Is there still room for additional common susceptibility alleles for venous thromboembolism ?

D-A Trégouët, A Delluc, David Roche, C. Derbois, Robert Olaso, Marie Germain, M. de Andrade, W. Tang, D. I. Chasman, Astrid van Hylckama Vlieg, et al.

Is there still room for additional common susceptibility alleles for venous thromboembolism ?

D-A. Trégouët, * †, A. Delluc, ‡ A. Roche, § C. Derbois, ¶ R. Olaso, ¶, M. Germain, * † M. de Andrade,** W. Tang, †† D. I. Chasman, ‡‡ A. van Hylckama Vlieg, §§ P. Reitsma, ¶¶, C. Kabrhel, *** N. Smith, ††† ‡‡‡ §§§ P-E. Morange, ¶¶¶ ****

* Sorbonne Universités, UPMC Univ. Paris 06, INSERM, UMR_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, Paris, France

† ICAN Institute for Cardiometabolism and Nutrition, Paris, France

‡ Université de Brest, EA3878 and CIC1412, 29238 Brest, France

§ Service de Pneumologie et soins intensifs respiratoires, AP-HP, Hôpital Européen Georges Pompidou, Paris, France, Université Paris Descartes, Sorbonne Paris Cité, France, Inserm UMR-S1140, Paris, France

¶ Centre National de Génotypage, Institut de Génomique, CEA, 91057 Evry, France.

** Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN 55905, USA

†† Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN 55454, USA

‡‡ Division of Preventive Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02215, USA

§§ Department of Thrombosis and Hemostasis, Department of Clinical Epidemiology, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands

¶¶ Einthoven Laboratory for Experimental Vascular Medicine, Department of Thrombosis and Hemostasis, Leiden University Medical Center, 2300 RC Leiden, the Netherlands;

*** Department of Emergency Medicine, Massachusetts General Hospital, Channing Network Medicine, Harvard Medical School, Boston, MA 2114, USA

††† Department of Epidemiology, University of Washington, Seattle, WA 98195, USA

‡‡‡ Group Health Research Institute, Group Health Cooperative, Seattle, WA 98101, USA;

§§§ Seattle Epidemiologic Research and Information Center, VA Office of Research and Development, Seattle, WA 98108, USA

¶¶¶ Laboratory of Haematology, La Timone Hospital, Marseille, France;

**** Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR_S) 1062, Nutrition Obesity and Risk of Thrombosis, Marseille, France; Aix-Marseille University

Running head title: Follow-up of the INVENT meta-analysis

Correspondence: Dr David-Alexandre Trégouët, INSERM UMR_S 1166, Room 405, 4th floor, 91 blvd de l'Hôpital , 75013 Paris , France; email: david.tregouet@upmc.fr: tel: +33 1 40 77 96 86; fax: + 33 1 40 77 97 28

Essentials

- Genetic architecture of venous thromboembolism (VTE) remains to be fully disentangled.
- 11 newly discovered candidate polymorphisms were genotyped in 3,019 VTE cases and 2,605 controls
- None of the 11 polymorphism were significantly associated with VTE risk
- Additional major efforts are needed to identify VTE-associated genetic variants

Abstract

Background Through a meta-analysis of twelve Genome-wide Association Studies, the INVENT consortium identified two novel susceptibility loci for venous thromboembolism (VTE). This project has also generated other candidates that need to be further replicated.

Objectives To assess the association with VTE of common SNPs that demonstrated strong statistical, but not genome-wide, significance in the INVENT cohorts.

Patients/Methods Eleven SNPs were genotyped and tested for association with VTE in three case-control studies totaling 3,019 patients and 2,605 healthy individuals.

Results and Conclusions None of the tested SNPs showed evidence for association with VTE. Different strategies are needed to decipher the whole spectrum of common and rare genetic variations associated with VTE risk.

