Genetic variation and susceptibility to venous thrombosis : Etiology and risk assessment

Bezemer, I.D.

Genetic Variation and Susceptibility to Venous Trombosis

Etiology And Risk Assessment

Irene D. Bezemer

Genetic Variation and Susceptibility to Venous Trombosis

Etiology And Risk Assessment

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 2 juni 2009
klokke 16.15 uur
door

Irene Dorothee Bezemer

geboren te Alphen aan den Rijn
in 1979

ISBN 978-90-9024216-3

PROMOTIECOMMISSIE

Promotor:	Prof. Dr. F.R. Rosendaal
Copromotor:	Dr. C.J.M. Doggen
Referent:	Prof. Dr. H.G. Brunner (Radboud Universiteit Nijmegen)
Overige leden:	Prof. Dr. H.R. Büller (Universiteit van Amsterdam) Prof. Dr. P.H. Reitsma

The work described in this thesis was performed at the department of Clinical Epidemiology of the Leiden University Medical Center in Leiden, the Netherlands.
The work was supported by grants from the Netherlands Heart Foundation (grants 89.063 and 98.113), the Dutch Cancer Foundation (grant RUL 99/1992), the Netherlands Organisation for Scientific Research (grant 912-03-033| 2003) and the Leducq Foundation, Paris, France for the development of Transatlantic Networks of Excellence in Cardiovascular Research (grant 04 CVD 02).
Financial support for the printing of this thesis by Celera Corporation and the J.E. Jurriaanse Stichting is gratefully acknowledged.

CONTENTS

Introduction

Chapter 1

INTRODUCTION

Venous thrombosis is the result of excessive blood coagulation in veins, most frequently in the deep veins of the leg. The thrombus impairs or obstructs the blood flow which leads to swelling of the affected limb, pain and red discoloration. Besides painful and disabling, venous thrombosis can be life-threatening when part of the thrombus breaks off. Via the inferior caval vein the thrombus travels to the lungs where it may obstruct branches of the lung artery; a pulmonary embolism. Venous thrombosis occurs in about one in every 1000 persons per year [1]. About two thirds of patients have deep vein thrombosis of the leg and one third has pulmonary embolism with or without concurrent deep vein thrombosis of the leg [1-3].

The traditional model for classification of the pathology of venous thrombosis is known as 'Virchow's triad', after the work by 19th century pathologist Rudolf Virchow [4,5]. According to the triad, venous thrombosis occurs as a result of (1) alterations in blood flow, (2) endothelial injury, or (3) alterations in blood constitution. Another classification of causes is often made into (1) genetic causes and (2) acquired (or environmental) causes: some individuals have a strong genetic predisposition to develop venous thrombosis while others experience venous thrombosis only after environmental triggers such as long-term immobilization, oral contraceptive use or advanced age [6]. In most cases, venous thrombosis presumably occurs after an environmental trigger on a background of increased susceptibility. Figure 1 (from [6]) illustrates how these risk factors might relate to the occurrence of venous thrombosis in an individual. The factor V Leiden mutation here represents genetic predisposition, which remains constant during life. Depending on the genetic make-up of an individual, the "genetic predisposition level" will be higher or lower. Combined with increasing risk of venous thrombosis with advancing age and risk due to the presence of possible environmental triggers, an individual's "thrombosis potential" might exceed the threshold for developing venous thrombosis.

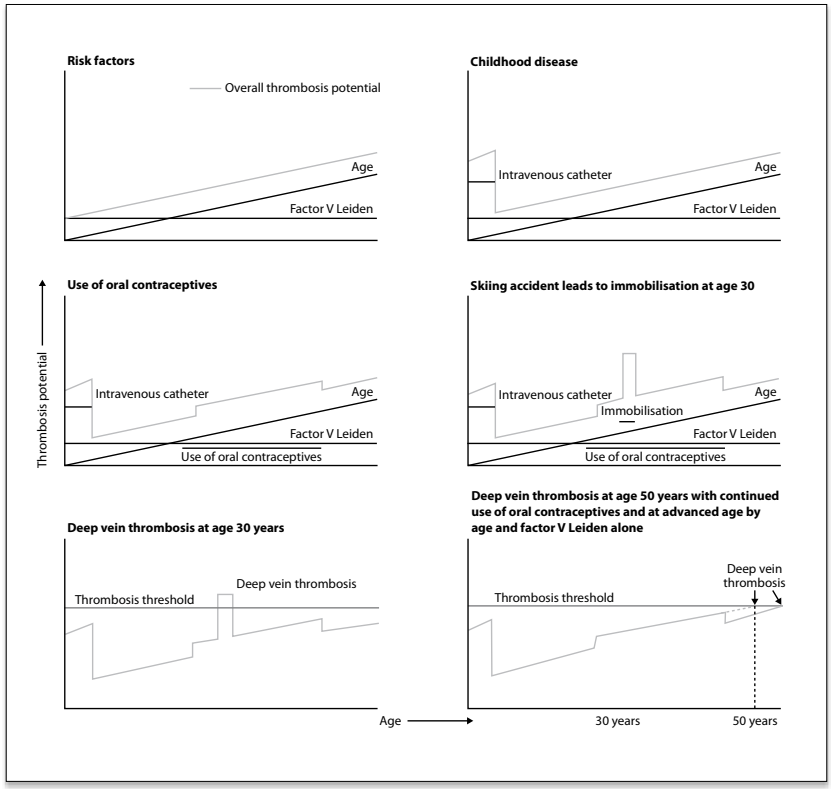

Figure 1: Models of thrombosis risk. In each panel, the figure shows the thrombosis (black) potential of each risk factor present during an individual's life and the resultant thrombosis potential (red). (Reprinted from The Lancet 6)

Genetic risk factors are generally classified as 'alterations in blood constitution', although they could also involve the vessel wall or stasis. A range of pro- and anticoagulant proteins act in the process that leads to thrombus formation; genetic alterations in any of these proteins may alter their level or function. For example, deficiencies of the anticoagulant proteins antithrombin, protein C and protein S increase the risk of venous thrombosis, as do increased levels of the procoagulant proteins fibrinogen, prothrombin, factor VIII, factor IX and factor XI [7].

The search for genes involved in common diseases was facilitated by the rapid evolution of the field of genetics in the last decade. With the completion of the Human Genome Project in 2003, a reference sequence of the 3 billion base pairs in the human genome has become publicly available (www.ornl.gov/sci/techresources/Human_Genome/home.shtml). In 2002, The International HapMap Project was initiated with the goal to make a catalogue of genetic variation between individuals [8]. The HapMap database, together with advances in genotyping technology, made it possible to perform genome-wide association studies for common diseases like type 2 diabetes [9-11] and breast cancer [12-14]. A genome-wide association study measures genetic variants throughout the genome and links them to the occurrence of disease [15].

A main source of genetic variation between individuals are single nucleotide polymorphisms (SNPs). In a SNP, two (for some SNPs three) allele options exist at one nucleotide position, for example an A-allele or a G-allele. The allele that is most prevalent in a population is called the major allele; the other is the minor allele. About 10 million "common" SNPs (minor allele frequency ≥0.01) exist in the human genome [16]; on average one in every 300 bases is a common SNP. SNPs that are located on the same chromosome are usually inherited together. This association between adjacent SNPs is known as linkage disequilibrium (LD), as opposed to the random combination (equilibrium) of different chromosomes during meiosis. Because of LD, it is possible to statistically predict the status of one SNP by genotyping a nearby SNP. SNPs that are genotyped to predict the status of other SNPs are known as "tag SNPs". However, even for SNPs on the same chromosome, loss of LD occurs. During meiosis, two homologous chromosomes often exchange sections. This may result in recombination of DNA stretches of the two chromosomes, which decreases the degree of linkage between the nucleotides in a population. The larger the distance between two nucleotides, the higher the likelihood of recombination between the two positions and the lower the degree of LD. The principle of LD is the rationale for haplotype analysis, and the construction of the HapMap and similar databases. Knowledge of LD patterns in a population allows to efficiently capture variation in a genetic

region by genotyping tag SNPs, instead of genotyping all known variants. These tag SNPs can be studied for association with disease.

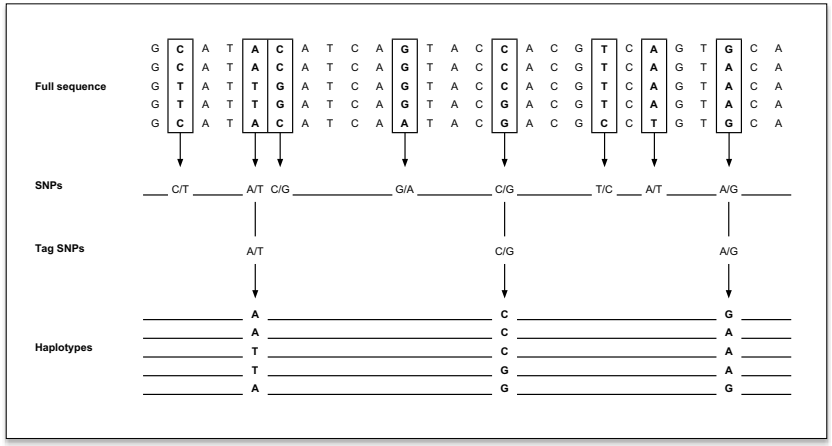

Figure 2: From full nucleotide sequence to tag SNPs and haplotypes. From 5 individuals the full sequence of a stretch of DNA (e.g. part of a gene) is shown. Most of the DNA sequence is identical between these individuals, but some positions vary: the SNPs. For some of these SNPs the allele is predicted from other SNPs because of LD. For example, the three SNPs on the left only occur in combinations CAC or TTG. With the allele of one SNP the other two are known as well. SNPs 4, 6 and 7 are also in perfect LD; their alleles are predicted from combining SNPs 5 and 8 (GG for the minor alleles, any other combination of SNP 5 and 8 for the major allele). A haplotype is the specific sequence of a DNA stretch and is identified by the tag SNPs. In the figure, each individual carries another haplotype.

STUDY POPULATIONS

Two case-control studies with similar design were used for the analysis of SNPs and the risk of venous thrombosis: the Leiden Thrombophilia Study and the Multiple Environmental and Genetic Assessment of risk factors for venous thrombosis. The Northwick Park Heart Study-II was used for the analysis of the *F9* Malmö SNP and activation of coagulation factor IX in chapter 6.

Leiden Thrombophilia Study (LETS)

The LETS was designed to identify genetic risk factors for venous thrombosis. Four hundred seventy four patients with a first venous thrombosis of the leg between January 1988 and December 1992 were recruited from the anticoagulation clinics of Leiden, Amsterdam and Rotterdam. Patients with malignant disorders, known to strongly increase the risk of venous thrombosis, were excluded. For each patient an unrelated control subject was recruited, matched on age and sex and having no history of venous thrombosis and no malignancy. Participants filled in a standard questionnaire on potential risk factors for venous thrombosis, and a blood sample was taken at least three months after discontinuation of anticoagulant therapy. Details of the LETS have been described previously [17,18].

Multiple Environmental en Genetic Assessment of risk factors for venous thrombosis (MEGA study)

The MEGA study was set up a decade after the LETS, and aimed to study combinations of genetic and environmental risk factors. The MEGA study included 4930 patients with a first venous thrombosis of the leg or arm or pulmonary embolism who attended the anticoagulation clinics of Leiden, Amsterdam, Rotterdam, Den Haag, Utrecht or Amersfoort between March 1999 and August 2004. Patients were asked to bring their partners as control subject. In addition, a population-based control group, frequency-matched to the patients on age and sex, was recruited by random digit dialing between January 2002 and December 2004. Overall 6287 control subjects were included in the MEGA study. Participants filled in a standard questionnaire on potential risk factors for venous thrombosis. A blood sample was taken at least three months after discontinuation of anticoagulant therapy from patients who were diagnosed before June 1, 2002, and their partners, and from the random population controls. Participants who refused to or were unable to provide a blood sample were offered the option of providing a buccal swab sample. Patients who were diagnosed from June 1, 2002, onwards and their partners received a cotton swab along with their questionnaire for collecting buccal cells. Details of the MEGA study have been described previously [19,20].

Northwick Park Heart Study-II (NPHS-II)
The NPHS-II is a cohort study of middle-aged men and was set up to study the association between coagulation and coronary heart disease. In total 2951 men aged 50-61 years registered with nine general medical practices in England and Scotland were included in the study. Exclusion criteria were a history of unstable angina or myocardial infarction; regular anti-platelet or anticoagulant therapy; cerebrovascular disease; malignancy; conditions exposing staff to risk or precluding informed consent. Details of the NPHS-II have been described previously [21].

OUTLINE OF THIS THESIS

The aim of the research presented in this thesis is to identify and evaluate common genetic variants that contribute to genetic susceptibility to venous thrombosis. We also studied the clinical importance of genetic variants in prediction of venous thrombosis.

Chapter 2 addresses the question whether the classical parameter to determine genetic predisposition, family history, remains of use now genetic predisposition can be measured at the molecular level.
In **Chapter 3** the recent history of genetic research in venous thrombosis is summarized; it describes genetic variants that were identified in the last decade and aims to put the findings in perspective. One of the genetic variants of which its involvement in the etiology of venous thrombosis has been long debated is the MTHFR 677 C>T SNP. **Chapter 4** describes the association between this SNP and venous thrombosis in the MEGA study.
Chapter 5 reports of a large-scale SNP association study in the LETS and MEGA studies. The study aimed to identify new genetic variants that contribute to the risk of venous thrombosis. Because the study design does not address the question of whether a single SNP association is causal, the chapter zooms in on the most strongly associated SNP in the *CYP4V2* gene. In the region of this SNP, fine-scale genotyping was performed in order

to determine whether the association was due to linkage to another, more strongly associated and potentially causal SNP. The study also aimed to gain insight into the mechanistic aspect of the association. On one of the other SNPs from the SNP association study, we reported in more detail in **Chapter 6**. The SNP in the *F9* gene, also known as *F9* Malmö, was studied for linkage with nearby SNPs, and for association with coagulation factor IX levels (LETS and MEGA) and factor IX activation (NPHS-II).
Finally, in **Chapter 7**, we incorporated knowledge from previous and present research into a model that aims to predict the risk of venous thrombosis based on associated SNPs. We aimed to explore to what extent we can predict venous thrombosis using the currently known thrombosis-associated SNPs.

The Value Of Family History As A Risk Indicator For Venous Thrombosis

Chapter 2

ID Bezemer
FJM van der Meer
HCJ Eikenboom
FR Rosendaal
CJM Doggen

Archives of Internal Medicine 2009;169(6):610-615

ABSTRACT

Background A positive family history of venous thrombosis may reflect the presence of genetic risk factors. Once a risk factor has been identified, it is not known whether family history is of additional value in predicting an individual's risk. We studied the contribution of family history to risk of venous thrombosis conditional on known risk factors.

Methods In the MEGA Study, a population-based case-control study, we collected blood samples and information about family history and environmental triggers from 1605 patients with a first venous thrombosis and 2159 control subjects.

Results 505 (31%) Patients and 373 (17%) control subjects reported having one or more affected first-degree relatives. A positive family history increased the risk of venous thrombosis more than twofold (odds ratio [95% confidence interval], 2.2 [1.9-2.6]) and up to fourfold (3.9 [2.7-5.7]) when more than one relative was affected. Family history corresponded poorly with known genetic risk factors. Both in those with and without genetic or environmental risk factors, family history remained a risk indicator. The risk increased with the number of risk factors identified; for those with a genetic and environmental risk factor and a positive family history, the risk was about 64-fold the risk of those with no known risk factor and a negative family history.

Conclusions Family history is a risk indicator for a first venous thrombosis, regardless of the risk factors identified. In clinical practice the family history may be more useful for risk assessment than thrombophilia testing.

