication stage.[11,20] In all studies, VTE (PE or DVT) was objectively diagnosed by physicians using different techniques including compression venous duplex ultrasonography, computed tomography, Doppler ultrasound, impedance plethysmography, magnetic resonance, venography, pulmonary angiography, and ventilation/perfusion lung scan. VTE events related to cancer, autoimmune disorders, or natural anticoagulant inhibitor deficiencies (protein C, protein S, antithrombin) were excluded in most studies. A detailed description of the design and the clinical characteristics of all VTE studies analyzed in this work is presented in Table S1. All participating studies were approved by their respective institutional review board and informed consent was obtained from studied individuals.

Genotyping and Imputation

Within each discovery study, DNA samples were genotyped with high-density SNP arrays and were imputed for SNPs available in the 1000 Genomes reference dataset. Summary descriptions of genotyping technologies, quality-control procedures, and imputation methods used for the discovery cohorts are shown in Table S1. In the replication studies, genotyping of the selected SNPs was performed by allele-specific PCR.

Association Analyses and Meta-Analysis for Discovery

Association analyses of imputed SNPs with VTE risk were performed separately in each study by using logistic or Cox-proportional regression analyses adjusted for study-specific covariates (Table S1). All SNPs with acceptable imputation quality ($r^2 > 0.3$)[21] in all 12 discovery studies and with estimated minor allele frequency greater than 0.005 were entered into a meta-analysis. For the meta-analysis, a fixed-effects model based on the inverse-variance weighting was employed as implemented in the METAL software.[22] A statistical threshold of 5×10^{-8} controlling for the number of independent tests across the genome was applied to declare genome-wide significance.[21,23,24] Heterogeneity of the SNP associations across studies was tested with the Cochran's Q statistic and its magnitude expressed by the I^2 index.

Table 1. Main Findings at Loci Demonstrating Genome-wide Significant Association with VTE in the Discovery Cohorts

				Discovery Stage (7,507 Case Subjects and 52,632 Control Subjects)				Replication Stage (3,009 Case Subjects and 2,586 Control Subjects)			Combined
Lead SNP	Chr.	Gene	Description	Risk Allele	Risk Allele Frequency	Allelic OR[a]	p	Risk Allele Frequency	Allelic OR	p	p[b]
Known Loci											
rs6025	1	*F5*	missense	T	0.033	3.25 (2.91–3.64)	1.10×10^{-96}	–	–	–	–
rs4524[c]	1	*F5*	missense	T	0.736	1.20 (1.14–1.26)	2.65×10^{-11}	–	–	–	–
rs2066865	4	*FGG*	3′ UTR	A	0.244	1.24 (1.18–1.31)	1.03×10^{-16}	–	–	–	–
rs4253417	4	*F11*	intronic	C	0.405	1.27 (1.22–1.34)	1.21×10^{-23}	–	–	–	–
rs529565	9	*ABO*	intronic	C	0.354	1.55 (1.48–1.63)	4.23×10^{-75}	–	–	–	–
rs1799963	11	*F2*	intronic	A	0.010	2.29 (1.75–2.99)	1.73×10^{-9}	–	–	–	–
rs6087685	20	*PROCR*	intronic	C	0.302	1.15 (1.10–1.21)	1.65×10^{-8}	–	–	–	–
New Loci											
rs4602861	8	*ZFPM2*	intronic	A	0.766	1.20 (1.13–1.27)	3.48×10^{-9}	0.714	1.02 (0.94–1.11)	0.631	5.04×10^{-7}
rs78707713	10	*TSPAN15*	intronic	T	0.878	1.28 (1.19–1.39)	5.74×10^{-11}	0.891	1.42 (1.24–1.62)	2.21×10^{-7}	1.67×10^{-16}
rs2288904[d]	19	*SLC44A2*	missense	G	0.785	1.19 (1.12–1.26)	1.07×10^{-9}	0.764	1.28 (1.16–1.40)	2.64×10^{-7}	2.75×10^{-15}

[a]Allelic odds ratio associated with the risk allele with its corresponding confidence interval. For ease of presentation, all confidence intervals shown in this table have been computed at $\alpha = 0.05$.
[b]The combined p value was derived from a fixed effect meta-analysis of the discovery and replication results.
[c]The rs4524 was associated with VTE risk independently of the rs6025 variant.
[d]The rs2288904 was not the lead SNP observed at the *SLC44A2* locus. However, because it is a non-synonymous polymorphism with strong evidence for functionality that is in strong LD ($r^2 \sim 1$) with the lead intronic rs2360742 ($p = 5.6 \times 10^{-10}$), it was the one taken forward for replication.

Association Analyses and Meta-Analysis for Replication

In the replication cohorts, association of tested SNPs with VTE risk was assessed by use of logistic regression under the assumption of additive allele effects, adjusting for age and sex. Results obtained in the replication cohorts were meta-analyzed via a fixed-effects model based on the inverse-variance weighting as implemented in the METAL software.[22] A statistical threshold of 0.05 divided by the number of replications performed was used to declare statistical replication. Heterogeneity of the SNP associations across studies and subgroups of individuals (e.g., PE versus DVT) was tested via the Cochran's Q statistic and its magnitude expressed by the I^2 index. For replicated SNPs, meta-analyses of all studies were performed to produce the most robust estimate of the effect size.

Conditional Analysis to Discover Independent Signals at Replicated Loci

To test for the presence of additional independent VTE-associated SNPs at each of the replicated loci, we re-analyzed in each discovery study a GWAS conditioning on the imputed allelic dose of each replicated SNP and meta-analyzed the results by the same strategy as the original GWAS analysis. Areas within 200 kb up- and downstream of the SNPs were examined.

Biologic Follow-up on Replicated Findings

The influence of replicated SNPs on the variability of hemostatic traits known to associate with VTE pathophysiology and measured in our available cohorts was assessed to learn more about biologic pathways. Investigated quantitative biomarkers were D-dimers, endogenous thrombin generation, plasma antigen or activity levels of fibrinogen, coagulation factors II, VII, VIII, IX, X, and XII, von Willebrand factor, antithrombin, protein C, protein S (total and free), protein Z, activated partial thromboplastin time, hemoglobin, and white blood cell and platelet counts. Associations of replicated SNPs with these quantitative hemostatic traits were investigated in five cohorts that were part of the discovery and replication stages by using linear regression analyses adjusted for age-, sex-, and cohort-specific covariates. The overall statistical evidence for association with a given phenotype across available cohorts was assessed by use of the Fisher combined test statistics to account for different measurement methods used across studies.