Keywords

genome-wide association studies, genetics, polymorphisms, risk factor, venous thromboembolism

Introduction

In 2015, the International Network against VENous Thrombosis (INVENT) consortium reported the results of a meta-analysis of twelve Genome-Wide Association studies (GWAS) on venous thromboembolism (VTE) [1]. In the discovery phase of this study which was composed of 7,507 VTE cases and 52,632 controls nine loci reached the genome-wide threshold of $5x10^{-8}$ for declaring statistical significance. Three of these loci, *SLC44A2*, *TSPAN15* and *ZFPM2*, had not previously been reported to associate with VTE risk. The first two replicated in three independent French case-control studies, MARTHA12, EDITH and FARIVE [1]. The association of *SLC44A2* and *TSPAN15* loci with VTE was also confirmed in a recent GWAS analysis of 6,135 self-reported VTE patients belonging to the 23andMe cohorts [2]. A meta-analysis conditioning on the well-established VTE-associated single nucleotide polymorphisms (SNPs) was also performed in the INVENT discovery cohorts but failed to identify any additional genome-wide significant associations [1]. However, at a lower statistical threshold of $1x10^{-5}$, 285 candidate SNP were identified in this conditional, 12-cohorts meta-analysis. After excluding associations at known VTE loci, and those at the *ZFPM2* locus that did not replicate previously, 125 SNPs mapping to 11 loci remained. We reasoned that there might be true positive VTE-associations among these 11 loci and decided to further genotype the lead SNPs in the three French case-control studies that served as replication for the INVENT meta-analysis.

Material and Methods

Study populations

Three French case-control studies for VTE were used in this work: EDITH, FARIVE and MARTHA12. These are exactly those that had been employed for replicating the main findings of the INVENT meta-GWAS [1].

EDITH - The EDITH study is a case–control study that was designed to test interactions between genetic and environmental risk factors of VTE [3]. Between May 2000 and December 2009, all consecutive unselected in- and outpatients seen at Brest (West of France) University Hospital for symptomatic VTE (deep vein thrombosis (DVT) and/or pulmonary embolism (PE)) were asked to participate in the study. Diagnosis of DVT was confirmed by the absence of full compressibility of a proximal or distal vein of the deep lower limb on compression ultrasonography (CUS). PE Diagnosis was done by a segmental or larger artery filling defect on chest computed tomography, or the combination of high pre-test clinical probability of PE with high probability ventilation–perfusion lung scan according to the PIOPED criteria, or proximal DVT on CUS in a patient with suspected PE. Controls were selected from the roster of patients hospitalized in the same ward in the following 12 months after the case's event date. Controls were not included if they had a past history of VTE or if they were receiving long-term anticoagulant therapy. DNA were available for 1,141 patients and 1,152 controls.

FARIVE - The FARIVE study is a multicenter case-control study of 607 patients with a first episode of proximal DVT and/or PE [4]. DVT was diagnosed by venography and CUS while PE was diagnosed by spiral computed tomography,high probability ventilation-perfusion lung scan and pulmonary angiography. Patients younger than 18 years, with previous VT event, that had a diagnosis of active cancer or a history of malignancy less than 5 years previously, or have a short life expectancy because of other causes, were excluded. The control group consists of age- and sex-matched individuals free of venous and arterial thrombotic disease. Potential control subjects with cancer, liver or kidney failure, or a history of venous and/or arterial thrombotic disease are ineligible. DNA were available for 698 cases and 672 controls.

MARTHA12 - The MARTHA12 study is composed of an independent sample of 1,245 VTE patients and 801 French healthy individuals. Patients have been recruited between 2010 and 2012 at the Thrombophilia center of La Timone hospital (Marseille, France) as part of the MARTHA study [5]. All patients had a history of a first VTE event documented by venography, Doppler ultrasound, angiography and/or ventilation/perfusion lung scan. They were all free of any chronic conditions and free of any well characterized genetic risk factors including anti-thrombin, protein C or protein S deficiency, homozygosity for FV Leiden or FII 20210A, and lupus anticoagulant. The control group was composed of two subsamples of healthy individuals: one consists of 475 healthy subjects recruited from the Marseille area, the second of 326 healthy heterozygous carriers of the FV Leiden or FII 20210A mutations selected from the national health examination centers of the French Social Security in collaboration with the Hemostasis and Thrombosis Study Group. DNA were available for 1,180 cases and 781 controls.