INTRODUCTION

A positive family history of venous thrombosis may reflect the presence of genetic risk factors in a family. Carriers of a genetic risk factor are at increased risk of a first venous thrombosis, particularly when exposed to environmental triggers. Factor V Leiden, for example, synergistically increases the risk of venous thrombosis in oral contraceptive users [22]. Since universal screening is not cost-effective [23,24], research efforts are focused on selection criteria that may be used to increase the chance of finding a genetic risk factor. Family history is an evident candidate.

Several authors have studied the value of family history as a surrogate of known genetic risk factors for venous thrombosis [25-29]. These studies have shown that the family history cannot be used to identify genetic risk factors because both positive predictive value and sensitivity are low.

Few have studied the association between family history and venous thrombosis [30,31]. In addition, it is not known whether family history is of additional value in predicting an individual's risk of venous thrombosis once a genetic risk factor is identified. We therefore estimated the relative risk of venous thrombosis when the family history is positive and studied the contribution of family history to risk in strata of known risk factors. Family history was evaluated in patients with venous thrombosis and control subjects from the Multiple Environmental and Genetic Assessment of risk factors for venous thrombosis (MEGA study), a large population-based case-control study.

METHODS

Study Population and Data Collection

Recruitment, data collection and ascertainment of venous thrombosis events in the MEGA study were described in detail previously [20,32]. Patients had experienced a first deep vein thrombosis of the leg or pulmonary embolism between March 1, 1999 and August 31, 2004. Control subjects were partners of patients or

random population control subjects. The random control subjects were recruited by random digit dialing [33] between January 1, 2002 and December 1, 2004, and frequency matched on sex and age to the patient group. All participants completed a questionnaire on risk factors for venous thrombosis and family history. A blood sample was taken three months after discontinuation of vitamin K antagonist therapy from patients who were diagnosed until June 1, 2002 and their partners. Patients with an indication for life-long treatment with vitamin K antagonists were invited for a blood draw one year after the index date. Patients who were diagnosed from June 1, 2002, onwards and their partners received a cotton swab along with their questionnaire for collecting buccal cells; these were not included in the present study. In the random population control group, blood samples were collected throughout the entire study period and after returning the questionnaire. Overall response rates were 83% in the patient group, 82% in the partner control group and 69% in the random population control group.

Family history

Participants were asked whether parents, brothers or sisters had experienced venous thrombosis and, if so, the age at the event. Because partners of patients were recruited as control subjects, offspring was not included in the family history definition. Family history was considered positive if at least one of these first-degree relatives had experienced venous thrombosis. Within this group, participants with a strong indication of genetic predisposition were defined as having at least one first-degree relative affected before the age of 50 years, or having multiple first-degree relatives affected regardless of their age. When none of the first-degree relatives had suffered a venous thrombosis, family history was defined negative. The answer 'I don't know' was also considered negative.

Environmental triggers

Environmental triggers were surgery, injury (any self-reported injury, such as muscle ruptures or sprain), immobilization (plaster cast, extended bed rest at home for at least 4 days, hospitalization) and pregnancy or puerperium within three months prior to the index date, use of oral contraceptives or hormone replacement therapy at the index date and diagnosis of malignancy within five

years before or within six months after the index date. The index date was defined as date of diagnosis for patients and their partners, and date of completing the questionnaire for random controls.

Genetic risk factors

Genetic risk factors were the factor V Leiden mutation, the prothrombin 20210A mutation, low antithrombin levels, low protein C levels and low protein S levels. Since many mutations in the genes encoding antithrombin, protein C and protein S may cause deficiency, protein levels served as a surrogate for genetic defects. A sample was classified as "low" when the protein level was below the reference value calculated in control subjects (geometric mean minus 2 standard deviations). For protein C and protein S, the reference values were calculated excluding vitamin K antagonist users. In addition, we compared protein C levels to factor VII levels, and protein S levels to factor II levels in order to discriminate between "isolated" low protein C or S levels and overall low coagulation factor levels. We calculated the expected protein C level by linear regression of protein C on factor VII, and calculated the observed over expected ratio for protein C [34]. For protein S the observed over expected ratio was calculated by regression on factor II. The observed protein C or S level was classified as "low" when both the absolute value and the observed over expected ratio were below the reference value calculated in control subjects (geometric mean minus 2 standard deviations). Specific reference values of protein C and protein S levels were calculated for vitamin K antagonist users that were included in sensitivity analyses; the ratios to factor VII and factor II are independent of vitamin K antagonist use.

For the present analysis we selected participants who provided complete information about family history and environmental triggers and donated a blood sample. Among 3033 patients who filled in the questionnaire, 2712 (89%) provided information about family history and of 1959 (65%) patients complete information about environmental triggers and a blood sample were available. In the control group, 4317 (88%) of 4887 participants provided information about family history and of 2438 (50%) control subjects complete information about environmental triggers and a blood sample were available.

During pregnancy and oral contraceptive use protein S levels are reduced and cannot be used as an indicator of a genetic defect of protein S. We therefore excluded women who were pregnant (0 participants) or used oral contraceptives (146 patients and 259 control subjects) at the time of the blood draw. We also excluded vitamin K antagonist users (208 patients and 20 control subjects) because protein C and protein S levels cannot be easily interpreted under these circumstances. After these exclusions 1605 patients and 2159 control subjects remained in the analyses.

Laboratory Analysis

Collection and processing of blood samples, subsequent DNA isolation and genotyping of factor V Leiden and the prothrombin 20210A mutation have been described previously [19]. Measurements of antithrombin and protein C levels were performed with a chromogenic assay and factor II and VII level measurements were based on a mechanical clotting time assay. These measurements were performed on a STA-R coagulation analyzer following the instructions of the manufacturer (Diagnostica Stago, Asnières, France). Total protein S levels were measured by an enzyme-linked immunosorbent assay (ELISA, Diagnostica Stago, Asnières, France). The mean intra- and inter-assay coefficients of variation were 1.7% and 2.6%, respectively, for antithrombin, 1.4% and 3.5% for protein C, 2.7% and 4.2% for factor II, 3.4% and 4.0% for factor VII and 5.0% and 3.5% for protein S.

Statistical Analysis

Odds ratios (OR) and 95% confidence intervals (CI) were computed to estimate the relative risk of venous thrombosis associated with a positive family history. Taking the group with a negative family history as reference, ORs were calculated for having any affected first-degree relative (with the exception of offspring), having a first-degree relative affected before the age of 50 years, and having multiple affected first-degree relatives. Adjustment for age (continuous) and sex was performed by logistic regression. Subgroup analyses were performed within strata of known risk factors and within 10-year age categories. We calculated the positive predictive value and sensitivity

of family history to identify genetic risk factors. For the positive predictive value and sensitivity estimates, binomial 95% CIs were calculated using the normal approximation.

RESULTS

Table 1. Distribution of age, sex and individual risk factors among patients with venous thrombosis and control subjects

	Patients (N=1605)	Control Subjects (N=2159)
Median age (5th – 95th percentile)	50 (27-68)	51 (28-67)
Men, N (%)	772 (48%)	1150 (53%)
Type of VTE, N (%)		
DVT	949 (59%)	NA
PE	510 (32%)	NA
DVT & PE	191 (9%)	NA
Environmental Risk Factor, Any, N (%)	1086 (68%)	425 (20%)
Surgery	276 (17%)	63 (3%)
Injury	266 (17%)	141 (7%)
Immobilisation	496 (31%)	136 (6%)
Pregnancy/puerperium*	68 (4%)	21 (1%)
Oral contraceptives / HRT*	456 (29%)	108 (5%)
Malignancy	100 (6%)	48 (2%)
Genetic Risk Factor, Any, N (%)	393 (25%)	243 (11%)
Factor V Leiden mutation	246 (15%)	102 (5%)
Prothrombin 20210A mutation	73 (5%)	37 (2%)
Low antithrombin	39 (2%)	56 (3%)
Low protein C	35 (2%)	23 (1%)
Low protein S	26 (2%)	36 (2%)

NA=not applicable

* The pregnancy and hormone use risk factor groups included women only, but the percentages are of the total study group including men and women.

Median age and distributions of sex and individual risk factors among the 1605 patients and 2159 control subjects are listed in Table 1. Family history of venous thrombosis was positive for 505 (31%) patients and 373 (17%) control subjects (Table 2). The overall OR of a positive relative to a negative family history was 2.2 (95% CI, 1.9-2.6). The association was stronger when only family members with venous thrombosis before the age of 50 years were considered positive (OR, 2.7; 95% CI 2.2-3.4) or when several relatives were affected (OR, 3.9; 95% CI 2.7-5.7). The OR for venous thrombosis when having several relatives affected and at least one of them before the age of 50 was 4.4 (95% CI 2.8-6.9, not shown). The median (25^{th} to 75^{th} percentile) number of relatives, i.e. parents and siblings, that was reported in the questionnaire was 5 (3 to 7) in the patient group and 5 (3 to 6) in the control group.

Table 2. Distribution of first-degree family history among 1605 patients with venous thrombosis and 2159 control subjects

Family history	Patients, N (%)	Control Subjects, N (%)	OR (95% CI)
negative	1100 (69%)	1786 (83%)	1 [Reference]
Positive, any relative	505 (31%)	373 (17%)	2.2 (1.9 -2.6)
Positive, relative < 50	240 (15%)	144 (7%)	2.7 (2.2 -3.4)
Positive, > 1 relative	97 (6%)	40 (2%)	3.9 (2.7 -5.7)

In 150 of 505 (30%) patients with a positive family history a genetic risk factor was identified. A higher number of affected relatives and a younger age at which the relative was affected increased the chance to find a genetic risk factor, up to 36% for patients with several affected relatives (positive predictive value, Table 3). The negative predictive value, i.e. the chance that known genetic risk factors are indeed absent when the family history is negative, was 78%. This indicates that 22% of patients were thrombophilic carriers despite a negative family history. In the control group genetic risk factors were less prevalent than among patients and the positive predictive values were lower. The ROC-curve for any relative affected, which represents

the accuracy of family history to identify genetic risk factors, had an area under the curve of only 54% in patients and 53% in the control group. When we took the presence of a genetic risk factor as the starting point, a positive family history was reported by 38% of patient carriers and by 22% of control carriers (sensitivity, Table 3). Thus, the majority of thrombophilic carriers did not have affected relatives.

Table 3. Family history and prevalence of genetic risk factors in patients and control subjects

Study Group	Family history [a]	Known genetic risk factor [b] yes	no	Predictive value (95% CI)	Sensitivity (95% CI)
Patients	Negative	243	857	78% (75% - 80%)	NA
	positive, any relative	150	355	30% (26% - 34%)	38% (33% - 43%)
	positive, relative < 50	80	160	33% (27% - 39%)	20% (15% - 26%)
	positive, > 1 relative	35	62	36% (27% - 46%)	9% (4% - 14%)
Control Subjects	Negative	190	1596	89% (88% - 91%)	NA
	positive, any relative	53	320	14% (11% - 18%)	22% (17% - 27%)
	positive, relative < 50	19	125	13% (8% - 19%)	8% (2% - 14%)
	positive, > 1 relative	9	31	23% (10% - 35%)	4% (-3% - 11%)

[a] History of venous thrombosis among parents, brothers and sisters

[b] Low protein levels of antithrombin, protein C or protein S, factor V Leiden mutation, prothrombin 20210A mutation

In order to study the value of family history as a risk indicator when known risk factors have been measured, we grouped patients and control subjects according to type of risk factor identified: none, environmental, genetic or both (Table 4). In all strata, patients more frequently reported to have affected relatives than control subjects. So, family history is a risk indicator regardless of the presence of known risk factors.

The relative risk associated with a positive family history was of similar magnitude as the risk associated with a genetic risk factor. In the absence of environmental triggers the ORs were 2.5 for family history and 2.3 for a

genetic risk factor. In the presence of environmental triggers the ORs were 16.4 for family history and 21.2 for a genetic risk factor. The OR increased with the number of risk factors identified; for those with a combination of any genetic and acquired risk factor the risk was about 60-fold the risk of those with no known risk factor and a negative family history.

To rule out that the higher prevalence of positive family histories in patients with genetic risk factors was the result of specific combinations or the number of genetic risk factors, we stratified this group by the specific genetic risk factors. In the group that carried factor V Leiden but no other genetic risk factor (40 patients and 22 control subjects), a positive family history further increased the risk of venous thrombosis; factor V Leiden carriers with a positive family history had a 2.9 fold (95% CI 1.5-5.7) higher risk than factor V Leiden carriers with a negative family history. When an affected relative was younger than 50 years, this OR was 5.4 (95% CI 2.0-14.6) and when at least two relatives were affected 17.8 (95% CI 2.2-143.1). The other strata of specific genetic risk factors included fewer patients and control subjects thereby precluding meaningful analysis.

Genetic risk factors might play the most prominent role at young age, when environmental triggers are less prevalent. We therefore calculated ORs for family history per 10-year age category. Family history was associated with the risk of venous thrombosis in all age groups. The relative risk slightly decreased with age; the ORs (95% CIs) for any relative affected were 3.2 (1.7-6.0) at age 18-29 years, 2.4 (1.6-3.6) at age 30-39 years, 2.1 (1.5-2.8) at age 40-49 years, 2.1 (1.6-2.8) at age 50-59 years and 2.2 (1.6-3.1) at age 60-69 years. Because thrombotic events in a family accumulate during life and the risk of venous thrombosis increases with age, we further studied whether age could have confounded our results. Adjustment for age did not change any of the estimates. We also adjusted for sex to assess the impact of possible associations between oral contraceptive use and family history, but again none of the estimates changed.

Table 4. Family history in strata of known risk factors.

Risk factor	Family history [a]	Patients	Control Subjects	OR (95% CI) [b]	OR (95% CI) [c]
No known risk factor					
	negative	261 (67%)	1286 (84%)	1 [Reference]	1 [Reference]
	positive, any relative	128 (33%)	252 (16%)	2.5 (1.9 - 3.2)	2.5 (1.9 - 3.2)
	positive, < 50 years	53 (14%)	98 (6%)	2.7(1.9 -3.8)	2.7 (1.9 - 3.8)
	positive, multiple relatives	23 (6%)	27 (2%)	4.2 (2.4 - 7.4)	4.2 (2.4 - 7.4)
Environmental factor only [d]					
	negative	596 (72%)	310 (82%)	1 [Reference]	9.5 (7.8 - 11.5)
	positive, any relative	227 (28%)	68 (18%)	1.7 (1.3 - 2.4)	16.4 (12.2 - 22.2)
	positive, < 50 years	107 (13%)	27 (7%)	2.1 (1.3 - 3.2)	19.5 (12.5 - 30.4)
	positive, multiple relatives	39 (5%)	4 (1%)	5.1 (1.8 - 14.3)	48.0 (17.0 - 135.6)
Genetic factor only [e]					
	negative	71 (55%)	150 (77%)	1 [Reference]	2.3 (1.7 - 3.2)
	positive, any relative	59 (45%)	46 (23%)	2.7 (1.7 - 4.4)	6.3 (4.2 - 9.5)
	positive, < 50 years	33 (25%)	15 (8%)	4.6 (2.4 - 9.1)	10.8 (5.8 - 20.2)
	positive, multiple relatives	14 (11%)	6 (3%)	4.9 (1.8 - 13.4)	11.5 (4.4 - 30.2)
Environmental and genetic factor					
	negative	172 (65%)	40 (85%)	1 [Reference]	21.2 (14.7 - 30.6)
	positive, any relative	91 (35%)	7 (15%)	3.0 (1.3 - 7.0)	64.1 (29.4 - 139.8)
	positive, < 50 years	47 (18%)	4 (9%)	2.7 (0.9 - 8.0)	57.9 (20.7 - 162.1)
	positive, multiple relatives	21 (8%)	3 (6%)	1.6 (0.5 - 5.7)	34.5 (10.2 - 116.5)

[a] History of venous thrombosis among parents, brothers and sisters

[b] Odds ratio per stratum of type of risk factors identified

[c] Odds ratio relative to the group with no known risk factor and a negative family history

[d] surgery, injury, immobilisation and pregnancy or puerperium within 3 months preceding the index date, use of oral contraceptives or hormone replacement therapy at the index date and diagnosis of malignancy within 5 years before or within six months after the index date

[e] Low protein levels of antithrombin, protein C or protein S, factor V Leiden mutation, prothrombin 20210A mutation.