Replicated SNPs were examined for association with the expression of their structural genes via publicly available genome-wide gene expression data from multiple cell lines and tissues. Impact of replicated lead SNPs on DNA methylation levels from peripheral blood DNA was also investigated.

Results

After applying quality-control measures, 6,751,884 SNPs were tested for association with VTE in a total of 7,507 case subjects and 52,632 control subjects. The Manhattan and Q-Q plots of the meta-analysis of GWAS results are shown in Figures S1–S4. A total of 1,060 SNPs clustered into nine chromosomal regions and reached the genome-wide significance level of $p < 5 \times 10^{-8}$.

Table 2. Replication Findings for Three SNPs that Reached Genome-wide Significance in the Discovery GWAS

	SLC44A2 rs2288904			*TSPAN15* rs78707713			*ZFPM2* rs4602861		
	GG	GA	AA	TT	TC	CC	AA	AG	GG
MARTHA12									
Control subjects	468 (60%)	258 (33%)	58 (7%)	624 (80%)	146 (19%)	5 (1%)	409 (53%)	314 (40%)	56 (7%)
Case subjects	764 (64%)	391 (33%)	35 (3%)	1,025 (86%)	162 (13%)	7 (<1%)	629 (53%)	472 (39%)	92 (8%)
RAF[a]	0.762 versus 0.806			0.899 versus 0.926			0.727 versus 0.745		
Allelic OR[b]	1.302 (1.115–1.519); $p = 8.67 \times 10^{-4}$			1.428 (1.136–1.798); $p = 2.31 \times 10^{-3}$			1.002 (0.866–1.157); p = 0.986		
FARIVE									
Control subjects	339 (59%)	204 (35%)	32 (6%)	452 (79%)	114 (20%)	9 (1%)	284 (50%)	231 (40%)	60 (10%)
Case subjects	370 (64%)	189 (32%)	22 (4%)	478 (82%)	102 (18%)	1 (<1%)	289 (50%)	243 (42%)	48 (8%)
RAF	0.767 versus 0.799			0.885 versus 0.910			0.695 versus 0.708		
Allelic OR	1.208 (0.989–1.475); p = 0.064			1.317 (1.000–1.733); p = 0.050			1.064 (0.891–1.271); p = 0.495		
EDITH									
Control subjects	680 (58%)	422 (36%)	66 (6%)	921 (79%)	236 (20%)	14 (1%)	564 (49%)	487 (42%)	110 (9%)
Case subjects	739 (63%)	394 (34%)	30 (3%)	993 (85%)	172 (15%)	8 (<1%)	570 (49%)	478 (41%)	110 (10%)
RAF	0.763 versus 0.805			0.887 versus 0.920			0.695 versus 0.699		
Allelic OR	1.294 (1.121–1.492); $p = 4.32 \times 10^{-4}$			1.457 (1.197–1.776); $p = 1.76 \times 10^{-4}$			1.014 (0.896–1.148); p = 0.821		
Combined allelic OR[c]	1.277 (1.164–1.403); $p = 2.64 \times 10^{-7}$			1.416 (1.241–1.616); $p = 2.21 \times 10^{-7}$			1.021 (0.939–1.110); p = 0.631		

[a]Risk allele frequencies in control subjects and in case subjects, respectively.
[b]Allelic odds ratio for disease associated with the risk allele of the studied polymorphisms derived from a logistic regression analysis adjusted for age and sex.
[c]Combined adjusted allelic OR derived from a standard fixed-effect meta-analysis of the results observed in the three replication studies. There was no evidence for heterogeneity across the three replication studies for the association of rs2288904 ($I^2 = 0.388$, p = 0.823), rs78707713 ($I^2 = 0.364$, p = 0.833), or rs4602861 ($I^2 = 0.285$, p = 0.867) with the disease.

Among the nine loci that were genome-wide significant in the discovery scan, six were already known to be associated with VTE (*ABO* [MIM 110300], *F2*, *F5*, *F11* [MIM 264900], *FGG*, and *PROCR* [MIM 600646]) whereas three had not been previously reported to be associated with VTE: *TSPAN15* (MIM 613140), *SLC44A2* (MIM 606106), and *ZFPM2* (MIM 603693) (Table 1). At loci with SNPs known to be associated with VTE, we did not identify additional VTE associations with variants. Detailed descriptions of findings at the known loci are in Figures S1 and S3.

For the three unknown loci, the region-specific lead SNPs were all common intronic variants: rs78707713 in *TSPAN15* with risk allele frequency (RAF) 0.88 and associated odds ratio (OR) for VTE of 1.28 ($p = 5.74 \times 10^{-11}$), rs2360742 in *SLC44A2* with RAF = 0.78 and OR = 1.20 ($p = 5.59 \times 10^{-10}$), and rs4602861 in *ZFPM2* with RAF = 0.77 and OR = 1.20 ($p = 3.48 \times 10^{-9}$). These associations showed little heterogeneity across studies, with Cochran's $Q = 16.8$, $I^2 = 0.34$, p = 0.11 for *TSPAN15* rs78707713. Corresponding values were Q = 5.64, $I^2 = 0.00$, p = 0.90 for *SLC44A2* rs2360742 and Q = 15.2, $I^2 = 0.28$, p = 0.17 for *ZFPM2* rs4602861. Of note, the intronic *SLC44A2* rs2360742 was in complete association ($r^2 = 1$) with the non-synonymous rs2288904 (p.Arg154Gln [c.461G>A]) that ranked ninth (in terms of p value) at this locus ($p = 1.07 \times 10^{-9}$). No coding SNPs in *TSPAN15* or *ZFPM2* were in high linkage disequilibrium (LD) with their lead SNPs.