Single nucleotide polymorphisms genotyping and association testing

Eleven SNPs were selected for genotyping in the EDITH, FARIVE and MARTHA12 samples. These SNPs were identified as the lead SNP (or best proxy for it) at the 11 loci that demonstrated significant association with VTE at $p < 10^{-5}$ in the INVENT meta-GWAS after adjusting for the known- and replicated- VTE associated SNPs [1]. The lead SNPs were *USP34* rs10177303, *CACNA2D3* rs72624844, *FABP6* rs72812220, *AKR1C1* rs114155856, *WWP2* rs151186597, *CPAMD8* rs10421800, *KRI1* rs6511703, *PPP5C* rs3795043, *LILRP2* rs8107460, *ADRA1D* rs78265041 and *ICOSLG* rs11701174 (Table 1). Genotyping was performed by taqman and sequenom technologies at the Centre National de Génotypage (CNG, Evry, France).

Hardy-Weinberg equilibrium was checked by means of a Chi-square test in case and control subjects from each study separately. Association of genotyped SNPs with VTE risk was assessed in each of the three studies separately by use of a logistic regression model

adjusted for age and sex. After having checked for the homogeneity of the associations across these studies using the Cochran's Q statistic, results were then meta-analyzed using a fixed-effects model based on the inverse-variance weighting method. Because the aim of the project is to validate the candidate associations generated in the INVENT cohorts, one-sided association tests were conducted.

Results and Discussion

The genotype distributions of all SNPs but one, the *KRI1* rs8113381, were compatible with the Hardy-Weinberg equilibirum (HWE). The distribution of the *KRI1* rs8113381 that served as a proxy for the rs6511703 significantly ($p < 10^{-3}$) deviating from HWE in all studies, both in cases and controls, this SNP was then excluded from the association analysis. The minimum per-study genotyping success rate over the 10 remaining SNPs was 0.976. Results of the association tests were detailed in Table 2. None of the tested SNPs was statistically associated with VTE in the combined samples, with the smallest p-value being p = 0.027 for the *AKR1C* rs115641535. Odds ratios for all SNPs were close to unity or in the opposite direction compared to the INVENT findings.

We also investigated whether these associations could be enhanced in specific subgroups of patients, and for this end patients were stratified according to the clinical manifestations of VTE (deep vein thrombosis or pulmonary embolism) or to the presence of strong genetic risk factors (*F2* rs1799963 or *F5* rs6025 mutations). Also in these analyses we did not observe evidence for preferential association of the tested SNPs in specific strata.

Despite an average power of ~80% to detect in the three case-control populations studied an odds ratio equal to or greater than the INVENT findings (Table 2), we were not able to replicate any of these associations. The lack of validation of the 11 SNPs does not preclude the existence of common variants with weaker effects or with similar effect size but in no linkage disequilibrium with SNPs assessed by imputation techniques, of common variants exercing haplotype or more complex effects, or of rare variants associated with much stronger genetics effects.

Our findings suggest that none of the 11 common SNPs we selected and that demonstrate suggestive statistical evidence at $p < 10^{-5}$ (which corresponds to allelic odds ratio greater than 1.12) in the INVENT GWAS tag for VTE-associated SNPs. Additional statistical power and possibly denser genotyping are needed to identify new associations with incident VTE.

Addendum

A. Delluc, A. Roche, M. De Andrade, W. Tang, D. Chasman, A. van Hylckama Vlieg, P.Reitsma, C. Kabrhel, N. Smith, D-A. Trégouët and P-E. Morange participated to data collection and study design. C. Derbois and R. Olaso organized the genotyping analyses. D-A Trégouët and M. Germain performed statistical data analysis. D-A. Trégouët drafted the document that were further reviewed by A. Delluc, M. De Andrade, W. Tang, D.Chasman, P. Reitsma, C. Kabrhel, A van Hylckama Vlieg, N.Smith and P-E Morange.

Acknowledgments:

This work was conducted as part of the INVENT consortium whose contributors list is given in the Appendix. This study recieved grant support for NHLBI R01HL116854 and NHLBI 1R01HL116854. This work was partially supported by the GenMed LABEX (ANR-10-LABX-0013), the French Clinical Research Infrastructure Network on Venous Thrombo-Embolism (F-CRIN INNOVTE) and the ICAN Institute for Cardiometabolism and Nutrition (ANR-10-IAHU-05). We thank Cécile Dulary and Fernando Dos Santos for their technical support in genotyping MARTHA12, FARIVE and EDITH samples.