Relatives for whom the answer to the question about family history was 'I don't know' were assumed to be negative. Among patients, 238 of 1605 (15%) had at least one relative with unknown venous thrombosis history while other relatives were negative (i.e. family history assumed negative), among controls 307 of 2159 (14%) answered 'I don't know' for at least one relative. Excluding these participants from the analysis led to slightly higher risk estimates for the family history.

All analyses were repeated including vitamin K antagonist users and oral contraceptive users. Including these users influenced the family history distributions by only a few percent.

COMMENT

In a large population-based case-control study we showed that a positive family history increased the risk of venous thrombosis more than twofold, regardless of the risk factors precipitating the thrombosis. A young age of the affected relative and in particular the number of affected relatives more strongly indicated a predisposition to develop venous thrombosis.

Family history and known genetic risk factors were poorly associated, as observed previously [25-27,29,35]. Both the positive predictive value and sensitivity of family history as a test for genetic risk factors were low, with ROC-curves hardly different from a random distribution. The poor predictive value either implies the existence of unknown genetic risk factors or clustering through household effects.

Patients more frequently had a positive family history than control subjects, even when known risk factors were similar. This indicates that an unwknown, probably genetic factor has caused their disease in concert with the risk factor identified. These findings suggest that most genetic risk factors have low penetrance. Only when additional risk factors are present, venous thrombosis

will develop [6,36]. The search for novel genetic risk factors should not be limited to patients without known thrombophilia, since genetic factors that interact with already known genetic risk factors might then not be found. As most carriers of a single genetic risk factor have a negative family history, the sensitivity of family history to identify a single genetic risk factor is low.
We selected low levels of antithrombin, protein C and protein S, the factor V Leiden mutation and the prothrombin 20210A mutation as genetic risk factors. These are clear and frequent genetic risk factors for venous thrombosis. Inclusion of more genetic risk factors will increase the positive predictive value at the cost of the negative predictive value, while sensitivity may remain low. More important is that our study confirms that venous thrombosis is a multi-gene disorder. Family history will be a better surrogate for multiple genetic risk factors, including those yet unknown, than for single defects.
Relatives generally underreport disease in their family [37-41]. We believe that also in our study family history may have been underreported. It does, however, correspond to clinical practice where physicians rely on the family information given by their patient and confirmation of all relatives' disease status is not feasible. Alternatively, we might have overestimated the prevalence of positive family histories because individuals might be more prone to participate in a study when their family history is positive. As selection is most likely in the control group, we might have underestimated the effect of family history.

Antithrombin, protein C and protein S levels were determined from one blood draw. In a clinical setting, low protein levels are confirmed by a second measurement before a patient is diagnosed as deficient. A previous study among patients with venous thrombosis and control subjects [42] reported that 5 of 20 (25%) patients who initially had antithrombin levels below the lower limit of normal were low at a second measurement. Confirmation of low protein C levels occurred in 15 of 22 (68%) patients and confirmation of low protein S levels in 5 of 8 (63%) patients. Confirmation occurred less frequently in control subjects. We acknowledge that the number of individuals with truly low levels of antithrombin, protein C and protein S will be lower than presented here.

We studied whether family history is of additional value in predicting an individual's risk of venous thrombosis once a genetic risk factor has been identified. We could also reverse the question and ask whether genetic testing provides additional prognostic value once the family history has been determined. This could guide decisions on starting oral contraceptive use or taking preventive measures during immobilization. Table 4 shows that environmental risk factors together with a positive family history strongly increase the risk of venous thrombosis. In the absence of a known genetic risk factor the risk is already increased more than 15-fold. Genetic testing to identify additional risk would then not seem useful. Moreover, the positive family history could well reflect unknown genetic risk factors. When the family history is negative, an environmental risk factor would increase the risk about 10-fold to 20-fold, depending on the identification of a genetic risk factor. Given the low chance of finding a genetic risk factor when the family history is negative, genetic testing does not seem to be cost effective in this situation.

It is important to note that the results from the current study apply to the risk of a first venous thrombosis, and may not be applicable to risk of recurrent venous thrombosis. In fact, previous studies have shown that neither genetic risk factors nor family history are predictive for recurrent venous thrombosis [43,44].

We conclude that family history is a risk indicator for a first venous thrombosis, even when a genetic risk factor has been identified. In clinical practice the family history may be more useful for risk assessment than thrombophilia tests. A positive family history represents increased susceptibility on top of the risk due to known genetic and environmental factors. This additional risk is due to unknown or unmeasured risk factors.

ACKNOWLEDGMENTS

Ms. I.D. Bezemer and Drs F.R. Rosendaal and C.J.M. Doggen had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. We thank the directors of the Anticoagulation Clinics of Amersfoort (M.H.H. Kramer, MD), Amsterdam (M. Remkes, MD), The Hague (E. van Meegen, MD), Rotterdam (A.A.H. Kasbergen, MD), and Utrecht (J. de Vries-Goldschmeding, MD) who made the recruitment of patients possible. The interviewers (J.C.M. van den Berg, B. Berbee, S. van der Leden, M. Roosen, and E.C. Willems of Brilman) performed the blood draws. We also thank I. de Jonge, MSc, R. Roelofsen, MSc, M. Streevelaar, L.M.J. Timmers, MSc, and J.J. Schreijer for their secretarial and administrative support and data management. The fellows J.W. Blom, MD, A. van Hylckama Vlieg, PhD, E.R. Pomp, MSc, L.W. Tick, MD, and K.J. van Stralen, MSc took part in every step of the data collection. C.J.M. van Dijk, R. van Eck, J. van der Meijden, P.J. Noordijk, and T. Visser performed the laboratory measurements. H.L. Vos supervised the technical aspects of DNA analysis. We express our gratitude to all individuals who participated in the MEGA study. This research was supported by the Netherlands Heart Foundation (NHS 98.113), the Dutch Cancer Foundation (RUL 99/1992) and the Netherlands Organisation for Scientific Research (912-03-033| 2003). The funding organizations did not play a role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript.

Predictive Genetic Variants For Venous Thrombosis: What's New?

Chapter 3

ID Bezemer
FR Rosendaal

Seminars in Hematology 2007; 44(2): 85-92

ABSTRACT

Various pathways lead to the development of venous thrombosis. Risk factors are common and can be genetic or acquired. Since the identification of factor V Leiden and prothrombin G20210A, the field of genetic epidemiology has developed rapidly and many new genetic variants have been described in the past decade. However, the association with venous thrombosis is often unclear and conflicting results have been reported in various studies. The aim of this review is to describe these candidate predictors of venous thrombosis and to put these in perspective.

Venous thrombosis is a complex condition in which genes and environment both contribute to the risk of disease. Many risk factors for venous thrombosis are common, and often, if not always, the coincidence of two or more risk factors is required to develop thrombosis. Heritable defects in factors that control the hemostatic balance have been identified since 1965 (for a comprehensive review, see Mannucci [45]). In that year, Egeberg described a family in which the incidence of venous thrombosis was higher and at a younger age than expected. It appeared that the affected family members had about 50% lower antithrombin levels than the non-affected family members. Deficiencies of protein C and protein S were identified in a similar manner during the 1980s. Families with a history of recurring venous thrombosis but no known hereditary abnormalities were studied to determine the role of plasma protein deficiencies. It appeared that affected family members had severely reduced protein levels of protein C or S. Today, numerous loss-of function mutations have been described in the genes encoding antithrombin, protein C, and protein S that lead to reduced plasma levels. In the heterozygous state these mutations lead to about halfnormal plasma levels. Homozygous mutations, especially in the antithrombin gene, are assumed to be incompatible with life.

In addition to these relatively rare protein deficiencies, two much more common genetic defects were described during the 1990s. Activated protein C (APC) resistance was identified as a risk factor for venous thrombosis in 1993. In 1994, Bertina et al found that factor V was involved in APC resistance. Subsequently, the factor V Leiden mutation was found by searching the factor V gene of APC resistant patients for mutations in APC binding- and cleavage sites. The second common genetic factor, prothrombin 20210 G>A, was identified in 1996 through screening the prothrombin gene for abnormalities among patients with a personal and family history of venous thrombosis. At present, the three plasma protein deficiencies and the two mutations are the main genetic risk factors for venous thrombosis. However, they still explain only part of venous thrombotic events [45]. The failure to identify a risk factor in many patients and the belief that genetic factors play

an important role in the development of venous thrombosis stimulate the search for novel predictive genetic variants.

Since the identification of prothrombin 20210 G>A in 1996, the field of genetic epidemiology has evolved rapidly and many genetic variants have been described that might influence the risk of venous thrombosis. In this review we will give an overview of the main genetic factors described in the past decade and put the findings in perspective. The main features are presented in Table 1.

NOVEL GENETIC VARIANTS

Fibrinogen

Fibrinogen is the precursor of fibrin, the fundamental constituent of the thrombus. The fibrinogen molecule consists of three chains: alpha (FGA), beta (FGB), and gamma (FGG), each encoded by a separate gene. High levels of fibrinogen increase the risk of venous thrombosis, mainly in elderly people [7]. The most frequently studied polymorphism is a G>A substitution at nucleotide -455 in the *FGB* gene. This genetic variant is associated with slightly increased fibrinogen levels but appears not to increase the risk of venous thrombosis [46,47]. Another, less frequently studied, polymorphism in the *FGA* gene, 4266 A>G (Thr312Ala), is associated with the risk of pulmonary embolism (PE) and is postulated to influence fibrin cross-linking [48]. The 312Ala genotype leads to reduced clot strength and a higher risk of embolization. Ko et al found an FGA haplotype, covering the Thr312Ala variant, to be related to both PE and deep vein thrombosis (DVT) in a Taiwanese population [49]. In another study that analyzed the association of haplotypes of the alpha, beta, and gamma genes and the risk of DVT, the only haplotype associated with DVT was a haplotype of the *FGG* gene, tagged by 10034 C>T (*FGG*-H2) [50] This polymorphism is located in a consensus sequence that is involved in cleavage and splicing of the *FGG* pre-mRNA. The *FGG*-H2 haplotype is associated with decreased levels of fibrinogen gamma' (FGG') and decreased FGG'/

total fibrinogen ratios. Interestingly, the *FGA* Thr312Ala polymorphism is strongly linked to this haplotype. Whether FGA Thr312Ala is a functional variant itself or only reflects the effect of the 10034C>T variation in *FGG*-H2 remains to be elucidated.

Prothrombin

Prothrombin is the inactive precursor of thrombin. A principal role of thrombin is to cleave fibrinogen to form fibrin. In addition, thrombin gives positive feedback to the coagulation cascade by activating other coagulation factors and negative feedback by activating protein C in a complex with thrombomodulin. It influences fibrinolysis also, through thrombin activatable fibrinolysis inhibitor (TAFI). The 20210 G>A mutation in the prothrombin gene increases the risk of venous thrombosis by increasing plasma prothrombin levels (first described by Poort et al [51]. Around the 20210 position several additional variants were described, but there is no clear evidence that these rare variants contribute to the risk of venous thrombosis [52]. Another common variant in the prothrombin gene, 19911 A>G, is also associated with slightly higher prothrombin levels [3,53,54]. Initially, two studies reported that 19911 A>G modulates the risk in 20210A carriers [53,55]. Recently, two larger studies reported an increased risk associated with 19911 A>G, independently of other genetic risk factors [3,54].

Factor V

Activated factor V (FVa) is a cofactor for factor Xa in the conversion of prothrombin to thrombin. In addition, FV is a cofactor of APC-mediated FVIII degradation. Factor V Leiden (1691 G>A, Arg506Gln) is a mutation in the major cleavage site of FVa by APC, which makes FVa more resistant to inactivation. Recently, all known *FV* missense mutations that are relatively common were reviewed by Vos [68]. The most important variant, apart from factor V Leiden, is a common haplotype including several polymorphisms throughout the *FV* gene, first described by Lunghi et al [69]. This haplotype, *FV* HR2, is associated with decreased cofactor activity in FVIII inactivation and a more procoagulant isoform of FV. The variant responsible for the *FV*

Table 1 Candidate Predictors of Venous Thrombosis

Gene	Nucleotide	Amino Acid	Allele Frequency*	Phenotype	Odds Ratio†
Procoagulant proteins					
FGA	4266 A>G	Thr312 Ala	0,26 (G)	reduced clot strength	1,8 [49]
FGG	10034 C>T (FGG-H2)		0,26 (T)	alternatively spliced protein	2,4 [50]
Prothrombin	19911 A>G		0,49 (G) [3]	higher protein level	1,4 [3]
FV	6755 A>G (HR2)	Asp2194Gly	0,07 (G)	decreased FV cofactor activity	1,2 [56] (carriers)
FVIII	94901 C>G	Asp1241Glu	0,15 (G)	lower protein level	0,6 [57] (carriers)
FVIII	HT1		0,14 (HT1) [58]	lower protein level	0,4 [58] (men)
FXII	46 C>T		0,15 (T)	lower protein level	not replicated
FXIII	G>T	Val34Leu	0,24 (T) [59]	higher protein activity	0,6 [59]
FXIII	8259 A>G	His95Arg	0,08 (T) [60]	higher protein activity	1,5 [60]
TF	1208 I/D		0,52 (D) [61]	lower protein level	0,7 [61]
FSAP	1601 G>A (Gly511Gly)		-	impaired fibrinolysis inhibition	not replicated
ACE	I/D		0,51 (D) [62]	higher protein level	not clear
Anticoagulant proteins					
TFPI	536 C>T	Pro151Leu	<0,01 (D) [63]		not replicated
TFPI	-33 T>C		0,36 (C)	higher protein level	0,6 [64]
EPCR	4600 A>G	Ser219Gly	0,08 (G)	higher sEPCR level	not clear
EPCR	4678 G>C		0,34 (C)	higher APC level	not clear
Antifibrinolytic proteins					
PAI-1	4G/5G (D/I)		0,48 (5G) [65]	lower protein level	not clear
TAFI	505 G>A	Ala147Thr	0,30 (T)	higher protein level	0,7
Other					
Blood group	O		0,43 (O) [66]	lower FVIII level	0,6 [66]
ZPI	728 C>T	Arg67Stop	<0,01(T)	lower protein level	3,3 [67] (carriers)
MTHFR	677 C>T	Ala222Val	0,24 (T)	lower protein activity	not clear

* Allele frequencies as calculated in the study population referred to. When no reference is given, the allele frequency for Caucasian populations (applies to almost all studies) was obtained from dbSNP.

† By default, the odds ratio for homozygous carriers. For FV HR2, FVIII Asp1241Glu, FXIII His95Arg, and ZPI Arg69Stop, the odds ratio for all carriers (heterozygous and homozygous) is given.

HR2 phenotype is probably 6755 A>G (Asp2194Gly) [68]. Whether the *FV* HR2 haplotype affects the risk of venous thrombosis is not clear. Individual studies reported conflicting results and in a meta-analysis a pooled odds ratio of 1.15 (95% confidence interval [CI], 0.98 to 1.36) was calculated [56].