We attempted replication of *TSPAN15* rs78707713, *SLC44A2* rs2288904, and *ZFPM2* rs4602861 SNPs in the meta-analysis of three independent studies totaling 3,009 VTE-affected individuals and 2,586 control subjects. We did not observe any evidence for association of *ZFPM2* rs4602861 with VTE risk, overall or in any of the three replication studies (Table 2), despite having 93% power in the combined replication studies to detect at the 0.017 (=0.05 / 3) statistical threshold the effect of a SNP with RAF 0.76 and associated with an OR of 1.20.[25] Conversely, we confirmed association for *TSPAN15* rs78707713 and *SLC44A2* rs2288904 SNPs (Table 2). In meta-analyzed replication results, the common *TSPAN15* rs78707713-T allele was associated with an OR for VTE of 1.42 ($p = 2.21 \times 10^{-7}$) (Figure S5), and the common *SLC44A2* rs2288904-G allele with an increased risk of 1.28 ($p = 2.64 \times 10^{-7}$) (Figure S6). When the results obtained in the discovery and the replication studies were meta-analyzed, the summary OR for VTE was 1.31 ($p = 1.67 \times 10^{-16}$) for *TSPAN15* rs78707713 (Figure S5) and 1.21 ($p = 2.75 \times 10^{-15}$) for *SLC44A2* rs2288904 (Figure S6). No heterogeneity across the 15 studies was present: $Q = 18.6$, $I^2 = 0.25$, and p = 0.18 for *TSPAN15* rs78707713 and Q = 7.40, $I^2 = 0.00$, and p = 0.92 for *SLC44A2* rs2288904.

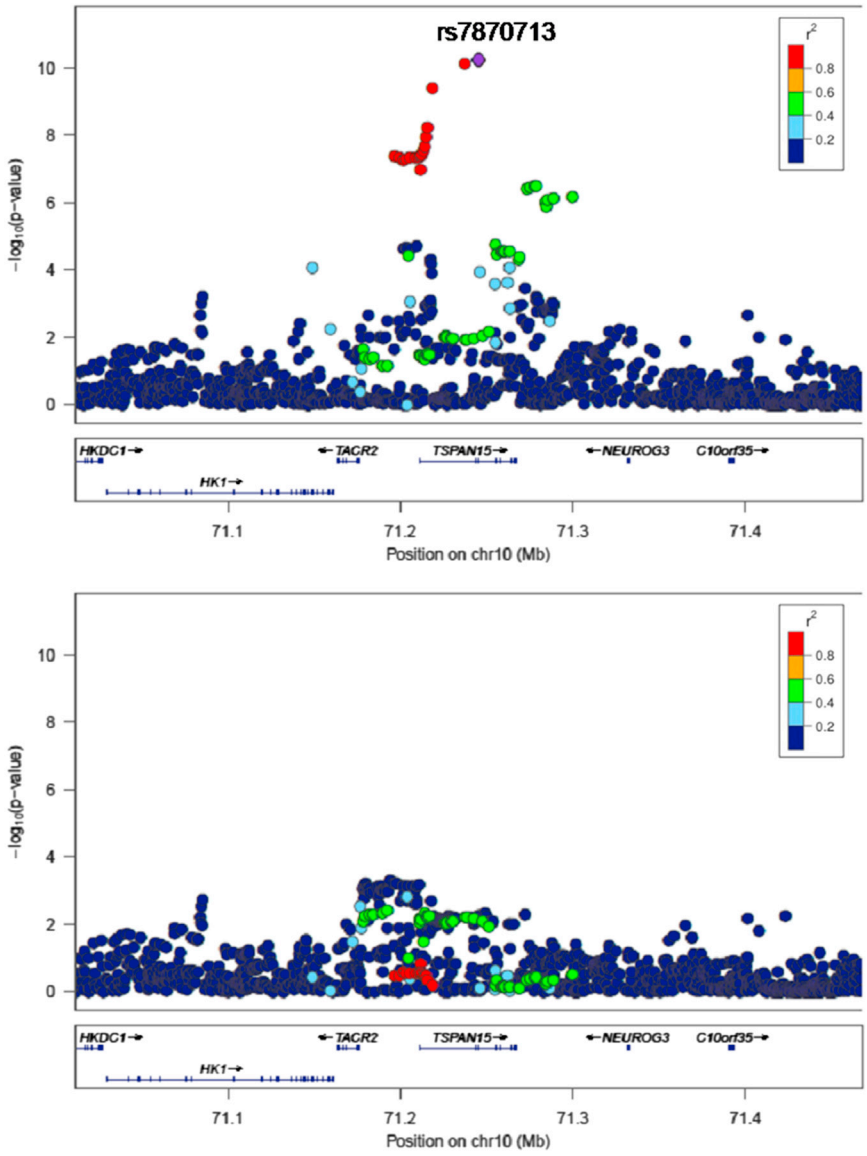

Figure 1. Regional Association Plot at the *TSPAN15* Locus

Association results were derived from the meta-analysis of 12 GWASs (top) that were further conditioned on the effect of *TSPAN15* rs7870713 (bottom). SNPs are colored according to their pairwise LD r^2 with the lead rs7870713. r^2 was estimated from 1000 Genomes (Mar 2012 Eur) database.

Visual inspection of the regional association plots at the *TSPAN15* (Figure 1) and *SLC44A2* (Figure 2) suggested that all SNPs at these loci with high level of significance for association with VTE were in strong LD with the lead SNPs. This was confirmed by the results of the conditional analysis that did not produce any other signal significant at the 5×10^{-8} statistical threshold. After adjusting for rs78707713, the lowest p value observed at *TSPAN15* was $p = 4.92 \times 10^{-4}$ for the intronic rs1072160 (Figure 1). Conditioning on rs2288904 abolished all associations at the *SLC44A2* locus, the lowest p value then being $p = 0.047$ for the *KEAP1* rs45524632 (Figure 2).

In subgroup analyses of the replication studies, the genotype distribution of *TSPAN15* rs78707713 and *SLC44A2* rs2288904 did not differ according to the clinical manifestations of VTE, either PE or DVT (Table S2). We did not observe heterogeneity in the effects of the two SNPs according to sex nor to *F5* rs6025 or *F2* rs1799963 mutations (Tables S3 and S4).

Although the six known VTE-associated loci affect VT risk through a modulation of known hemostatic traits (i.e., levels of von Willebrand factor and Factor VIII for *ABO*, of FXI for *F11*, of endogenous thrombin potential for *F2*, of resistance to activated protein C for *F5*, of fibrinogen for *FGG*, and of protein C for *PROCR*[26–33]), the loci of the two discovered genetic associations were not in or near genes that are currently known to influence hemostasis. We then explored whether these SNPs were associated with well-characterized hemostasis phenotypes that are associated with thrombotic propensity. We did not find any evidence for an association with 25 plasma biomarkers (Table S5) even though the sample size of the investigated studies were sufficiently statistically powered (~95%) to detect the additive allele effect of any SNP that would explain 1% of the variability of a quantitative trait.[34] By contrast, the anticipated effects of the six known VTE-associated loci were observed in these studies (Table S6).