References

1 Germain M, Chasman DI, de Haan H, Tang W, Lindström S, Weng LC, de Andrade M, de Visser MC, Wiggins KL, Suchon P, Saut N, Smadja DM, Le Gal G, van Hylckama Vlieg A, Di Narzo A, Hao K, Nelson CP, Rocanin-Arjo A, Folkersen L, Monajemi R, Rose LM, et al. Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet.* 2015;**96**:532-42

2 Hinds DA, Buil A, Ziemek D, Martinez-Perez A, Malik R, Folkersen L, Germain M, Mälarstig A, Brown A, Soria JM, Dichgans M, Bing N, Franco-Cereceda A, Souto JC, Dermitzakis ET, Hamsten A, Worrall BB, Tung JY; METASTROKE Consortium, INVENT Consortium, Sabater-Lleal M. Genome-wide association analysis of self-reported events in 6135 individuals and 252 827 controls identifies 8 loci associated with thrombosis. *Hum Mol Genet.* 2016 Feb 9. pii: ddw037. [Epub ahead of print].

3 Oger E, Lacut k, Le Gal G, Couturaud F, Guénet D, Abalain JH, Roguedas AM, Mottier D, EDITH COLLABORATIVE STUDY GROUP. 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.* 2016; **4**: 793–799

4 Trégouët DA, Heath S, Saut N, Biron-Andreani C, Schved JF, Pernod G, Galan P, Drouet L, Zelenika D, Juhan-Vague I, Alessi MC, Tiret L, Lathrop M, Emmerich J, Morange PE. Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. *Blood* 2009; **113,** 5298–5303

5 Oudot-Mellakh T, Cohen W, Germain M, Saut N, Kallel C, Zelenika D, Lathrop M, Trégouët DA, Morange PE. Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol.* 2012;**157**:230-239

Table 1 Selected single nucleotide polymorphisms that demonstrated suggestive statistical evidence at $p < 10^{-5}$ for association with VTE in the INVENT meta-analysis.

					Discovery Meta-Analysis INVENT consortium Ncases = 7,507 Ncontrols=52,632		
Chr	Locus	Lead SNP	Alleles	MAF	OR[1]	P[2]	P[3]
2	USP34	rs10177303	G/A	0.37	0.889 [0.845 - 0.934]	$p = 4.44\ 10^{-6}$	p = 0.65
3	CACNA2D3	rs72624844	A/G	0.04	1.410 [1.239 - 1.605]	$p = 2.00\ 10^{-7}$	p = 0.21
5	FABP6	rs72812220	G/A	0.05	1.327 [1.173 - 1.500]	$p = 1.28\ 10^{-6}$	p = 0.86
10	AKR1C1	rs114155856	C/G	0.02	1.489 [1.258 - 1.763]	$p = 3.75\ 10^{-6}$	p = 0.10
16	WWP2	rs151186597	T/C	0.06	0.736 [0.643 - 0.841]	$p = 6.95\ 10^{-6}$	p = 0.98
19	CPAMD8	rs10421800	G/A	0.05	1.315 [1.182 - 1.465]	$p = 5.70\ 10^{-7}$	p = 0.07
19	KRI1	rs6511703	T/C	0.23	0.866 [0.816 - 0.919]	$p = 2.13\ 10^{-6}$	p = 0.71
19	PPP5C	rs3795043	G/A	0.24	0.869 [0.817 - 0.920]	$p = 2.83\ 10^{-6}$	p = 0.52
19	LILRP2	rs8107460	C/T	0.13	1.206 [1.116 1.302]	$p = 1.98\ 10^{-6}$	p = 0.47
20	ADRA1D	rs78265041	C/T	0.03	1.567 [1.297 - 1.893]	$p = 3.33\ 10^{-6}$	p = 0.11
21	ICOSLG	rs11701174	C/T	0.28	0.877 [0.829 - 0.929]	$p = 7.98\ 10^{-6}$	p = 0.34

[1] Allelic Odds Ratio (OR) [95%Confidence Interval] associated with the minor allele.

[2] Association p-value resulting from the meta-analysis of INVENT GWAS conditioning on known VTE associated SNPs

[3] p-value of the test for homogeneity of association (Cochran 's Q statistic) across the twelve discovery cohorts of the INVENT consortium.