Factor VIII

The factor VIII (*FVIII*) gene has been studied extensively because there is a clear association between FVIII levels and the risk of venous thrombosis. Several studies screened cleavage sites, promoter and polyadenylation regions of the *FVIII* gene but found no mutation that corresponded with either FVIII levels or thrombosis [70-72]. However, two other studies reported lower FVIII levels associated with a 94901 C>G (Asp1241Glu) polymorphism that was therefore possibly protective for venous thrombosis [57,73]. In a recent study of the haplotypes carrying the 1241Glu variant (HT1, HT3, HT5), the protective effect of 1241Glu and lower levels of FVIII were confirmed, but seemed limited to male carriers of HT1 [58]. This indicates that Asp1241Glu is not likely to be functional but is probably linked to a functional variant. In addition, the risk reduction was only partially dependent on the lower levels of FVIII, which suggests that not only FVIII levels influence risk but also protein function may contribute.

Factor XII

Factor XII (FXII) is involved in both the intrinsic coagulation pathway and the fibrinolytic pathway. Although it has been suggested that deficiency of FXII may lead to an increased risk of thrombosis, there is not enough evidence to confirm an association. In 1998, Kanaji et al [74] described the *FXII* 46 C>T polymorphism that was associated with decreased plasma FXII levels. After finding *FXII* 46 C>T associated with FXII levels and venous thrombosis in a linkage study, Tirado et al reported an odds ratio of 4.8 (95% CI, 1.5 to 15.6) for the TT genotype [75]. However, other studies could not confirm this finding [76-78].

Factor XIII

Factor XIII (FXIII) stabilizes the fibrin clot by cross-linking fibrin monomers. A genetic variant that interferes with clot stabilization is the *FXIII* Val34Leu polymorphism. The *FXIII* 34Leu variant is more rapidly activated and thus more rapidly cross-links fibrin fibers. However, these fibers are thinner and the fibrin clots are more solid [79]. The 34Leu variant was found to be protective when first studied in relation to venous thrombosis [80]. Subsequent studies mostly failed to confirm this finding, perhaps due to small study sizes, but in a meta-analysis a pooled odds ratio of 0.63 (95% CI, 0.46 to 0.86) was computed [81]. Recently, this protective effect was shown to depend on fibrinogen levels [82]. High levels of fibrinogen also lead to less porous fibrin clots, and seem to act synergistically with *FXIII* 34Leu, being protective against venous thrombosis. Another polymorphism was reported in the FXIII B-subunit. This is a carrier protein for FXIII and dissociates upon activation of FXIII. The 8259 A>G (His95Arg) variant leads to increased dissociation and moderately increases the risk of venous thrombosis [60].

Tissue Factor

Tissue factor (TF) initiates coagulation by activating factor VII. TF is expressed by most cells and organs and a soluble form of TF might also circulate in plasma. The relationship between circulating TF levels and venous thrombosis is not clear. Arnaud et al [61] screened the promoter region of the *TF* gene in blood donors and identified a deletion/insertion polymorphism, 1208 D/I, of which the 1208 D variant was associated with lower circulating TF levels and a lower risk of venous thrombosis in a subsequent case-control study. The coding region of the *TF* gene was screened in a case-control setting by Zawadzki et al [83]. They found a C>T (Arg200Trp) variant that might lower monocyte TF concentrations, but the variant was too rare to study its effect on the risk of thrombosis.

Factor VII–Activating Protease

Factor VII-activating protease (FSAP) stimulates coagulation by activating FVII. It also stimulates fibrinolysis by activating urokinase precursor. The

Marburg I polymorphism (1601 G>A, Gly511Gly) impairs the profibrinolytic activity of FSAP but has no effect on its FVII-activating function [84]. Initially, an increased risk of venous thrombosis was reported for carriers of the Marburg I polymorphism, mainly for idiopathic cases [85]. However, this finding could not be confirmed by others [86,87].

Blood Group

ABO blood group is associated with both FVIII and von Willebrand factor (VWF) levels, and a reduced risk of venous thrombosis for phenotypic blood group O was already recognized by Jick et al [88] in 1969. The protective effect of blood group O is mainly explained by FVIII levels [66,89]. In 2005, several investigators reported on the association between blood group genotype and venous thrombosis [66,90-92]. All studies confirmed the lower risk for blood group O (genotypes O^1O^1 and O^1O^2), and the highest risk for carriers of the A^1 allele. Blood group O genotypes were associated with lower FVIII and VWF levels [91,92]. The protective effect is restricted to homozygous carriers of O alleles; hence individuals with phenotypic blood group O [90]. There is some evidence that the risk for non-OO genotypes is higher in carriers of factor V Leiden [66].

Angiotensin-Converting Enzyme

Angiotensin-converting enzyme (ACE) stimulates platelet activation and regulates fibrinolysis. The relationship between ACE levels and the risk of venous thrombosis is, however, not clear. Variation in ACE levels is largely explained by a 287-bp insertion/deletion (I/D) polymorphism in the *ACE* gene. Individuals with the DD genotype have higher plasma ACE than individuals with the II genotype [93]. Initially, the DD genotype was reported to increase the risk of venous thrombosis [94,95]. Subsequent studies reported varying associations between the I/D variant and venous thrombosis, as summarized by Okumus et al [96] and Ekim et al [97]. An important conclusion from both reviews was that studies were heterogeneous in patient selection, with regard to etiology and ethnicity. The most recent and largest study found a slightly protective effect of the DD genotype, which was restricted to women. However, the authors concluded that the I/D polymorphism itself is unlikely to be functional [62].

Tissue Factor Pathway Inhibitor

The first step in the extrinsic coagulation pathway is inhibited by tissue factor pathway inhibitor (TFPI). Low TFPI levels are associated with an increased risk of venous thrombosis [98]. Five polymorphisms have been described in the *TPFI* gene. The -399 C>T polymorphism, reported by Miyata et al [99] was not associated with TFPI concentrations or with venous thrombosis. Kleesiek et al [63] screened the coding region of the gene and found a 536 C>T (Pro151Leu) variant that was associated with an increased risk of venous thrombosis but not with TFPI plasma activity and concentration. The association with venous thrombosis was not confirmed by others [100-102]. Three other polymorphisms were reported by Moatti et al [103,104]. The 874 G>A (Val264Met) variant might influence TFPI levels, but it was not associated with venous thrombosis [98,105]. The intron variant -33 T>C and the promoter polymorphism -287 T>C were both associated with higher TFPI levels [64,104]. Accordingly, -33C was found to be protective for venous thrombosis [64]. The -287 T>C variant has not yet been studied in venous thrombosis patients.

Endothelial Protein C Receptor

The endothelial protein C receptor (EPCR) is expressed on the endothelium of large vessels. Binding of protein C stimulates protein C activation by the thrombin–thrombomodulin complex. A soluble form of EPCR (sEPCR) also binds protein C but inhibits protein C activity. Low levels of sEPCR have been found to reduce the risk of venous thrombosis [106]. Biguzzi et al screened the *EPCR* gene in venous thrombosis patients and found a 23-bp insertion (4031ins23) that impaired EPCR function, and four promoter polymorphisms that did not affect transcription in vitro [107,108]. These variants were rare in both patients and control subjects and therefore the effect on thrombosis risk could not be established [109-111]. España et al [112] reported more common polymorphisms in exon 4 (4600 A>G, Ser219Gly) and in the 3′UTR (4678 G>C). The 4600G genotype was associated with increased sEPCR levels and the 4678C genotype was associated with increased APC levels. Conflicting results were published about the association with venous thrombosis [106,113-115]. Two studies constructed haplotypes of the *EPCR* gene. Saposnik et al [116] found

three haplotypes of which the haplotype corresponding to the 4600G genotype (H3) was associated with increased sEPCR levels and an increased risk of venous thrombosis. Uitte de Willige et al [106] identified an additional haplotype (H4) that also contained 4600G and was part of H3 in the study of Saposnik et al. In this study, carriers of H4 (not H3) had a slightly increased risk of venous thrombosis, which suggests that 4600G itself is not a functional variant.

Plasminogen Activator Inhibitor-1

Plasminogen activator inhibitor-1 (PAI-1) indirectly inhibits the fibrinolytic activity of plasminogen. There is no clear relationship between PAI-levels and venous thrombosis [117]. Dawson et al [118] described a guanine deletion/insertion (4G/5G) upstream of the *PAI-1* gene that influences PAI-1 activity. A literature review regarding 4G/5G and venous thrombosis by Francis [117] showed that although most studies find higher plasma levels of PAI-1 in individuals with 4G/4G, the effect on the risk of venous thrombosis is not clear. There is some evidence that 4G/4G increases the risk of venous thrombosis in subgroups with additional genetic risk factors [65,119], but these findings need to be confirmed. The fact that the 4G/5G variant is associated with PAI-I levels but not with venous thrombosis suggests that PAI-I levels have no major effect on the risk of venous thrombosis.

Thrombin-Activatable Fibrinolysis Inhibitor

During fibrinolysis, partially degraded fibrin is a cofactor for plasminogen activation. Thrombin-activatable fibrinolysis inhibitor (TAFI) inhibits fibrinolysis by suppressing this cofactor activity. Increased TAFI levels were found to be associated with a slightly increased risk of venous thrombosis. Zhao et al [120] described the 505 G>A polymorphism (Ala147Thr) that was associated with higher TAFI concentrations. In addition, several promoter polymorphisms (-438 A>G, 1102 T>G, 1690 G>A) and an additional polymorphism in the coding region (1040 C>T, Thr325Ile) were described that influenced TAFI levels and possibly the risk of venous thrombosis, but these variants were strongly linked to the 505 polymorphism [121,122]. Martini et al [122] analyzed haplotypes constructed of -438 G>A, 505 G>A and 1040

C>T and found that only 505A was associated with a reduced risk of venous thrombosis. This was an unexpected finding because 505A is the variant associated with higher TAFI levels. Further research is needed to unravel the relationship between TAFI and venous thrombosis.

5,10-Methylenetetrahydrofolate Reductase

Increased homocysteine has been associated with venous thrombosis in many studies but the mechanism by which homocysteine affects coagulation is not clear. The common 677 C>T (Ala222Val) polymorphism in the 5,10-methylenetetrahydrofolate reductase (*MTHFR*) gene, leading to a thermolabile variant of the enzyme, slightly increases plasma homocysteine levels and has been studied extensively. Odds ratios computed from a large meta-analysis indicate that *MTHFR* 677 C>T is a weak risk factor for venous thrombosis [123]. However, we have recently studied *MTHFR* 677 C>T in a large case-control study but found no evidence of an association with venous thrombosis [124].

Protein Z–Dependent Protease Inhibitor

Although in vitro studies suggest that protein Z–dependent protease inhibitor (ZPI) might influence coagulation by inhibiting FXa and FXIa [125], ZPI plasma levels do not seem to affect the risk of venous thrombosis [126]. Van de Water et al [125] screened the coding region of the *ZPI* gene for mutations and identified two premature stop codons at Arg67 and Trp303 that were associated with venous thrombosis. However, in subsequent studies, the Trp303 stop codon was not detected [127] or the presence of either Arg67Stop or Trp303Stop was not associated with venous thrombosis [67]. Recently, however, Corral et al [128] studied haplotypes of the *ZPI* gene and found the haplotype covering Arg67Stop to be associated with venous thrombosis with a 3.3-fold increased risk. Since Arg67Stop is a rare mutation, large studies are needed to confirm this finding.

Factors VII, IX, X, XI, von Willebrand Factor, and Thrombomodulin

Despite a clear association between plasma levels and the occurrence of

venous thrombosis, no predictive genetic variants are known in the genes encoding factor IX and factor XI. The promoter region of the *FXI* gene has been screened and two single-nucleotide polymorphisms (SNPs) were identified [129], but the association with plasma concentrations of FXI and venous thrombosis has not yet been studied. Plasma levels of factor X and von Willebrand factor are also related to the risk of venous thrombosis, but the association is dependent on levels of other coagulation factors [66,89,130]. For both factors, gene variants were studied in relation to plasma concentrations and the risk of venous thrombosis, but no association was found [72,130]. Plasma concentrations of factor VII and thrombomodulin are not related to the risk of venous thrombosis. Some genetic variants were identified that were related to levels but not to the risk of venous thrombosis [46,47,131-134].

FUTURE PERSPECTIVES

Research Goals

All variants that were initially found to be associated with venous thrombosis are listed in Table 1. For each factor, the magnitude of the association with venous thrombosis is given in the right column. However, most of these findings have not yet been replicated, and others have failed to replicate. Failure of replication may occur for several reasons.First, many genetic variants of modest effect are expected to contribute to the risk of venous thrombosis. To detect these modest effects large sample sizes are needed. All of the genetic variants discussed above are at most weak risk factors and were studied in samples of at most several hundreds of study subjects.

Another obstacle for replication is phenotype definition. Venous thrombosis is a multicausal disease and genetic variants that influence the risk of disease are expected to affect only one of the causal pathways. The intermediate phenotype of that causal pathway (for example, protein levels) will be more strongly correlated to the genetic variant. On the other hand, when the

intermediate phenotype is not associated with venous thrombosis, which is the case with FVII, FX, FXII, von Willebrand factor, and thrombomodulin, finding an association between venous thrombosis and the genetic variant is less likely.

Alternatively, the initial finding may have been a false-positive result. As the amount of genetic variants tested increases, the possibility of finding false-positive results also increases when multiple testing is not correctly accounted for. The problem of false-positive findings probably accounts for many of the non-replicated associations.

Once a finding is replicated, the knowledge from previous studies can be used to further explore the relationship between the replicated variant and disease. This can be done by constructing haplotypes in the genes where association was found, as was the case in relation to the fibrinogen [50], *FVIII* [58], and *EPCR* [116] genes. In these studies, haplotypes were identified that increased the risk of thrombosis and probably explained the association observed before with a single nucleotide polymorphism. Knowledge from previous research can further be used to study interaction of genetic variants that seem to affect the same causal pathway. The synergistic effect of the *FXIII* Val34Leu variant and fibrinogen levels observed by Vossen et al [82] is an example of this approach.

Clinical Implications

The ultimate goal of genetic epidemiologic research is to be able to predict each individual's risk of disease on the basis of genetic and acquired factors. With a genetic profile it would then be possible to give lifestyle recommendations or prophylaxis in high-risk situations to maximally reduce the risk of disease.

However, venous thrombosis is a complex disease and prediction requires an extensive model including many acquired and genetic factors. To be clinically useful, the genetic variants described above should be evaluated together with many other, as yet unknown, risk factors. In addition, the outcome of testing the genetic variant should influence treatment or preventive measures. The

various possible or weak genetic risk factors for venous thrombosis that have been identified in the past years are not of clinical importance.

CONCLUSION

In the past decade, many factors have been described that might influence the risk of venous thrombosis. Some of these, like ABO blood group, FXIII Val34Leu, or haplotypes of fibrinogen, *FV*, and *FVIII*, are possible novel predictive variants. Others have failed to replicate in subsequent studies. More well-designed, large studies are needed to unravel the various pathways that lead to the development of venous thrombosis.

No Association Between The Common MTHFR 677C>T Polymorphism And Venous Thrombosis

Chapter 4

ID Bezemer
CJM Doggen
HL Vos
FR Rosendaal

Archives of Internal Medicine 2007;167(5):497-501

ABSTRACT

Background Increased homocysteine levels are related to the occurrence of venous thrombosis, but whether this relation is causal is unclear. The T-variant of the common methylenetetrahydrofolate reductase (MTHFR) 677CT polymorphism mildly increases homocysteine levels. Meta-analyses have demonstrated a weak effect of the MTHFR 677TT genotype on risk but are sensitive to selective publication of positive results. The aim of the present study was to evaluate the effect of the MTHFR genotype on the risk of venous thrombosis, overall and in subgroups of known risk factors, in a single large study.