We interrogated genome-wide gene expression in ten tissues and DNA methylation studies in blood to determine whether the *TSPAN15* and *SLC44A2* SNPs influenced the regulation of their associated genes. We observed significant associations of the *TSPAN15* rs78707713 with *TSPAN15* DNA methylation measured from peripheral blood DNA and *TSPAN15* expression in macrophages, endothelial cells, and esophagus mucosa (Table 3). The *SLC44A2* rs2288904 was found significantly associated with *SLC44A2* gene expression in monocytes, macrophages, and whole blood (Table 4). However, for none of the interrogated bio-resources for gene expression or DNA methylation levels did the identified disease-associated SNPs show the strongest locus-wide effects. Of note, most SNPs that demonstrated the greatest influence on gene regulation at the replicated loci showed weak or null associations with VTE risk (Tables 3 and 4). Further

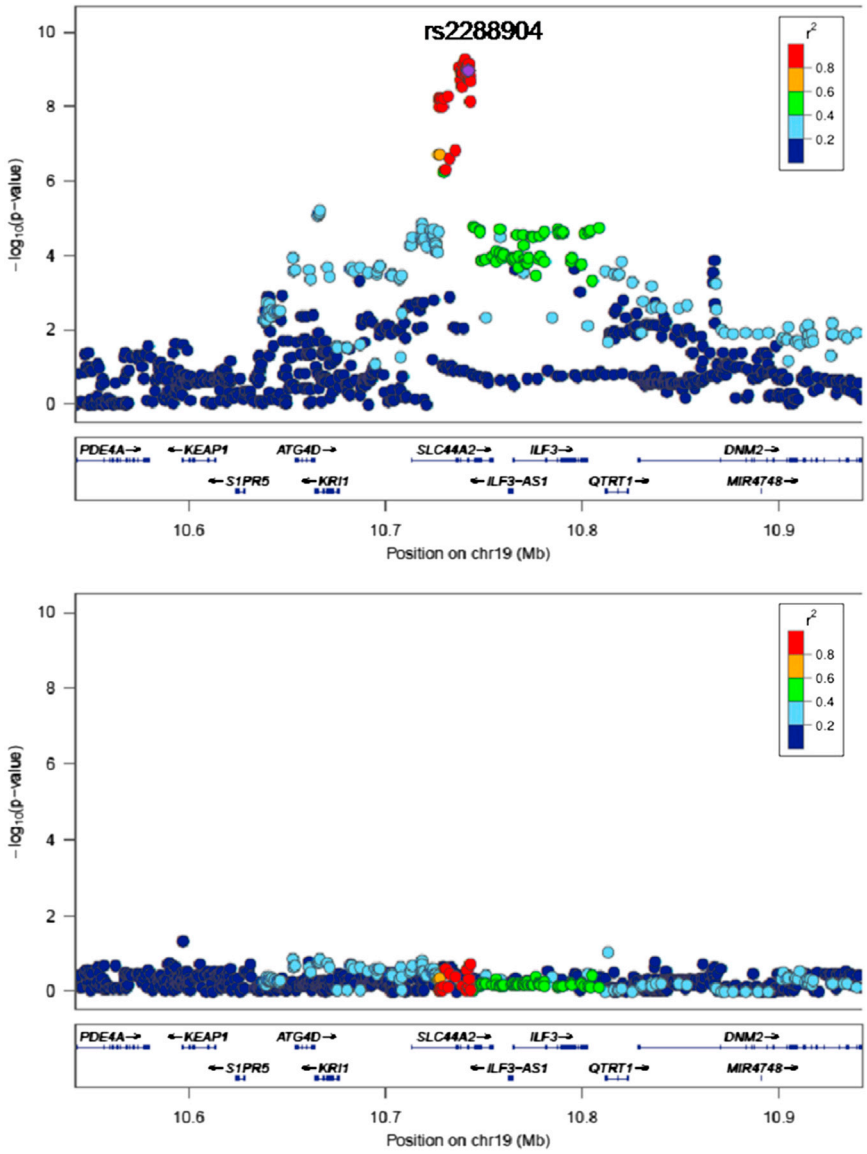

Figure 2. Regional Association Plot at the *SLC44A2* Locus

Association results were derived from the meta-analysis of 12 GWASs (top) that were further conditioned on the effect of *SLC44A2* rs2288904 (bottom). SNPs are colored according to their pairwise LD r^2 with the lead rs2288904. r^2 was estimated from 1000 Genomes (Mar 2012 Eur) database.

analyses revealed that the observed effects of *TSPAN15* rs78707713 and *SLC44A2* rs2288904 on gene expression and DNA methylation were probably due to their linkage disequilibrium with other regulatory variants with stronger effects (data not shown).

Discussion

By using data from more than 7,000 case subjects and more than 50,000 control subjects, which represents a 4-fold increase in the number of VTE events used in our discovery effort compared with that used in the latest meta-analysis of GWASs for VTE, we identified and then replicated two associations for common variants in *TSPAN15* and *SLC44A2*. The strengths of the associations were modest in size with ORs of 1.3 or less, but the statistical evidence was robust in the discovery and replication stages. Further, we observed that the variants were not associated with dozens of hemostatic markers characterizing the coagulation/fibrinolysis balance. The identified VTE-associated SNPs map to genes that are not in conventional pathways to thrombosis that have marked most of the genetic associations to date, suggesting that these genetic variants represent novel biological pathways leading to VTE.