Table 2 Association of selected candidate SNPs with VTE in the MARTHA12, FARIVE and EDITH replication studies

	USP34 rs2694642[1]	CACNA2D3 rs72624844	AKR1C1 rs115641535[1]	FABP6 rs72812220	WWP2 rs80215330 [1]	CPAMD8 rs10421800	PPP5C rs3795043	LILRP2 rs7249176 [1]	ADRA1D rs78265041	ICOSLG rs4818892 [1]
MARTHA12										
Controls (N = 781)	0.385	0.064	0.028	0.042	0.052	0.056	0.245	0.121	0.023	0.341
Cases (N = 1180)	0.344	0.064	0.022	0.030	0.069	0.063	0.254	0.111	0.016	0.317
Allelic OR [95%CI] [2]	0.84 [0.74-0.96]	1.02 [0.79-1.32]	0.78 [0.52-1.17]	0.70 [0.49-1.00]	1.32 [1.00-1.73]	1.13 [0.86-1.48]	1.05 [0.90-1.22]	0.91 [0.75-1.12]	0.71 [0.44-1.13]	0.89 [0.77-1.02]
FARIVE										
Controls (N = 672)	0.326	0.063	0.034	0.047	0.079	0.065	0.243	0.121	0.017	0.321
Cases (N = 698)	0.355	0.062	0.022	0.035	0.065	0.052	0.253	0.140	0.018	0.315
Allelic OR [95%CI] [2]	1.14 [0.97-1.33]	0.98 [0.72-1.33]	0.63 [0.39-1.01]	0.72 [0.49-1.07]	0.80 [0.60-1.07]	0.78 [0.57-1.08]	1.06 [0.89-1.26]	1.19 [0.96-1.50]	1.05 [0.59-1.88]	0.97 [0.82-1.14]
EDITH										
Controls (N = 1152)	0.392	0.056	0.026	0.048	0.056	0.056	0.249	0.111	0.019	0.289
Cases (N = 1141)	0.384	0.047	0.022	0.060	0.054	0.057	0.239	0.101	0.021	0.294
Allelic OR [95%CI] [2]	0.97 [0.86-1.09]	0.82 [0.63-1.07]	0.85 [0.58-1.25]	1.29 [0.99-1.69]	0.96 [0.75-1.25]	1.01 [0.79-1.29]	0.95 [0.83-1.09]	0.89 [0.74-1.08]	0.75[0.25-2.19]	1.03 [0.90-1.17]
All										
Allelic OR [95%CI] [3]	0.96[0.89-1.04]	0.94[0.79-1.09]	0.77[0.60-0.97]	0.95[0.79-1.15]	1.01[0.87-1.19]	0.98[0.84-1.15]	1.00[0.93-1.10]	0.91[0.80-1.04]	0.82[0.58-1.16]	0.96[0.89-1.04]
P [4]	0.292	0.405	0.027	0.626	0.866	0.849	0.824	0.178	0.267	0.351
P [5]	0.0165	0.493	0.609	0.007	0.047	0.227	0.497	0.866	0.570	0.314
Post-Hoc power[6]	67%	91%	83%	76%	94%	80%	71%	82%	98%	70%

[1] rs2694642, rs115641535, rs80215330, rs7249176 and rs4818892 were used as proxies for rs10177303, rs114155856, rs151186597, rs8107460 and rs11701174, respectively. Data for the KRI1 rs8113381 that served as a proxy for the lead rs6511703 were not shown as its distribution did not follow Hardy-Weinberg equilibrium.

[2] Allelic Odds Ratio (OR) [95%Confidence Interval] associated with the minor allele derived from a logistic regression analysis adjusted for age and sex.

[3] Meta-analyzed adjusted allelic OR derived from a standard fixed-effect meta-analysis of the results observed in the three replication studies.

[4] P-value of the combined allelic OR derived from a standard fixed-effect meta-analysis of the results observed in the three replication studies.

[5] P-value of the test for homogeneity of associations (Cochran 's Q statistic) across the 3 studies.

[6] Power of the combined replication studies to replicate, at the 0.05/10 significance level, the allelic effect observed in the INVENT discovery cohorts (one-sided test).