Methods In the Multiple Environmental and Genetic Assessment of risk factors for venous thrombosis (MEGA Study), a population-based case-control study, we collected DNA from 4375 patients with a first deep vein thrombosis of the leg or pulmonary embolism and from 4856 control subjects. Information about risk factors for venous thrombosis was obtained from questionnaires.

Results MTHFR 677CT was not associated with the risk of venous thrombosis (odds ratio [95% confidence interval], 0.99 [0.91-1.08] for the CT genotype and 0.94 [0.81-1.08] for the TT genotype). Stratification by known risk factors for venous thrombosis provided no evidence of an association in specific groups.

Conclusions In a single large study, MTHFR 677CT was not associated with the risk of venous thrombosis, and the narrow confidence interval excludes even a small effect. Therefore, mildly elevated homocysteine levels as a result of MTHFR 677TT do not seem to cause venous thrombosis. There is no rationale for measuring the MTHFR 677CT variant for clinical purposes.

INTRODUCTION

Venous thrombosis is a common disease with an annual incidence of 1 to 3 in 1000 individuals and is caused by the joint effect of environmental and genetic risk factors [135]. One of the genetic factors that have been extensively studied over the past decade is a polymorphism in the gene encoding 5,10-methylenetetrahydrofolate reductase (*MTHFR*). MTHFR is an enzyme involved in homocysteine metabolism by converting folate, a cofactor for homocysteine conversion, into its major circulating form 5-methyltetrahydrofolate. A common C>T substitution at nucleotide 677 converts an alanine to a valine residue [136] and causes thermolability of the enzyme at 37°C. Homozygotes have more than 50% reduced enzyme activities, but the effect of reduced *MTHFR* activity on homocysteine levels is dependent on folate intake. Homocysteine levels are about 25% higher in homozygous carriers only when plasma folate concentration is low [137].

Hyperhomocysteinemia is associated with venous thrombosis [138] and therefore *MTHFR* 677C>T has been one of the candidate genetic risk factors for venous thrombosis. However, most case-control and cohort studies that assessed the association between *MTHFR* 677C>T and venous thrombosis reported either a weak association or no relationship at all. Because these studies were small and often underpowered to detect weak effects, several meta-analyses have been performed [123,139,140]. The most recent and largest meta-analysis, including 8364 cases and 12 468 controls, found a small increase in risk for *MTHFR* 677TT carriers (odds ratio [OR], 1.20; 95% confidence interval [CI], 1.08-1.32) [123]. This risk increase is in line with the expected risk based on the association of the variant with homocysteine levels and the effects of hyperhomocysteinemia. The major disadvantage of meta-analyses, however, compared with a single large study, is the possibility of publication bias. Meta-analyses are based on published studies and rely on the quality of the collected studies. Publication bias is present when studies with "positive" results have a higher probability of being published than studies with "negative" results. It has been shown repeatedly that

publication bias is common in the medical literature [141]. Publication bias leads to an overestimate of the risk. Although no evidence of publication bias was found in the meta-analysis mentioned previously [123], it cannot be ruled out that studies that found no association, irrespective of sample size, were underrepresented. The alternative is a single large study, in which publication bias obviously can play no role.

Several studies have suggested that the effect of *MTHFR* 677C>T on venous thrombosis is only visible in specific subgroups with other predisposing genetic or environmental factors or, on the contrary, in subgroups in which conventional risk factors for venous thrombosis are absent [142-150]. To adequately investigate an effect in subgroups, a large study is needed because only a relatively small proportion of study subjects will carry both the risk factor and the *MTHFR* 677TT genotype.

In this article, we report on the association between *MTHFR* 677C>T and venous thrombosis in a single large study. The analysis included 4375 patients with a first venous thrombotic event, either deep vein thrombosis of the leg or pulmonary embolism, and 4856 control subjects from the Multiple Environmental and Genetic Assessment of risk factors for venous thrombosis (MEGA Study).

METHODS

Study Population

Between March 1, 1999, and August 31, 2004, consecutive patients aged 18 to 70 years with a first venous thrombosis of the leg or arm or a pulmonary embolism were recruited from 6 anticoagulation clinics in the Netherlands. Partners of patients were invited as control subjects. An additional control group was recruited between January 1, 2002, and December 1, 2004, using a random digit dialing (RDD) method [33]. The random control group was age- and sex frequency matched to the group of patients that provided a

blood sample. For practical reasons individuals with severe psychiatric problems or individuals who did not speak Dutch were excluded.

All participants were asked to fill in a questionnaire on acquired risk factors for venous thrombosis, family history of venous thrombosis, and vitamin B supplementation. The date of diagnosis as reported by the patient in the questionnaire or, when missing, the date of the first visit at the anticoagulation clinic served as the index date for patients and their partners. The index date for RDD control subjects was the date on which the questionnaire was completed or, when missing, the date on which the completed questionnaire was returned.
A blood sample was taken approximately 3 months after discontinuation of anticoagulant therapy. If patients continued their anticoagulant therapy, blood was drawn 1 year after the index date. Partner controls were invited for a blood draw along with their partner; RDD controls were invited after returning their questionnaire. Blood samples were taken from patients who were diagnosed before June 1, 2002, and their partners. Patients who were diagnosed from June 1, 2002, onwards and their partners received a cotton swab along with their questionnaire for collecting buccal cells. In the RDD group, blood samples were collected throughout the entire study period. Participants who refused to or were unable to provide a blood sample were offered the option of providing a buccal swab sample. Ideally, participants filled in a full questionnaire and provided a DNA sample, but a minor proportion provided only DNA. When the full questionnaire was not returned, we attempted to collect information about acquired risk factors (but not family history and vitamin supplementation) through a miniquestionnaire by telephone.

Of 6333 eligible patients with deep vein thrombosis of the leg or a pulmonary embolism, 358 died before inclusion and 271 could not be contacted. Of the remaining 5704 patients, 5053 (89%) participated. A DNA sample was donated by 4379 patients, among whom 4257 full questionnaires and 99 miniquestionnaires were collected.

Partners of participating patients were invited as control subjects. Of 3655 eligible partners, 1 died before inclusion and 10 could not be contacted. Of the remaining 3644 partners, 2984 (82%) participated. A DNA sample was donated by 2602 partners. In addition, we collected DNA from 240 controls whose partner had venous thrombosis of the arm (n = 104), eventually refused to participate (n = 10), or was excluded (n = 126). Among these 2850 partners, 2800 full questionnaires and 34 miniquestionnaires were collected.

The RDD method yielded 4350 eligible control subjects, but 4 died before inclusion and 88 control subjects could no longer be contacted despite repeated efforts. Of the remaining 4258, 3000 (70%) participated. A DNA sample was provided by 2023 RDD controls, among whom 2011 full questionnaires and 11 miniquestionnaires were collected.

To study the association between *MTHFR* 677C>T and venous thrombosis in subgroups, participants were stratified according to acquired risk factors as reported in the questionnaire. Participants were also stratified according to 2 common genetic risk factors, factor V Leiden, and prothrombin 20210G>A. Complete genetic data on *MTHFR* 677C>T, factor V Leiden, and prothrombin 20210G>A was available from 4375 patients and 4856 control subjects, and these were included in the present analysis.

Laboratory Analysis

Collection and processing of blood samples and buccal swabs and subsequent DNA isolation has been described previously [19]. Assessment of *MTHFR* 677C>T (rs1801133), factor V Leiden (rs6025), and prothrombin 20210G>A (rs1799963) in DNA retrieved from whole blood or DNA from buccal swabs was initially performed by restriction fragment length polymorphism analysis after conventional polymerase chain reaction (PCR). The presence of the *MTHFR* 677T allele was assessed by incubation with the restriction enzyme *Hinf*I. Factor V Leiden and prothrombin 20210G>A were analyzed in a combined method using *Mnl*I and *Hind*III restriction enzymes. Later, all 3 polymorphisms were determined by a 5' nuclease (Taqman;

Applied Biosystems, Foster City, Calif) assay using a standard PCR reaction mix (Eurogentec, Seraing, Belgium) and allele-specific fluorescent probes equipped with a minor groove binding moiety (Applied Biosystems).

Statistical Analysis

Odds ratios and 95% CIs were computed as an estimate of the risk of venous thrombosis associated with *MTHFR* 677CT and TT genotypes relative to CC. Adjustment for age and sex was performed by logistic regression.

The association between *MTHFR* genotype and venous thrombosis was further explored through stratification by known risk factors and computing ORs for the *MTHFR* 677TT genotype in strata of the risk factor under study, relative to the combined 677CT or CC genotype. Strata were made for the factor V Leiden, prothrombin 20210G>A, vitamin B supplementation, age group, family history of venous thrombosis, and the presence of predisposing factors for venous thrombosis.

Age was categorized as younger than 50 years (18-49 years) or 50 years and older (50-70 years). Family history was positive if at least 1 parent or sibling had venous thrombosis or a pulmonary embolism before the age of 50 years. For 2212 (24%) of 9055 participants with a full questionnaire, family history status could not be determined because of incomplete data on 1 or more family members. Predisposing factors were surgery, immobilization, pregnancy or puerperium within the year preceding the index date, and diagnosis of malignancy before or within 6 months after the index date. Participants who did not have complete information on these variables (239 [3%] of 9199 miniquestionnaires) were not included in the subgroup analysis. Vitamin B supplementation was defined as the self-reported use of vitamin supplementation that contained pyridoxine hydrochloride (vitamin B_6), folic acid (vitamin B_{11}), or cyanocobalamin (vitamin B_{12}), which are all cofactors for homocysteine conversion. Among 383 (4%) of 9055 participants, no information on vitamin B use was available. All statistical analyses were performed with SPSS for Windows, release 12.0.1 (SPSS Inc, Chicago, Ill).

RESULTS

Patients included in the analysis were diagnosed as having a first deep vein thrombosis of the leg (n = 2519), a first pulmonary embolism (n = 1315), or both (n = 541). In total, 4375 patients and 4856 control subjects were included. Median age (5th-95th percentile) at the index date was 50 years (26-68 years) for patients and 49 years (27-67 years) for control subjects. Slightly more women than men were included in both groups (54% of patients and 53% of control subjects).

The MTHFR 677TT genotype was present in 440 patients (10%) and 517 control subjects (11%), and the 677CT genotype in 1891 patients (43%) and in 2094 control subjects (43%).
Since genotype distributions did not differ between these 2 groups, there was no excess risk associated with the T allele: ORs (95% CIs) of venous thrombosis when carrying the 677T allele were 0.99 (0.91-1.08) for heterozygous and 0.94 (0.81-1.08) for homozygous carriers, relative to 677CC (Table 1).

Table 1. MTHFR Genotype Distribution Among Patients With Venous Thrombosis and Control Subjects

MTHFR 677 C>T	Cases, No (%) (n=4375)	Control subjects, No (%) (n=4856)	Odds Ratio (95% Confidence Interval)
CC	2044 (47)	2245 (46)	1 [Reference]
CT	1891 (43)	2094 (43)	0.99 (0.91-1.08)
TT	440 (10)	517 (11)	0.94 (0.81-1.08)

Factor V Leiden was present in 685 cases (16%) and 256 control subjects (5%) (Table 2). The association between MTHFR 677C>T and venous thrombosis did not differ between strata of factor V Leiden. Among patients, 224 individuals (5%) carried the prothrombin 20210A mutation and 94 controls (2%) were carrier. In prothrombin 20210A carriers,

the risk associated with MTHFR 677TT was somewhat higher than in noncarriers (OR, 1.63) but the 95% CI was wide (0.75-3.56).

Table 2. Association of Venous Thrombosis With MTHFR 677TT Relative to 677CC/677CT in Subgroups of Coexisting Risk Factors

	No.*	Odds Ratio (95% Confidence Interval)		No.*	Odds Ratio (95% Confidence Interval)
Factor V Leiden GG	8290	0.91 (0.79-1.06)	Factor V Leiden GA/AA	941	0.94 (0.61-1.45)
Prothrombin 20210 GG	8913	0.91 (0.79-1.05)	Prothrombin 20210 GA/AA	318	1.63 (0.75-3.56)
Age 18-49	4775	1.04 (0.87-1.26)	Age 50-70	4456	0.84 (0.69-1.02)
Negative family history	5999	0.97 (0.82-1.15)	Positive family history	844	0.98 (0.61-1.57)
No predisposing factors	5956	0.91 (0.76-1.08)	Predisposing factors	3004	1.05 (0.81-1.35)
Vitamin B supplementation	2441	0.96 (0.74-1.25)	No vitamin B supplementation	6407	0.93 (0.79-1.10)

* Total number of subjects in indicated group

The association between MTHFR 677C>T was also studied in a subgroup of participants who did not take vitamin supplements containing folic acid (vitamin B11), vitamin B6, or vitamin B12. The use of these vitamin supplements was more frequently reported by control subjects than by patients (29% vs 26%). In the subgroup without vitamin B supplementation, no effect of MTHFR genotype was observed (OR, 0.93; 95% CI, 0.79-1.10).

The association between MTHFR 677C>T and venous thrombosis was further explored by stratifying patients and control subjects according to age at index date (age 50 or <50 years), family history of venous thrombosis, and

the presence of predisposing factors for venous thrombosis. In none of these strata was an effect of MTHFR 677TT observed.

Adjustment for age or sex, stratifying patients according to diagnosis (thrombus in the leg, pulmonary embolism, or both), excluding study subjects with cancer, or restricting the control group to either partners of patients or RDD control subjects did not change these observations.

COMMENT

In the MEGA Study, a very large population-based case-control study, *MTHFR* 677C>T was not associated with the risk of venous thrombosis. Stratification by factor V Leiden, prothrombin 20210G>A, family history, age, presence of predisposing factors, and vitamin B supplementation did not provide evidence of an association in specific groups.

These results should be seen in the light of many conflicting study results. As summarized in 3 meta-analyses, the majority of previous single studies found no association. From these meta-analyses mildly elevated ORs of 1.29 (95% CI, 1.08-1.54) [140], 1.2 (95% CI, 1.1-1.4) [139], and 1.20 (95% CI, 1.08-1.32) were calculated [123]. These 3 meta-analyses mostly included the same single studies. The most recent one was the largest and included 20 832 study subjects from 53 studies [123]. When these single studies were grouped by population, a slightly increased risk of venous thrombosis was associated with the *MTHFR* 677TT genotype in European populations (OR, 1.15; 95% CI, 1.02-1.28). In total, 30 European case-control studies were included, with 1280 subjects in the largest single study [151]. If this meta-analysis [123] were repeated to include the MEGA Study, the risk estimate for European studies would decrease to 1.06 (95% CI, 0.96-1.16). The absence of any association in the present analysis, which is many times larger than any previously published case-control study, suggests that there may have been an overrepresentation in the literature of studies with positive results.

MTHFR 677C>T has been a candidate genetic risk factor for venous thrombosis because its phenotype, elevated serum homocysteine level, is associated with venous thrombosis. In the present study, homocysteine levels were not measured, but elevated levels were observed in homozygous carriers of the T variant in many studies, including studies in the Dutch population [143]. The mechanism by which homocysteine would affect thrombotic risk is unknown, and therefore it is still a matter of debate whether the relation is causal or whether homocysteine is a marker of other causal risk factors or the consequence of venous thrombosis. The study of the *MTHFR* genotype offers the possibility to investigate these various hypotheses, since a genotype cannot be a marker of another risk factor or a post hoc phenomenon. So, when elevated levels of homocysteine cause thrombosis, *MTHFR* 677TT is expected to be related to thrombotic risk. According to this reasoning, the absence of an association between *MTHFR* genotype and venous thrombosis suggests that the association between elevated homocysteine levels and venous thrombosis is not causal.