The common T allele, with frequency ~0.89, of the identified *TSPAN15* rs78707713 was associated with an increased risk of 1.31-fold. The *TSPAN15* rs78707713 is intronic and the interrogation of several gene expression databases as well as the application of prediction/annotation tools[46–48] did not suggest any regulatory elements supporting a functional role of this SNP. It is likely that the SNP is in strong LD with yet unidentified culprit variant(s). For instance, rs78707713 is in strong LD (pairwise $r^2 = 0.89$) with the intronic *TSPAN15* rs17490626 predicted[46,47] to map an enhancer domain, the significance of the VTE association of the latter SNP being $p = 3.74 \times 10^{-10}$. No association of *TSPAN15*-coding SNPs with VTE risk was observed in the discovery meta-analysis (smallest p value = 0.49). *TSPAN15* codes for tetraspanin 15, a member of the tetraspanin superfamily that act as scaffolding proteins, anchoring multiple proteins to the cell membrane.[49] Members of the tetraspanin family have roles in cells that regulate hemostasis. *TSPAN24* (CD151)-[50] and *TSPAN32* (TSSC6)-[51] deficient mice exhibit a bleeding phenotype with impaired "outside-in" signaling through $\alpha_{IIb}\beta_3$, the major platelet integrin. CD63 (*TSPAN30*) facilitates the release of von Willebrand factor (VWF) from endothelial cell Weibel-Palade bodies through transient enhancement of fusion between the Weibel-Palade body membrane and the plasma membrane,[52] which probably contributes to the leucocyte attachment to the endothelium.[53]

The risk allele at the second identified locus, *SLC44A2* rs2288904-G, was also common, with frequency ~0.77. It

Table 3. *TSPAN15* SNPs Showing the Strongest Influence on *TSPAN15* Gene Expression and DNA Methylation Levels in Various Human Bio-resources and Their Relation with the *TSPAN15* VTE-Associated rs78707713

Bio-resource	Cell Type or Tissue	Interrogated Bio-resources	Best *cis* eQTL/mQTL SNP			rs78707713	r^2 between rs78707713 and Best rsID[a]
			rsID	eQTL/mQTL p Value	VTE Association p Value	eQTL/mQTL p Value	
	B cells	Fairfax et al.[35]	rs7897621	0.0053	0.267	0.129	0.001
	endothelial cells	Erbilgin et al.[36]	rs768498	6.87×10^{-27}	0.807	2.80×10^{-11}	0.180
	esophagus mucosa	GTEx Consortium[37]	rs28463525	2.2×10^{-17}	5.64×10^{-9}	2.4×10^{-16}	0.841
	heart, left ventricle	GTEx Consortium[37]	rs10823376	5.9×10^{-9}	0.911	NA	0.217
	heart	Folkersen et al.[38]	rs768498	9.25×10^{-10}	0.807	0.038[b]	0.180
	intestine	Kabakchiev and Silverberg[39]	rs4565792	2.28×10^{-5}	0.639	0.373	0.242
	nerve-tibial	Folkersen et al.[38]	rs34187097	5.8×10^{-10}	0.507	NA	0.000
Gene expression	liver	Folkersen et al.[38]	rs2084274	9.02×10^{-4}	0.195	0.092[b]	0.011
	liver	Schadt et al.[40]	rs972570	3.16×10^{-30}	3.31×10^{-4}	NA	0.196
	macrophages	Garnier et al.[41]	rs768498	1.85×10^{-39}	0.807	1.12×10^{-6}	0.180
	monocytes	Fairfax et al.[35]	rs2812541	0.012	0.356	0.991	0.013
	monocytes	Garnier et al.[41]	rs10128334	8.05×10^{-3}	0.867	0.978	0.004*
	2 hr LPS-stimulated monocytes	Fairfax et al.[42]	rs10823371	3.50×10^{-4}	0.701	0.233	0.217
	24 hr LPS-stimulated monocytes	Fairfax et al.[42]	rs1052179	2.90×10^{-3}	0.073	0.693	0.148
	2 hr IFN-stimulated monocytes	Fairfax et al.[42]	rs5030949	1.65×10^{-3}	0.025	0.307	0.004
DNA methylation	whole blood	Dick et al.[43]	rs12416520	2.24×10^{-187}	0.035	3.53×10^{-46}	0.489

[a]Pairwise r^2 was derived from the SNAP database[45] except for result noted with asterisk (*) where linkage disequilibrium was estimated from the MARTHA GWAS imputed genotypes.
[b]In the ASAP study, the rs12242391 served as a proxy ($r^2 = 0.84$) for rs78707713 that was not typed.

was associated with a relative risk of 1.21 and is probably the functional variant. Indeed, the observed risk allele (G) codes for the Arg154 isoform of the choline transporter-like protein 2 (CTL-2). CTL-2 has been associated with several human diseases,[54] including transfusion-related acute lung injury (TRALI). TRALI is a life-threatening complication of blood transfusion and the leading cause of transfusion-associated mortality in developed countries. Severe TRALI is due to antibodies in blood components directed against the human neutrophil alloantigen-3a (HNA-3a), which is determined by the Arg154 isoform.[55,56] Greinacher et al.[55] found that alloantibodies targeting CTL-2 lead to leucocyte activation and aggregation. Recently, human anti-HNA-3a antibodies were shown to directly interact with endothelial CTL-2 to disturb pulmonary endothelial barrier function that would lead to severe TRALI.[57] It might be that carrying the HNA-3a antigen, determined by the Arg154 isoform, favors activation of leucocytes/neutrophils and endothelial cells in some triggering circumstances.

We did not observe any difference in *TSPAN15* and *SLC44A2* VTE-associated SNP allele frequencies between DVT- and PE-affected individuals in the replication populations, the two main clinical manifestations of VTE. This observation suggests that the underlying pathophysiological mechanisms are more likely to be involved in thrombus formation rather than its rupture and its migration toward the pulmonary vein. About 20% of persons with unprovoked VTE (i.e., occurring without clear external factors like surgery, trauma, immobilization, hormone use, or cancer) will experience a recurrent event, even after a 6-month course of anticoagulant prophylaxis. Because we had little information about recurrence follow-up in our studies, we were not able to assess whether the identified SNPs can discriminate between individuals that will and those that will not face a recurrent event. Investigating whether these SNPs can help improve the secondary prevention of VTE would definitively be warranted.

Despite having gathered the largest GWAS samples of VTE-affected individuals, our approach was not well powered to identify common SNPs associated with more modest effects than those observed for *SLC44A2* and *TSPAN15*, i.e., with OR < 1.20, or extremely rare mutations (e.g., with frequency $< 1\%$) that cannot be efficiently tagged by the tested SNPs (Table S7). Heterogeneity in the design

Table 4. *SLC44A2* SNPs Showing the Strongest Influence on *SLC44A2* Gene Expression in Various Human Bio-resources and Their Relation with the *SLC44A2* VTE-Associated rs2288904

Cell Type or Tissue	Interrogated Bio-resources	Best *cis* eQTL SNP			rs2288904	R^2 between rs2288904 and Best eSNP[a]
		rsID	eQTL p Value	VTE Association p Value	eQTL p Value	
B cells	Fairfax et al.[35]	rs8106664	3.99×10^{-4}	6.12×10^{-9}	2.72×10^{-3}	0.906
Heart	Folkersen et al.[38]	rs3760648	8.59×10^{-4}	0.293	0.868	0.000
Intestine	Kabakchiev and Silverberg[39]	rs11672431	5.71×10^{-5}	9.70×10^{-3}	0.050	0.116
Liver	Schadt et al.[40]	rs7251213	4.01×10^{-6}	3.57×10^{-3}	NA	0.160
Liver	Folkersen et al.[38]	rs11672431	5.29×10^{-4}	9.70×10^{-3}	0.115	0.116
Macrophages	Garnier et al.[41]	rs3859514	1.92×10^{-12}	1.34×10^{-4}	1.36×10^{-9}	0.393
Monocytes	Garnier et al.[41]	rs62129987	4.85×10^{-25}	1.64×10^{-5}	2.71×10^{-12}	0.440*
Monocytes	Fairfax et al.[35]	rs7252007	7.39×10^{-30}	6.98×10^{-3}	5.01×10^{-9}	0.345
24 hr LPS-stimulated monocytes	Fairfax et al.[42]	rs7252007	4.03×10^{-10}		8.47×10^{-5}	
2 hr LPS-stimulated monocytes	Fairfax et al.[42]	rs7252007	1.06×10^{-5}		4.70×10^{-4}	
2 hr IFN-stimulated monocytes	Fairfax et al.[42]	rs12609501	1.92×10^{-12}	9.23×10^{-3}	1.57×10^{-7}	0.072
Whole blood	Westra et al.[44]	rs892078	3.64×10^{-93}	2.11×10^{-5}	3.85×10^{-70}	0.438

Of note, no CpG probe targeting the *SLC44A2* locus satisfied the adopted quality-control procedures, preventing us from efficiently testing for the effect of *SLC44A2* rs2288904 on DNA methylation levels at this locus in the interrogated bio-resources.
[a]Pairwise r^2 was derived from the SNAP database[45] except for result noted with asterisk (*) where linkage disequilibrium was estimated from the MARTHA GWAS imputed genotypes.

and clinical characteristics of the studied populations might also have contributed to slightly attenuate the power of our study to detect additional genome-wide significant associations. Further investigations deserve to be conducted to better characterize the 202 suggestive statistical associations with $p < 10^{-5}$ in our discovery meta-analysis and to identify additional susceptibility loci for VTE.

Though additional work is needed, including the identification of the functional *TSPAN15* variant(s), our results demonstrated that *SLC44A2* and *TSPAN15* are two susceptibility loci for VTE. The products of the two genes, expressed by cells that are central to the pathophysiology of thrombosis (monocytes/macrophages and endothelial cells), are not known to have central roles in the traditional hemostasis pathways nor to other cardiovascular diseases. These results pave the way for novel mechanistic concepts of VTE pathophysiology, new biomarkers for the disease, and novel therapeutic perspectives.

Supplemental Data

Supplemental Data include six figures, ten tables, Supplemental Acknowledgments, funding information for each cohort, and a list of members of the INVENT consortium and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2015.01.019.

Acknowledgments

This work is the product of the International Network against Thrombosis (INVENT) collaboration. This work was financially supported by the French National Institute for Health and Medical Research; French Ministry of Health; French Medical Research Foundation; French Agency for Research; National Heart, Lung, and Blood Institute (USA); National Institute for Health Research (USA); National Human Genome Research Institute (USA); National Cancer Institute (USA); the Donald W. Reynolds Foundation; Fondation Leducq; The European Commission; Netherlands Organisation for Scientific Research; Netherlands Heart Foundation; Dutch Cancer Foundation; Canadian Institutes of Health Research; Heart and Stroke Foundation of Canada; British Heart Foundation; and ICAN Institute for Cardiometabolism and Nutrition. The lists of grants supporting this research can be found in the Supplemental Data.

Received: September 8, 2014
Accepted: January 29, 2015
Published: March 12, 2015

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
OMIM, http://www.omim.org/
RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq

References

1. Naess, I.A., Christiansen, S.C., Romundstad, P., Cannegieter, S.C., Rosendaal, F.R., and Hammerstrøm, J. (2007). Incidence and mortality of venous thrombosis: a population-based study. J. Thromb. Haemost. *5*, 692–699.