Another way to disentangle causal and noncausal effects is to perform an experiment. If hyperhomocysteinemia causes thrombosis, lowering homocysteine level is expected to protect those with the *MTHFR* 677TT genotype from developing thrombosis. Several randomized trials in which homocysteine level was lowered by vitamin B supplementation have been performed or are still ongoing, both in arterial and venous disease. Three trials on the effect of lowering homocysteine level on arterial thrombosis have been completed [152-154], and 1 trial examined the effect in venous thrombosis [155]. Despite a decrease in homocysteine levels in the vitamin treatment group, none of these trials showed a beneficial effect on disease outcome. Thus, the results of these trials do not support the hypothesis that high levels of homocysteine cause thrombosis. It should be noted, however, that these trials examined the effect of lowering homocysteine level on recurrent thrombosis, not on a first event.

Alternatively, the effect of thermolabile MTHFR on homocysteine levels may be too small to cause thrombosis on its own. In the Leiden Thrombophilia Study [138], the risk of thrombosis was only increased when homocysteine

concentrations were above 2.43 mg/L (>18 µmol/L), compared with a reference level below 1.62 mg/L (<12 µmol/L), which corresponds to at least 50% higher homocysteine concentrations. The 25% increase in homocysteine concentrations that is generally observed in individuals with the *MTHFR* 677TT genotype may therefore not be enough in most cases to cause thrombosis.

The *MTHFR* 677TT genotype increases homocysteine levels only when combined with low vitamin B levels. Sufficient intake of folic acid (vitamin B_{11}), vitamin B_6, or vitamin B_{12} normalizes serum homocysteine in *MTHFR* 677TT carriers [156,157]. Therefore, the use of vitamin supplements could mask a possible association between *MTHFR* genotype and venous thrombosis. In the MEGA Study, the association between *MTHFR* genotype and venous thrombosis did not depend on vitamin B supplementation as reported in the questionnaire.

If *MTHFR* 677C>T is only a weak risk factor for venous thrombosis, it may only be discernible in individuals with a specific predisposition for developing venous thrombosis. A number of studies have evaluated the joint effect of *MTHFR* 677C>T and predisposing genetic factors, mainly factor V Leiden and prothrombin 20210G>A. Factor V Leiden was most frequently studied, but owing to the small numbers in strata of the combined factors, risk estimates in these studies had wide confidence intervals [142,144,145,148,149]. Only 1 study showed an association between *MTHFR* 677C>T in carriers of factor V Leiden [142]. None of the studies that focused on the coexistence of *MTHFR* 677C>T and prothrombin 20210G>A found evidence of effect modification [144,146]. The MEGA Study confirms these negative results.

Other factors previously studied in relation to the *MTHFR* genotype and venous thrombosis were age, family history of venous thrombosis, and presence of acquired risk factors such as recent surgery, immobilization, malignancy, and pregnancy. In previous studies, different results were reported about specific subgroups in which an effect of *MTHFR* genotype

was observed. Some studies found an association in individuals in whom other known risk factors were absent [147], while others suggested that *MTHFR* 677C>T mainly affects the risk of venous thrombosis in cooperation with genetic or acquired risk factors [142,143,150]. Again, small study sizes might account for these conflicting results. The MEGA Study was large enough to make sufficiently large strata. No association between *MTHFR* genotype and venous thrombosis was observed within any of these strata.

Taken together, no evidence was found for an association between *MTHFR* 677C>T and the risk of venous thrombosis. There is no rationale for measuring the *MTHFR* 677C>T variant for clinical purposes.

ACKNOWLEDGMENTS

We thank the directors of the Anticoagulation Clinics of Amersfoort (M. H. H. Kramer, MD), Amsterdam (M. Remkes, MD), Leiden (F. J. M. van der Meer, MD), The Hague (E. van Meegen, MD), Rotterdam (A. A. H. Kasbergen, MD), and Utrecht (J. de Vries-Goldschmeding, MD), who made the recruitment of patients possible. The interviewers (J. C. M. van den Berg, B. Berbee, S. van der Leden, M. Roosen, and E. C. Willems of Brilman) performed the blood draws. We also thank I. de Jonge, MSc, R. Roelofsen, MSc, M. Streevelaar, L. M. J. Timmers, MSc, and J. J. Schreijer for their secretarial and administrative support and data management. The fellows J. W. Blom, MD, A. van Hylckama Vlieg, PhD, E. R. Pomp, MSc, L. W. Tick, MD, and K. J. van Stralen, MSc, took part in every step of the data collection. C. J. M. van Dijk, R. van Eck, J. van der Meijden, P. J. Noordijk, and T. Visser performed the laboratory measurements. We also thank M. den Heijer, MD, for providing the individual study data from his meta-analysis. We express our gratitude to all individuals who participated in the MEGA Study.

Gene Variants Associated With Deep Vein Thrombosis

Chapter 5

ID Bezemer
LA Bare
CJM Doggen
AR Arellano
C Tong
CM Rowland
J Catanese
BA Young
PH Reitsma
JJ Devlin
FR Rosendaal

JAMA. 2008;299(11):1306-1314

ABSTRACT

Context The genetic causes of deep vein thrombosis (DVT) are not fully understood.

Objective To identify single-nucleotide polymorphisms (SNPs) associated with DVT.

Design, Setting, and Patients We used 3 case-control studies of first DVT. A total of 19 682 gene-centric SNPs were genotyped in 443 cases and 453 controls from the Leiden Thrombophilia Study (LETS, 1988-1992). Twelve hundred six SNPs associated with DVT were reinvestigated in the Multiple Environmental and Genetic Assessment of Risk Factors for Venous Thrombosis study (MEGA-1, 1999-2004) in a subset of 1398 cases and 1757 controls. Nine SNPs associated with DVT in both LETS and MEGA-1 were investigated a third time in 1314 cases and 2877 controls from MEGA-2, a second subset of MEGA. Additional SNPs close to one SNP in *CYP4V2* were genotyped in LETS and MEGA-1.

Main Outcome Measure Odds ratios (ORs) for DVT were estimated by logistic regression. False discovery rates served to investigate the effect of multiple hypothesis testing.

Results Of 9 SNPs genotyped in MEGA-2, 3 were strongly associated with DVT ($P < .05$; false discovery rate ≤.10): rs13146272 in *CYP4V2* (risk allele frequency, 0.64), rs2227589 in *SERPINC1* (risk allele frequency, 0.10), and rs1613662 in *GP6* (risk allele frequency, 0.84). The OR for DVT per risk allele was 1.24 (95% confidence interval [95%CI], 1.11-1.37) for rs13146272, 1.29 (95% CI, 1.10-1.49) for rs2227589, and 1.15 (95% CI, 1.01-1.30) for rs1613662. In the region of *CYP4V2*, we identified 4 additional SNPs (in *CYP4V2*, *KLKB1*, and *F11*) that were also associated with both DVT (highest OR per risk allele, 1.39; 95% CI, 1.11-1.74) and coagulation factor XI level (highest increase per risk allele, 8%; 95% CI, 5%-11%).

Conclusions We identified SNPs in several genes that were associated with DVT. We also found SNPs in the region around the SNP in *CYP4V2* (rs13146272) that were associated with both DVT and factor XI levels. These results show that common genetic variation plays an important role in determining thrombotic risk.

INTRODUCTION

The incidence of deep vein thrombosis (DVT) is 1 per 1000 person-years [1]. The 10-year recurrence risk is 30% [158]. Deep vein thrombosis can lead to life-threatening pulmonary embolism [159]. Deep vein thrombosis is caused by acquired and genetic risk factors. Acquired risk factors include age, hospitalization, cancer, pregnancy, hormone therapy, and surgery [158]. Family and twin studies indicate that genetics accounts for about 60% of the risk for DVT [160,161]. Deficiencies of natural anticoagulants antithrombin, protein C, and protein S are strong risk factors for DVT; however, the variants causing these deficiencies are rare and explain only about 1% of all DVTs [6]. Two more common genetic variants, Factor V Leiden (FVL) and prothrombin G20210A, have been consistently found to be associated with DVT [51,162] but still only explain a fraction of the DVT events [6]. It has been suggested that 2 or more risk factors are needed for thrombosis [6,36,163].

The identification of additional common gene variants associated with DVT will improve the ability to predict risk for DVT and increase understanding of this disease. Therefore, we investigated whether any of 19 682 primarily missense single-nucleotide polymorphisms (SNPs) were associated with DVT in 3 large case-control studies.

METHODS

Study Populations and Data Collection

The 3 studies (LETS, MEGA-1 and MEGA-2) in the present analysis are derived from 2 large population-based case-control studies: the Leiden Thrombophilia Study (LETS) [17] and the Multiple Environmental and Genetic Assessment of Risk Factors for Venous Thrombosis (MEGA study) [19]. These studies were approved by the Medical Ethics Committee of the Leiden University Medical Center, Leiden, the Netherlands. All participants gave oral informed consent for LETS and written for MEGA to participate.

LETS Population

Collection and ascertainment of DVT events in LETS has been described previously.[11] Briefly, 474 consecutive patients, 70 years or younger, without a known malignancy were recruited between January 1, 1988, and December 30, 1992, from 3 anticoagulation clinics in the Netherlands. For each patient, an age- and sex-matched control participant without a history of DVT was enrolled. Participants completed a questionnaire on risk factors for DVT and provided a blood sample. No ethnicity information was collected. After exclusion of 52 participants due to inadequate sample, 443 cases and 453 controls remained in the analyses.

MEGA-1 and MEGA-2 Studies

Collection and ascertainment of DVT events in MEGA has been described previously [19,20]. MEGA enrolled consecutive patients aged 18 to 70 years who presented with their first diagnosis of DVT or pulmonary embolism (PE) at any of 6 anticoagulation clinics in the Netherlands between March 1, 1999, and May 31, 2004. Control subjects included partners of patients and random population control subjects frequency-matched on age and sex to the patient group. Participants completed a questionnaire on risk factors for DVT and provided a blood or buccal swab sample. The questionnaire included an item on parent birth country as a proxy for ethnicity.

For the present analyses, we split the MEGA study to form 2 case-control studies, based on recruitment date and sample availability (blood or buccal swab). We excluded those with isolated pulmonary embolism or a history of malignant disorders to obtain a study population similar to that of the LETS population. The first subset, MEGA-1, included 1398 cases and 1757 controls who all donated a blood sample. The remaining 1314 cases and 2877 controls who donated either a blood sample or a buccal swab sample were included in MEGA-2.

SNP Association Study

The 19 682 SNPs tested in this study are located in 10 887 genes and were

selected because of their potential to affect gene function or expression [164]. Most SNPs (69%) are missense. Another 24% of the SNPs are located in transcription factor binding sites or in untranslated regions of mRNA, which could affect messenger RNA expression or stability. Ninety-one percent of the SNPs studied have minor allele frequencies of at least 5% in whites. Information on all SNPs tested and primer sequences are available on request.

The design of the SNP association study is presented in the Figure. First, all 19 682 SNPs were tested in pooled DNA samples of LETS (http://www.ncbi.nlm.nih.gov/projects/SNP). Single-nucleotide polymorphisms that were associated with DVT ($P \leq .05$) were tested in pooled DNA samples of MEGA-1. Single-nucleotide polymorphisms that were associated in both LETS and MEGA-1 pools ($P \leq .05$) were confirmed by genotyping individual samples of LETS and MEGA-1. Single-nucleotide polymorphism genotypes consistently associated with DVT in LETS and MEGA-1 ($P \leq .05$) were genotyped in MEGA-2.

Allele Frequency and Genotype Determination

DNA concentrations were standardized to 10 ng/µL using PicoGreen (Molecular Probes, Invitrogen Corp, Carlsbad, California) fluorescent dye. DNA pools, typically of 30 to 100 samples, were assembled based on case-control status, sex, age, and factor V Leiden status. DNA pools were made by mixing equal volumes of standardized DNA solution from each individual sample. Each allele was amplified separately by polymerase chain reaction (PCR) using 3 ng of pooled DNA. In the pooled stage, we used 6 case pools and 4 control pools for LETS, and 13 case pools and 18 control pools for MEGA-1. Allele frequencies in pooled DNA were determined by kinetic polymerase chain reaction (kPCR) [165]. Duplicate kPCR assays were run for each allele and the amplification curves from these assays were used to calculate the allele frequencies of the SNP [165]. Genotyping of individual DNA samples was similarly performed using 0.3 ng of DNA in kPCR assays [165] or using multiplexed oligo ligation assays [166]. Genotyping accuracy of the multiplex method and kPCR has been assessed in

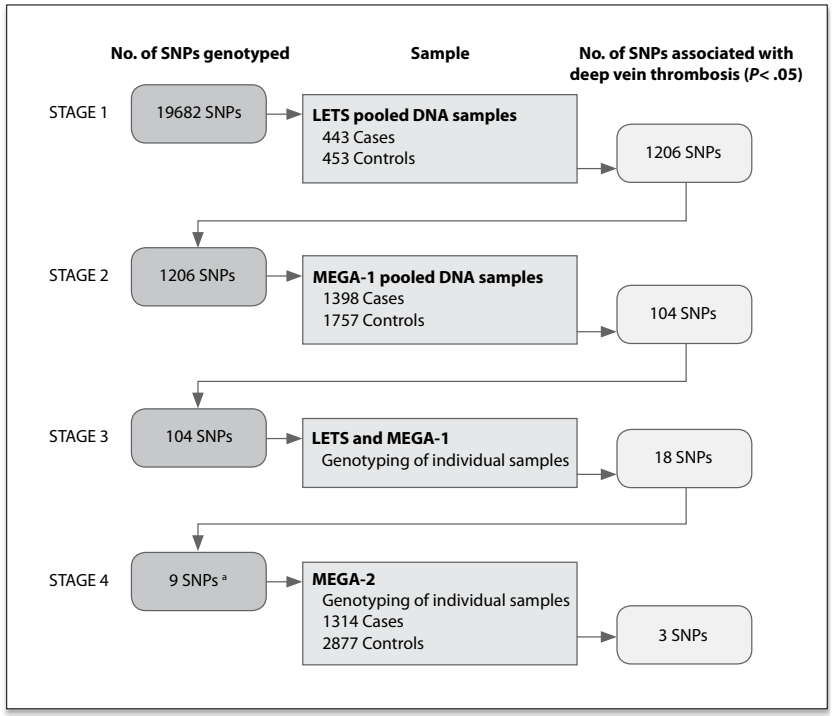

Figure. Flowchart of the Approach Used to Identify SNPs Associated With Deep Vein Thrombosis

[a] Only 9 SNPs were subsequently tested in MEGA-2 because assays for the other 9 were not available.

3 previous studies, and the overall concordance of the genotype calls from these 2 methods was greater than 99% [164,167,168]. The SNPs associated with DVT in MEGA-2 were successfully genotyped in more than 95% of the participants in LETS, MEGA-1, and MEGA-2.

Gene Variants and DVT Risk in the *CYP4V2* Region

The rs13146272 SNP in the gene *CYP4V2* was most strongly associated with DVT in the SNP association study. To investigate whether other SNPs in this region are associated with DVT, we used results from the HapMap Project [169] to identify a region surrounding rs13146272 (chromosome 4:187,297,249-

187,467,731). This region contained 149 SNPs with allele frequencies of more than 2% (HapMap NCBI build 36). Allele frequencies and linkage disequilibrium were calculated from the SNP genotypes in the HapMap Centre d'Etude du Polymorphisme Humain (CEPH) population, which includes Utah residents with ancestry from northern and western Europe. We selected 48 of these 149 SNPs for genotyping, as surrogates for 142 of the 149 SNPs in this region that were either directly genotyped or in strong linkage disequilibrium (r^2>0.8) with at least 1 of the 48 genotyped SNPs (the remaining 7 of the 149 SNPs were in low-linkage disequilibrium with rs13146272 (r^2<0.2) and therefore not likely to be the cause of the observed association). The 48 SNPs were chosen using pairwise tagging in Tagger (implemented in Haploview [170]). The 48 SNPs were initially investigated in LETS, and SNPs that were equally or more strongly associated with DVT than rs13146272 were investigated in MEGA-1.