2. Virchow, R. (1856). Gesammelte Abhandlungen zur Wissenschaftlichen Medizin (Frankfurt: Meidinger).
3. Heit, J.A., Phelps, M.A., Ward, S.A., Slusser, J.P., Petterson, T.M., and De Andrade, M. (2004). Familial segregation of venous thromboembolism. J. Thromb. Haemost. *2*, 731–736.
4. Zöller, B., Li, X., Sundquist, J., and Sundquist, K. (2011). Age- and gender-specific familial risks for venous thromboembolism: a nationwide epidemiological study based on hospitalizations in Sweden. Circulation *124*, 1012–1020.
5. Tang, W., Teichert, M., Chasman, D.I., Heit, J.A., Morange, P.-E., Li, G., Pankratz, N., Leebeek, F.W., Paré, G., de Andrade, M., et al. (2013). A genome-wide association study for venous thromboembolism: the extended cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. Genet. Epidemiol. *37*, 512–521.
6. Germain, M., Saut, N., Oudot-Mellakh, T., Letenneur, L., Dupuy, A.-M., Bertrand, M., Alessi, M.C., Lambert, J.C., Zelenika, D., Emmerich, J., et al. (2012). Caution in interpreting results from imputation analysis when linkage disequilibrium extends over a large distance: a case study on venous thrombosis. PLoS ONE *7*, e38538.
7. Morange, P.E., and Trégouët, D.A. (2013). Current knowledge on the genetics of incident venous thrombosis. J. Thromb. Haemost. *11* (1), 111–121.
8. The ARIC investigators (1989). The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am. J. Epidemiol. *129*, 687–702.
9. Fried, L.P., Borhani, N.O., Enright, P., Furberg, C.D., Gardin, J.M., Kronmal, R.A., Kuller, L.H., Manolio, T.A., Mittelmark, M.B., Newman, A., et al. (1991). The Cardiovascular Health Study: design and rationale. Ann. Epidemiol. *1*, 263–276.
10. Tell, G.S., Fried, L.P., Hermanson, B., Manolio, T.A., Newman, A.B., and Borhani, N.O. (1993). Recruitment of adults 65 years and older as participants in the Cardiovascular Health Study. Ann. Epidemiol. *3*, 358–366.
11. Trégouët, D.-A., Heath, S., Saut, N., Biron-Andreani, C., Schved, J.-F., Pernod, G., Galan, P., Drouet, L., Zelenika, D., Juhan-Vague, I., et al. (2009). Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. Blood *113*, 5298–5303.
12. de Visser, M.C.H., van Minkelen, R., van Marion, V., den Heijer, M., Eikenboom, J., Vos, H.L., Slagboom, P.E., Houwing-Duistermaat, J.J., Rosendaal, F.R., and Bertina, R.M. (2013). Genome-wide linkage scan in affected sibling pairs identifies novel susceptibility region for venous thromboembolism: Genetics In Familial Thrombosis study. J. Thromb. Haemost. *11*, 1474–1484.
13. Psaty, B.M., Heckbert, S.R., Koepsell, T.D., Siscovick, D.S., Raghunathan, T.E., Weiss, N.S., Rosendaal, F.R., Lemaitre, R.N., Smith, N.L., Wahl, P.W., et al. (1995). The risk of myocardial infarction associated with antihypertensive drug therapies. JAMA *274*, 620–625.
14. Germain, M., Saut, N., Greliche, N., Dina, C., Lambert, J.-C., Perret, C., Cohen, W., Oudot-Mellakh, T., Antoni, G., Alessi, M.C., et al. (2011). Genetics of venous thrombosis: insights from a new genome wide association study. PLoS ONE *6*, e25581.
15. Heit, J.A., Armasu, S.M., Asmann, Y.W., Cunningham, J.M., Matsumoto, M.E., Petterson, T.M., and De Andrade, M. (2012). A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. J. Thromb. Haemost. *10*, 1521–1531.
16. Blom, J.W., Doggen, C.J.M., Osanto, S., and Rosendaal, F.R. (2005). Malignancies, prothrombotic mutations, and the risk of venous thrombosis. JAMA *293*, 715–722.
17. Hankinson, S.E., Colditz, G.A., Hunter, D.J., Manson, J.E., Willett, W.C., Stampfer, M.J., Longcope, C., and Speizer, F.E. (1995). Reproductive factors and family history of breast cancer in relation to plasma estrogen and prolactin levels in postmenopausal women in the Nurses' Health Study (United States). Cancer Causes Control *6*, 217–224.
18. Tworoger, S.S., Sluss, P., and Hankinson, S.E. (2006). Association between plasma prolactin concentrations and risk of breast cancer among predominately premenopausal women. Cancer Res. *66*, 2476–2482.
19. Ridker, P.M., Chasman, D.I., Zee, R.Y.L., Parker, A., Rose, L., Cook, N.R., and Buring, J.E.; Women's Genome Health Study Working Group (2008). Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. Clin. Chem. *54*, 249–255.
20. Oger, E., Lacut, K., Le Gal, G., Couturaud, F., Guénet, D., Abalain, J.-H., Roguedas, A.M., and Mottier, D.; EDITH COLLABORATIVE STUDY GROUP (2006). Hyperhomocysteinemia and low B vitamin levels are independently associated with venous thromboembolism: results from the EDITH study: a hospital-based case-control study. J. Thromb. Haemost. *4*, 793–799.
21. Johnson, E.O., Hancock, D.B., Levy, J.L., Gaddis, N.C., Saccone, N.L., Bierut, L.J., and Page, G.P. (2013). Imputation across genotyping arrays for genome-wide association studies: assessment of bias and a correction strategy. Hum. Genet. *132*, 509–522.
22. Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190–2191.
23. Li, M.-X., Yeung, J.M.Y., Cherny, S.S., and Sham, P.C. (2012). Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. Hum. Genet. *131*, 747–756.
24. Panagiotou, O.A., and Ioannidis, J.P.A.; Genome-Wide Significance Project (2012). What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. Int. J. Epidemiol. *41*, 273–286.
25. Skol, A.D., Scott, L.J., Abecasis, G.R., and Boehnke, M. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat. Genet. *38*, 209–213.
26. Smith, N.L., Chen, M.H., Dehghan, A., Strachan, D.P., Basu, S., Soranzo, N., Hayward, C., Rudan, I., Sabater-Lleal, M., Bis, J.C., et al.; Wellcome Trust Case Control Consortium (2010). 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 *121*, 1382–1392.
27. Li, Y., Bezemer, I.D., Rowland, C.M., Tong, C.H., Arellano, A.R., Catanese, J.J., Devlin, J.J., Reitsma, P.H., Bare, L.A., and Rosendaal, F.R. (2009). Genetic variants associated with deep

vein thrombosis: the F11 locus. J. Thromb. Haemost. *7*, 1802–1808.
28. Rocanin-Arjo, A., Cohen, W., Carcaillon, L., Frère, C., Saut, N., Letenneur, L., Alhenc-Gelas, M., Dupuy, A.M., Bertrand, M., Alessi, M.C., et al.; CardioGenics Consortium (2014). A meta-analysis of genome-wide association studies identifies ORM1 as a novel gene controlling thrombin generation potential. Blood *123*, 777–785.
29. Oudot-Mellakh, T., Cohen, W., Germain, M., Saut, N., Kallel, C., Zelenika, D., Lathrop, M., Trégouët, D.A., and Morange, P.E. (2012). Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. Br. J. Haematol. *157*, 230–239.
30. Sabater-Lleal, M., Huang, J., Chasman, D., Naitza, S., Dehghan, A., Johnson, A.D., Teumer, A., Reiner, A.P., Folkersen, L., Basu, S., et al.; VTE Consortium; STROKE Consortium; Wellcome Trust Case Control Consortium 2 (WTCCC2); C4D Consortium; CARDIoGRAM Consortium (2013). Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. Circulation *128*, 1310–1324.
31. Smith, N.L., Huffman, J.E., Strachan, D.P., Huang, J., Dehghan, A., Trompet, S., Lopez, L.M., Shin, S.Y., Baumert, J., Vitart, V., et al. (2011). Genetic predictors of fibrin D-dimer levels in healthy adults. Circulation *123*, 1864–1872.
32. Tang, W., Schwienbacher, C., Lopez, L.M., Ben-Shlomo, Y., Oudot-Mellakh, T., Johnson, A.D., Samani, N.J., Basu, S., Gögele, M., Davies, G., et al. (2012). Genetic associations for activated partial thromboplastin time and prothrombin time, their gene expression profiles, and risk of coronary artery disease. Am. J. Hum. Genet. *91*, 152–162.
33. Tang, W., Basu, S., Kong, X., Pankow, J.S., Aleksic, N., Tan, A., Cushman, M., Boerwinkle, E., and Folsom, A.R. (2010). Genome-wide association study identifies novel loci for plasma levels of protein C: the ARIC study. Blood *116*, 5032–5036.
34. Gauderman, W.J. (2002). Sample size requirements for matched case-control studies of gene-environment interaction. Stat. Med. *21*, 35–50.
35. Fairfax, B.P., Makino, S., Radhakrishnan, J., Plant, K., Leslie, S., Dilthey, A., Ellis, P., Langford, C., Vannberg, F.O., and Knight, J.C. (2012). Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat. Genet. *44*, 502–510.
36. Erbilgin, A., Civelek, M., Romanoski, C.E., Pan, C., Hagopian, R., Berliner, J.A., and Lusis, A.J. (2013). Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. J. Lipid Res. *54*, 1894–1905.
37. GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. *45*, 580–585.
38. Folkersen, L., van't Hooft, F., Chernogubova, E., Agardh, H.E., Hansson, G.K., Hedin, U., Liska, J., Syvänen, A.C., Paulsson-Berne, G., Franco-Cereceda, A., et al.; BiKE and ASAP study groups (2010). Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. Circ Cardiovasc Genet *3*, 365–373.
39. Kabakchiev, B., and Silverberg, M.S. (2013). Expression quantitative trait loci analysis identifies associations between genotype and gene expression in human intestine. Gastroenterology *144*, 1488–1496, e1–e3.
40. Schadt, E.E., Molony, C., Chudin, E., Hao, K., Yang, X., Lum, P.Y., Kasarskis, A., Zhang, B., Wang, S., Suver, C., et al. (2008). Mapping the genetic architecture of gene expression in human liver. PLoS Biol. *6*, e107.
41. Garnier, S., Truong, V., Brocheton, J., Zeller, T., Rovital, M., Wild, P.S., Ziegler, A., Munzel, T., Tiret, L., Blankenberg, S., et al.; Cardiogenics Consortium (2013). Genome-wide haplotype analysis of cis expression quantitative trait loci in monocytes. PLoS Genet. *9*, e1003240.
42. Fairfax, B.P., Humburg, P., Makino, S., Naranbhai, V., Wong, D., Lau, E., Jostins, L., Plant, K., Andrews, R., McGee, C., and Knight, J.C. (2014). Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. Science *343*, 1246949.
43. Dick, K.J., Nelson, C.P., Tsaprouni, L., Sandling, J.K., Aïssi, D., Wahl, S., Meduri, E., Morange, P.-E., Gagnon, F., Grallert, H., et al. (2014). DNA methylation and body-mass index: a genome-wide analysis. Lancet *383*, 1990–1998.
44. Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., et al. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. *45*, 1238–1243.
45. Johnson, A.D., Handsaker, R.E., Pulit, S.L., Nizzari, M.M., O'Donnell, C.J., and de Bakker, P.I. (2008). SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics *24*, 2938–2939.
46. Ward, L.D., and Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. *40*, D930–D934.
47. Barenboim, M., and Manke, T. (2013). ChroMoS: an integrated web tool for SNP classification, prioritization and functional interpretation. Bioinformatics *29*, 2197–2198.
48. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. *22*, 1790–1797.
49. Berditchevski, F. (2001). Complexes of tetraspanins with integrins: more than meets the eye. J. Cell Sci. *114*, 4143–4151.
50. Lau, L.-M., Wee, J.L., Wright, M.D., Moseley, G.W., Hogarth, P.M., Ashman, L.K., and Jackson, D.E. (2004). The tetraspanin superfamily member CD151 regulates outside-in integrin alphaIIbbeta3 signaling and platelet function. Blood *104*, 2368–2375.
51. Goschnick, M.W., Lau, L.-M., Wee, J.L., Liu, Y.S., Hogarth, P.M., Robb, L.M., Hickey, M.J., Wright, M.D., and Jackson, D.E. (2006). Impaired "outside-in" integrin alphaIIbbeta3 signaling and thrombus stability in TSSC6-deficient mice. Blood *108*, 1911–1918.
52. Bailey, R.L., Herbert, J.M., Khan, K., Heath, V.L., Bicknell, R., and Tomlinson, M.G. (2011). The emerging role of tetraspanin microdomains on endothelial cells. Biochem. Soc. Trans. *39*, 1667–1673.
53. Poeter, M., Brandherm, I., Rossaint, J., Rosso, G., Shahin, V., Skryabin, B.V., Zarbock, A., Gerke, V., and Rescher, U. (2014). Annexin A8 controls leukocyte recruitment to activated endothelial cells via cell surface delivery of CD63. Nat. Commun. *5*, 3738.

54. Traiffort, E., O'Regan, S., and Ruat, M. (2013). The choline transporter-like family SLC44: properties and roles in human diseases. Mol. Aspects Med. *34*, 646–654.
55. Greinacher, A., Wesche, J., Hammer, E., Fürll, B., Völker, U., Reil, A., and Bux, J. (2010). Characterization of the human neutrophil alloantigen-3a. Nat. Med. *16*, 45–48.
56. Curtis, B.R., Cox, N.J., Sullivan, M.J., Konkashbaev, A., Bowens, K., Hansen, K., and Aster, R.H. (2010). The neutrophil alloantigen HNA-3a (5b) is located on choline transporter-like protein 2 and appears to be encoded by an R>Q154 amino acid substitution. Blood *115*, 2073–2076.
57. Bayat, B., Tjahjono, Y., Sydykov, A., Werth, S., Hippenstiel, S., Weissmann, N., Sachs, U.J., and Santoso, S. (2013). Anti-human neutrophil antigen-3a induced transfusion-related acute lung injury in mice by direct disturbance of lung endothelial cells. Arterioscler. Thromb. Vasc. Biol. *33*, 2538–2548.