Factor XI Assays

Factor XI antigen measurements in LETS were described previously [171]. In MEGA, factor XI levels were measured on a STA-R coagulation analyzer (Diagnostica Stago, Asnières, France). STA calcium chloride solution was used as an activator, STA Unicalibrator was used as a reference standard, and Preciclot plus I (normal factor XI range) was used as control plasma. The intraassay coefficient of variation was 5.8% (10 assays). The interassay coefficient of variation was 8.7% (48 assays).

Statistical Analysis

Deviations from Hardy-Weinberg expectations were assessed using an exact test in controls [172]. For pooled DNA analysis, a Fisher exact test was used to evaluate allele frequency differences between cases and controls. For the final set of SNPs, logistic regression models were used to calculate the odds ratio (OR), 95% confidence interval (95% CI), and 2-sided *P* value for the association of each SNP with DVT and to adjust for age and sex. For each SNP, we calculated the OR per genotype relative to noncarriers of the risk allele, and the risk allele OR from an additive model. This risk allele OR

can be interpreted as the risk increase per copy of the risk allele, and the corresponding *P* value was used to decide whether the SNP was associated with DVT ($P \leq .05$). For SNPs on the X chromosome the analysis was conducted separately in men and women.

The OR (95% CI) for SNPs in the *CYP4V2* region was estimated by logistic regression with adjustment for factor XI levels and other SNPs in the region. Differences in factor XI level between groups were tested with *t* tests, and changes in factor XI level per allele were estimated by linear regression.
Analyses were done using SAS version 9 (SAS Institute Inc, Cary, North Carolina) and SPSS for Windows, 14.0.2 (SPSS Inc, Chicago, Illinois).

False Discovery Rate

Studies of thousands of SNPs can lead to false-positive associations. Therefore, we performed 2 replications after the initial discovery stage in LETS and calculated the false discovery rate for the SNPs genotyped in MEGA-2. The false discovery rate estimates the expected fraction of false positives among a group of SNPs; and is a function of the *P* values and the number of tests [173]. False discovery rates were estimated using the 2-sided, unadjusted *P* value from the additive model. We used a false discovery rate of 0.10 as a criterion for further analysis (for a false discovery rate of 0.10, one would expect 10% of the SNPs in the group considered associated to be false positives).

RESULTS

Baseline characteristics of the participants are presented in Table 1.

SNPs Associated With DVT in LETS and MEGA-1

In LETS, we investigated 19 682 SNPs by comparing the allele frequencies of patients and controls using pooled DNA samples [165]. We found that 1206 of these 19 682 SNPs were associated ($P \leq .05$) with DVT. These 1206 SNPs were then investigated in patients and controls from MEGA-1 using pooled

Table 1. Characteristics of Cases and Controls in LETS, MEGA-1, and MEGA-2

	LETS		MEGA-1		MEGA-2	
	Cases (n=443)	Controls (n=453)	Cases (n=1398)	Controls (n=1757)	Cases (n=1314)	Controls (n=2877)
Men, No. (%)	190 (43)	192 (42)	652 (47)	843 (48)	633 (48)	1348 (47)
Age, Mean (SD)	45 (14)	45 (14)	47 (13)	48 (12)	48 (13)	47 (12)
Both parents born in North-West Europe, N(%)[a]	-	-	1247 (91)	1609 (92)	1149 (90)	2527 (89)

[a] No information on birth country was collected in LETS.

DNA samples. The SNPs that were associated with DVT in both LETS and MEGA-1 were confirmed by genotyping in both studies, and we found that 18 SNPs were consistently (with the same risk allele) associated with DVT ($P \leq .05$) in both LETS and MEGA-1 (Table 2).

SNPs Associated with DVT in MEGA-2

Nine of these 18 SNPs were subsequently tested in MEGA-2 for association with DVT (Table 3); assays for the other 9 SNPs were not available at the time. The genotypes of these 9 SNPs did not deviate from the Hardy-Weinberg equilibrium ($P \leq .01$) in the LETS and MEGA controls.

To account for the many tests, we estimated the false discovery rate for the SNPs tested in MEGA-2. In Table 2, factor V Leiden and the prothrombin G20210A mutation are presented for reference. Because these variants were not included in the SNP association study, we did not calculate their false discovery rate. For the SNP in *F9* (rs6048), we only included men because in women, no association with DVT was observed in LETS and MEGA-1. We found that 3 SNPs were again associated with DVT in MEGA-2 ($P \leq .05$), with false discovery rates ≤ .10. These 3 SNPs were in the genes *CYP4V2, SERPINC1*, and *GP6*. The 4 SNPs with the next lowest *P* values (ranging from .06-.15) also had low false discovery rates (≤ .20). These SNPs were in the genes *RGS7, NR1I2, NAT8B*, and *F9*. The risk allele frequencies for

Table 2. Association of 18 SNPs From the SNP Association Study and Factor V Leiden and Prothrombin G20210A With Deep Vein Thrombosis in the LETS and MEGA-1 Studiesa

Chr	Gene	SNP ID	SNP Type[b]	Study	Risk Allele	No. (%) of Alleles Cases	Controls	OR (95% CI)[c]	P-Value
3	NR1I2	rs1523127	5′UTR	LETS	C	373 (42)	300 (33)	1.44 (1.19-1.73)	<.001
				MEGA-1		1185 (42)	1373 (39)	1.15 (1.04-1.27)	.008
19	GP6	rs1613662	Ser219Pro	LETS	A	749 (85)	725 (80)	1.36 (1.07-1.74)	.01
				MEGA-1		2318 (84)	2823 (81)	1.21 (1.06-1.38)	.004
17	APOH	rs1801690	Ser335Trp	LETS	C	850 (97)	852 (95)	1.65 (1.02-2.68)	.04
				MEGA-1		2676 (96)	3312 (94)	1.42 (1.12-1.79)	.004
2	NAT8B	rs2001490	Ala112Gly	LETS	C	382 (43)	348 (38)	1.23 (1.01-1.49)	.04
				MEGA-1		1118 (40)	1301 (37)	1.14 (1.03-1.26)	.01
1	SERPINC1	rs2227589	Intronic	LETS	T	105 (12)	78 (9)	1.42 (1.04-1.94)	.03
				MEGA-1		303 (11)	313 (9)	1.24 (1.05-1.47)	.01
7	MET	rs2237712	Intronic	LETS	G	45 (5)	27 (3)	1.68 (1.05-2.70)	.03
				MEGA-1		119 (4)	110 (3)	1.38 (1.06-1.80)	.02
11	EPS8L2	rs3087546	Leu101Leu	LETS	T	522 (60)	487 (54)	1.26 (1.04-1.52)	.02
				MEGA-1		1637 (59)	1964 (56)	1.12 (1.01-1.24)	.03
6	CASP8AP2	rs369328	Lys93Lys	LETS	A	461 (52)	406 (45)	1.35 (1.11-1.63)	.002
				MEGA-1		1420 (51)	1680 (48)	1.13(1.02-1.24)	.02
1	SELP	rs6131	Asn331Ser	LETS	T	196 (22)	161 (18)	1.29 (1.03-1.62)	.03
				MEGA-1		589 (21)	636 (18)	1.21 (1.06-1.36)	.003
19	ZNF544	rs6510130	Asp203His	LETS	G	33 (4)	13 (1)	2.54 (1.34-4.83)	.004
				MEGA-1		78 (3)	64 (2)	1.56 (1.11-2.18)	.01
1	RGS7	rs670659	Intronic	LETS	C	617 (70)	584 (64)	1.27 (1.04-1.54)	.02
				MEGA-1		1864 (67)	2249 (64)	1.13 (1.01-1.25)	.03
2	TACR1	rs881	3′UTR	LETS	C	745 (85)	713 (80)	1.38 (1.07-1.77)	.01
				MEGA-1		2356 (85)	2894 (83)	1.15 (1.01-1.32)	.04
4	CYP4V2	rs13146272	Lys259Gln	LETS	A	611 (69)	588 (65)	1.22 (1.00-1.49)	.05
				MEGA-1		1896 (68)	2245 (64)	1.19 (1.07-1.32)	.001
1	F5	rs4524	Arg858Lys	LETS	T	708 (80)	671 (74)	1.36 (1.09-1.69)	.006
				MEGA-1		2184 (79)	2608 (74)	1.26 (1.12-1.42)	<.001
1	SMOYKEEBO/F5	rs6016	Ile736Ile	LETS	G	704 (80)	668 (74)	1.35 (1.09-1.69)	.006
				MEGA-1		2188 (79)	2615 (75)	1.27 (1.13-1.43)	< .001
1	C1orf114	rs3820059	Ser172Phe	LETS	A	320 (36)	269 (30)	1.34 (1.10-1.64)	.004
				MEGA-1		1065 (38)	1169 (33)	1.22 (1.10-1.35)	<.001
X	F9	rs6048	Ala194Thr	LETS Men	A	146 (77)	128 (67)	1.74 (1.10-2.74)	.02
				LETS Women		225 (70)	238 (68)	1.09 (0.84-1.42)	.50
				MEGA-1 Men		464 (73)	566 (68)	1.26 (1.00-1.58)	.05
				MEGA-1 Women		674 (72)	818 (71)	1.04 (0.90-1.21)	.61
X	ODZ1	rs2266911	Intronic	LETS Men	C	161 (85)	147 (77)	1.66 (0.99-2.79)	.06
				LETS Women		422 (83)	418 (80)	1.26 (0.91-1.73)	.17
				MEGA-1 Men		556 (85)	671 (80)	1.47 (1.12-1.94)	.006
				MEGA-1 Women		1234 (82)	1430 (78)	1.25 (1.05-1.48)	.01
1	F5 (Leiden)	rs6025	Arg534Gln	LETS	A	95 (11)	14 (2)	7.19 (4.05-(12.77)	<.001
				MEGA-1		291 (10)	96 (3)	4.10 (3.23-5.21)	<.001
11	F2 (G20210A)	rs1799963	3′UTR	LETS	A	28 (3)	10 (1)	2.98 (1.43-6.20)	<.001
				MEGA-1		81 (3)	37 (1)	2.89 (1.94-4.29)	<.001

[a] All gene symbols, rs numbers, SNP types and chromosome numbers are from NCBI build 36.

[b] The first amino acid corresponds to the non risk allele.

[c] ORs were estimated by logistic regression using an additive model. Sex was included as a covariate in logistic regression models containing markers residing on the X chromosome and the number of risk alleles for these SNPs were coded as 0 or 1 for males and 0, 1 or 2 for females.

Table 3. Associations of SNPs From the SNP Association Study With Deep Vein Thrombosis in MEGA-2[a]

				No. (%) of genotypes					
Chromo-some	Gene	SNP	Risk Allele[b]	Genotype[c]	Cased	Controld	OR (95% CI)	P value	FDR[e]
4	CYP4V2	rs13146272	A	CC	121 (10)	352 (13)	1 [Reference]		
				CA	478 (41)	1178 (45)	1.18 (0.94-1.49)		
				AA	561 (48)	1094 (42)	1.49 (1.19-1.88)		
				Additive	(69)	(64)	1.24 (1.11-1.37)	<.001	<.001
1	SERPINC1	rs2227589	T	CC	1001 (77)	2325 (82)	1 [Reference]		
				CT	278 (21)	483 (17)	1.34 (1.13-1.58)		
				TT	15 (1)	28 (1)	1.24 (0.66-2.34)		
				Additive	(12)	(11)	1.29 (1.10-1.49)	<.001	0.004
19	GP6	rs1613662	A	GG	29 (2)	89 (3)	1 [Reference]		
				GA	355 (27)	835 (29)	1.31 (0.84-2.02)		
				AA	915 (70)	1924 (68)	1.46 (0.95-2.24)		
				Additive	(84)	(82)	1.15 (1.01-1.30)	0.03	0.10
1	RGS7	rs670659	C	TT	129 (10)	355 (13)	1 [Reference]		
				TC	615 (48)	1326 (47)	1.28 (1.02-1.60)		
				CC	548 (42)	1153 (41)	1.31 (1.04-1.64)		
				Additive	(66)	(64)	1.10 (1.00-1.22)	0.06	0.13
3	NR1I2	rs1523127	C	AA	480 (37)	1097 (39)	1 [Reference]		
				AC	598 (46)	1340 (47)	1.02(0.88-1.18)		
				CC	220 (17)	409 (14)	1.23 (1.01-1.50)		
				Additive	(40)	(38)	1.09 (0.99-1.20)	0.07	0.13
2	NAT8B	rs2001490	C	GG	490 (38)	1122 (39)	1 [Reference]		
				GC	603 (46)	1334 (47)	1.04 (0.90-1.19)		
				CC	205 (16)	394 (14)	1.19 (0.98-1.45)		
				Additive	(39)	(37)	1.08 (0.98-1.19)	0.12	0.18
X	F9 (men)	rs6048	A	Additive	(73)	(70)	1.17 (0.94-1.45)	0.15	0.20
X	F9 (women)	rs6048	A	GG	56 (8)	148 (10)	1 [Reference]		
				GA	275 (41)	615 (41)	1.18 (0.84-1.66)		
				AA	343 (51)	752 (50)	1.21 (0.86-1.68)		
				Additive	(71)	(70)	1.07 (0.93-1.23)	0.37	NA
19	ZNF544	rs6510130	G	CC	1192 (95)	2626 (95)	1 [Reference]		
				CG	60 (5)	137 (5)	0.97 (0.71-1.32)		
				GG	0 (0)	4 (0)	-		
				Additive	(2)	(3)	0.91 (0.67-1.24)	0.56	0.63
11	MET	rs2237712	G	AA	1183 (93)	2528 (93)	1 [Reference]		
				AG	86 (7)	184 (7)	1.00 (0.77-1.30)		
				GG	3 (0)	3 (0)	2.14 (0.43-10.6)		
				Additive	(4)	(4)	1.03 (0.80-1.33)	0.79	0.79
1	F2	rs1799963	A	GG	1219 (94)	2794 (98)	1 [Reference]		
				GA	76 (6)	55 (2)	3.17 (2.22-4.51)		
				AA	0 (0)	0 (0)	-		
				Additive	(3)	(1)	3.17 (2.22-4.51)	<.001	NA
1	F5	rs6025	A	GG	1029 (81)	2646 (95)	1 [Reference]		
				GA	235 (18)	140 (5)	4.32 (3.46-5.39)		
				AA	8 (0)	2 (0)	10.30 (2.18-48.52)		
				Additive	(10)	(3)	4.24 (3.42-5.26)	<.001	NA

Abbreviations: NA, not applicable, not in FDR analysis.

[a] All gene symbols, rs numbers, SNP types and chromosome numbers are from NCBI build 36.

[b] Risk increasing allele identified in LETS and MEGA-1.

[c] In the additive model, the increase in risk per copy of the risk allele is calculated

[d] For the additive model, only the allele frequency is presented, not the count.

[e] P value from the additive model was used for FDR estimation. Factor V Leiden and the prothrombin 20210A mutation are presented for reference. Because these variants were not included in the SNP association study, we did not calculate their FDR. F9 FDR was calculated for men only

these 7 SNPs ranged from 11% to 82% among the controls. The OR for homozygous carriers, compared with homozygotes of the other allele, ranged from 1.19 to 1.49. The 2 SNPs most strongly associated with DVT were in *CYP4V2* (rs13146272, $P < .001$, false discovery rate 0.0006) and *SERPINC1* (rs2227589, $P < .001$, false discovery rate, 0.004).

For the 2 SNPs on chromosome 1 (rs2227589 and rs670659), we investigated linkage disequilibrium with FVL. The SNP (rs2227589) in *SERPINC1*, which encodes antithrombin, is 4.37 megabases away from the FVL variant. The SNP in *RGS7* (rs670659) is 71.48 megabases from FVL. Each was in weak linkage disequilibrium with FVL (r^2<.01). Restricting analyses to noncarriers of FVL did not appreciably change the risk estimate of either SNP (data not shown).

SNPs in *CYP4V2* Region and DVT Risk

The SNP with the strongest association with DVT was rs13146272, located in the gene encoding a member of the cytochrome P450 family 4 (*CYP4V2*). We genotyped 48 SNPs in this region in the LETS population (Table 4) and estimated the OR for DVT per copy of the risk-increasing allele. For many of the 48 SNPs, including rs13146272, the common allele was the risk allele. In LETS, rs13146272 had an OR for DVT of 1.22 (95% CI, 1.00-1.49). Higher ORs were observed for 9 of the other SNPs tested in this region. These SNPs were located in the *CYP4V2*, *KLKB1* (coding for prekallikrein), and *F11* (coding for coagulation factor XI) genes.

We then selected the 9 of the 48 SNPs that had an OR of more than 1.22 (the OR of rs13146272) and investigated them in MEGA-1. We found that, in addition to rs13146272, four of these SNPs were associated with DVT in both LETS and MEGA-1: rs3087505, rs3756008, rs2036914, and rs4253418 (Table 5). The rs3087505 SNP in *KLKB1* had the highest risk estimate: OR 3.61 (95% CI, 1.48-8.82) for the major allele homozygotes vs minor allele homozygotes. Mutual adjustment among these 5 SNPs did not indicate that any of these 5 associations were explained by the other 4 SNPs (data not shown).

Table 4. 48 SNPs in *CYP4V2* Region Genotyped in the Leiden Thrombophilia Study[a]

rs number	Gene[b]	SNP Type[c]	Risk allele (%) Case	Control	OR (95% CI)[d]	*P*
rs7686244	-	intergenic	39	37	1.11 (0.91 - 1.34)	*0.30*
rs4862650	*DKFZP564J102*	Gly41Lys	11	10	1.10 (0.82 - 1.49)	*0.52*
rs4862653	*DKFZP564J102*	Gly146Lys	11	10	1.09 (0.80 - 1.48)	*0.58*
rs2276922	*DKFZP564J102*	Pro241Leu	27	25	1.10 (0.89 - 1.36)	*0.38*
rs2276921	*DKFZP564J102*	intronic	49	48	1.04 (0.87 - 1.26)	*0.65*
rs2276920	*DKFZP564J102*	intronic	21	19	1.09 (0.86 - 1.38)	*0.47*
rs1877321	*DKFZP564J102*	intronic	79	76	1.19 (0.95 - 1.48)	*0.13*
rs2276919	*DKFZP564J102*	intronic	74	73	1.04 (0.85 - 1.28)	*0.70*
rs13141433	*DKFZP564J102*	intronic	87	85	1.12 (0.87 - 1.45)	*0.37*
rs11733307	*DKFZP564J102*	intronic	46	43	1.13 (0.94 - 1.35)	*0.21*
rs2241818	*DKFZP564J102*	intronic	21	19	1.16 (0.92 - 1.46)	*0.22*
rs6552959	*DKFZP564J102*	intronic	35	33	1.08 (0.89 - 1.30)	*0.46*
rs10017419	-	intergenic	42	41	1.04 (0.86 - 1.24)	*0.70*
rs7676755	*CYP4V2*	intronic	19	18	1.08 (0.84 - 1.38)	*0.55*
rs13146272	*CYP4V2*	Gln259Lys	69	65	1.22(1.00 - 1.49)	*0.05*
rs7687961	*CYP4V2*	intronic	83	81	1.19 (0.93 - 1.51)	*0.16*
rs3817184	*CYP4V2*	splice site	47	44	1.15 (0.96 - 1.39)	*0.13*
rs3736456	*CYP4V2*	Cys282Cys	96	94	1.49 (0.96 - 2.32)	*0.08*
rs2276917	*CYP4V2*	intronic	64	63	1.07 (0.89 - 1.30)	*0.46*
rs3733402	*KLKB1*	Ser143Asn	56	55	1.04 (0.87 - 1.25)	*0.67*
rs4253259	*KLKB1*	intronic	95	94	1.27 (0.84 - 1.92)	*0.25*
rs4253260	*KLKB1*	intronic	85	84	1.14 (0.88 - 1.47)	*0.31*
rs4253301	*KLKB1*	Ala381Ser	89	88	1.15 (0.86 - 1.55)	*0.35*
rs2292423	*KLKB1*	intronic	47	43	1.16 (0.97 - 1.40)	*0.11*
rs3775302	*KLKB1*	intronic	89	86	1.24 (0.95 - 1.64)	*0.12*
rs4253325	*KLKB1*	Gln560Arg	92	89	1.37 (1.01 - 1.86)	*0.04*
rs925453	*KLKB1*	Asn587Asn	71	71	1.00 (0.81 - 1.23)	*0.99*
rs3087505	*KLKB1*	3'UTR	92	90	1.26 (0.91 - 1.76)	*0.17*
rs3822055	*KLKB1*	3'near gene	20	18	1.14 (0.89 - 1.46)	*0.30*
rs6844764	-	intergenic	60	56	1.16 (0.96 - 1.40)	*0.13*
rs13135645	-	intergenic	86	83	1.22 (0.94 - 1.58)	*0.14*
rs3756008	*F11*	5'near gene	47	42	1.22 (1.02 - 1.46)	*0.03*

rs number	Gene[b]	SNP Type[c]	Risk allele (%) Case	Control	OR (95% CI)[d]	*P*
rs3822056	*F11*	5'near gene	90	89	1.07 (0.79 - 1.46)	*0.65*
rs3733403	*F11*	5'near gene	90	89	1.10 (0.81 - 1.50)	*0.53*
rs2036914	*F11*	intronic	60	54	1.25 (1.04 - 1.51)	*0.02*
rs4253408	*F11*	intronic	10	7	1.43 (1.01 - 2.01)	*0.04*
rs1593	*F11*	intronic	90	88	1.16 (0.85 - 1.57)	*0.36*
rs4253414	*F11*	intronic	3	3	1.07 (0.61 - 1.87)	*0.82*
rs4253418	*F11*	intronic	96	95	1.42 (0.88 - 2.27)	*0.15*
rs5974	*F11*	Thr267Thr	87	86	1.01 (0.77 - 1.32)	*0.93*
rs4253423	*F11*	intronic	85	83	1.13 (0.88 - 1.46)	*0.33*
rs5971	*F11*	Arg604Arg	96	96	1.01 (0.64 - 1.59)	*0.97*
rs4253430	*F11*	3'near gene	67	65	1.13 (0.92 - 1.37)	*0.24*
rs11938564	-	intergenic	81	79	1.14 (0.91 - 1.43)	*0.27*
rs13136269	-	intergenic	76	73	1.18 (0.95 - 1.46)	*0.13*
rs10025152	-	intergenic	85	85	1.01 (0.78 - 1.32)	*0.93*
rs12500826	-	intergenic	67	64	1.12 (0.93 - 1.36)	*0.24*
rs13133050	-	intergenic	71	68	1.13 (0.93 - 1.38)	*0.22*

[a] Gene symbols, rs numbers, SNP types, and chromosome numbers are from National Center for Biotechnology Information build 36.

[b] Some SNPs were located between genes, indicated in the "SNP type" column as "intergenic."

[c] The first amino acid corresponds to the nonrisk allele.

[d] Odds ratios were estimated by logistic regression using an additive model.

SNPs in *CYP4V2* Region and Factor XI Levels

Because the *F11* gene is located close to rs13146272 and because factor XI levels have been previously reported to be associated with DVT in the LETS population [171], we investigated whether an association between SNPs and factor XI levels explained the association between the SNPs and DVT. In LETS, factor XI levels above the 90th percentile had been shown to be associated with a 2-fold increased risk of DVT [171]. We found that high factor XI levels (>90th percentile) were also associated with DVT in MEGA (OR, 1.9; 95% CI, 1.6-2.3).

The 5 SNPs from the *CYP4V2* region that were associated with DVT were all associated with factor XI levels in LETS and MEGA-1, with higher factor XI levels for those who carried the risk-increasing alleles (Table 5). We investigated whether factor XI levels mediate the association between these 5 SNPs and DVT by adjusting for factor XI levels in the combined LETS and MEGA-1 studies. For all 5 SNPs, adjustment for factor XI levels weakened the association with DVT but none of the associations disappeared. Interestingly, the 5 SNPs that were not associated with DVT in the combined analysis of LETS and MEGA-1 (rs3736456, rs4253259, rs4253408, rs4253325, and rs3775302) were also not associated with factor XI levels in LETS.

All analyses were performed with and without adjustment for age and sex, and analyses in MEGA-1 and MEGA-2 were performed with and without restriction to the group with both parents born in northwestern Europe. Because neither influenced the results, we presented the unadjusted OR.

COMMENT

We identified 7 SNPs that were associated with DVT in 3 large, well-characterized populations including 3155 cases and 5087 controls. The evidence was strongest for the 3 SNPs in the *CYP4V2, SERPINC1*, and *GP6* genes. It is interesting to note that these SNPs are in or near genes that have a clear role in blood coagulation. This may indicate that the coagulation system is well characterized.

Testing 19 682 SNPs will result in false-positive associations. Therefore, we investigated the SNPs in 3 large studies and estimated the false discovery rate for the SNPs tested in the third study. The 3 SNPs in genes *CYP4V2*, *SERPINC1*, and *GP6* were associated with DVT with a false discovery rate of less than 10%, which means that less than 10% of these 3 SNPs would be expected to be false positive. Relaxing the false discovery rate to less than 20% would add 4 SNPs, in *RGS7, NR1I2, NAT8B,* and *F9* as associated with DVT.

Table 5. Association of 10 SNPs in *CYP4V2* Region With Deep Vein Thrombosis and Factor XI Levels in the Combined LETS and MEGA-1 Studies.

SNP	Gene	Risk Allele	Genotype	Risk allele, No. (%) Case	Control	Factor XI[a] % Difference (95% CI)	Deep Vein Thrombosis OR (95% CI)	OR[b] (95% CI)
rs13146272	*CYP4V2*	A	CC	181 (10)	296 (13)	[Reference]	1 [Reference]	1 [Reference]
			CA	808 (44)	995 (45)	3 (1 to 6)	1.32 (1.07 to 1.62)	1.26 (1.03 to 1.56)
			AA	850 (46)	919 (42)	7 (4 to 9)	1.50 (1.22 to 1.84)	1.36 (1.10 to 1.68)
			Additive	(68)	(64)	3 (2 to 4)	1.20 (1.09 to 1.31)	1.14(1.04 to 1.25)
rs3736456	*CYP4V2*	T	CC	7 (0)	0 (0)			
			CT	163 (9)	222 (10)	[Reference]	1 [Reference]	1 [Reference]
			TT	1663 (91)	1973 (90)	1 (-1 to 4)	1.16 (0.93 to 1.43)	1.15 (0.93 to 1.42)
			Additive	(95)	(95)	1 (-2 to 4)	1.06 (0.86 to 1.30)	1.05 (0.85 to 1.28)
rs3087505	*KLKB1*	C	TT	6 (0)	25 (1)	[Reference]	1 [Reference]	1 [Reference]
			TC	317 (17)	438 (20)	11 (6 to 16)	3.02 (1.22 to 7.44)	2.59 (1.05 to 6.40)
			CC	1509 (82)	1743 (79)	19 (14 to 24)	3.61 (1.48 to 8.82)	2.81 (1.15 to 6.89)
			Additive	(91)	(89)	8 (6 to 10)	1.27 (1.09 to 1.47)	1.15 (0.99 to 1.34)
rs4253259	*KLKB1*	C	AA	5 (0)	6 (0)	[Reference]	1 [Reference]	1 [Reference]
			AC	168 (9)	219 (10)	0 (-18 to 18)	0.92 (0.28 to 3.07)	0.95 (0.28 to 3.20)
			CC	1652 (91)	1978 (90)	0 (-19 to 18)	1.00 (0.31 to 3.29)	1.03 (0.31 to 3.43)
			Additive	(95)	(95)	0 (-3 to 2)	1.08 (0.88 to 1.32)	1.08 (0.88 to 1.32)
rs4253408	*F11*	A	GG	1526 (83)	1869 (85)	[Reference]	1 [Reference]	1 [Reference]
			GA	293 (16)	317 (14)	4 (2 to 6)	1.13 (0.95 to 1.35)	1.06 (0.89 to 1.27)
			AA	15 (2)	17 (1)	13 (-1 to 27)	1.08 (0.54 to 2.17)	0.97 (0.48 to 1.98)
			Additive	(9)	(8)	5 (3 to 7)	1.11 (0.95 to 1.30)	1.05 (0.89 to 1.23)
rs4253325	*F11*	G	AA	21 (1)	23 (1)	[Reference]	1 [Reference]	1 [Reference]
			AG	308 (17)	392 (18)	5 (0 to 16)	0.86 (0.47 to 1.58)	0.86 (0.46 to 1.59)
			GG	1507 (82)	1785 (81)	8 (-3 to 13)	0.93 (0.51 to 1.68)	0.88 (0.48 to 1.62)
			Additive	(90)	(90)	3 (1 to 5)	1.05 (0.91 to 1.21)	1.01 (0.87 to 1.17)
rs3775302	*KLKB1*	A	GG	1418 (77)	1686 (77)	[Reference]	1 [Reference]	1 [Reference]
			GA	380 (21)	481 (22)	-1 (-3 to 1)	0.94 (0.81 to 1.09)	0.97 (0.83 to 1.13)
			AA	38 (2)	34 (2)	-4 (-11 to 3)	1.33 (0.83 to 2.12)	1.34 (0.83 to 2.14)
			Additive	(12)	(12)	-1 (-3 to 0)	1.00 (0.87 to 1.14)	1.02 (0.89 to 1.16)
rs3756008	*F11*	T	AA	526 (29)	788 (36)	[Reference]	1 [Reference]	1 [Reference]
			AT	903 (49)	1032 (47)	7 (6 to 9)	1.31 (1.14 to 1.51)	1.21 (1.04 to 1.39)
			TT	408 (22)	384 (17)	15 (12-17)	1.59 (1.33 to 1.90)	1.32 (1.09 to 1.59)
			Additive	(47)	(41)	7 (6 to 8)	1.27 (1.16 to 1.38)	1.16 (1.05 to 1.27)

SNP	Gene	Risk Allele	Genotype	Risk allele, No. (%) Case	Control	Factor XI[a] % Difference (95% CI)	Deep Vein Thrombosis OR (95% CI)	OR[b] (95% CI)
rs2036914	*F11*	C	TT	302 (17)	505 (23)	[Reference]	1 [Reference]	1 [Reference]
			TC	895 (49)	1081 (49)	7 (5 to 9)	1.38 (1.17 to 1.64)	1.27 (1.07 to 1.51)
			CC	633 (35)	620 (28)	14 (12 to 16)	1.71 (1.43 to 2.05)	1.43 (1.19 to 1.73)
			Additive	(59)	(53)	7 (6 to 8)	1.30 (1.19 to 1.42)	1.19 (1.08 to 1.30)
rs4253418	*F11*	G	AA	3 (0)	4 (0)	[Reference]	1 [Reference]	1 [Reference]
			AG	120 (7)	199 (9)	14 (10 to 19)	0.80 (0.18 to 3.65)	0.69 (0.15 to 3.14)
			GG	1710 (93)	2000 (91)	22 (18 to 26)	1.14 (0.26 to 5.10)	0.88 (0.20 to 3.94)
			Additive	(97)	(95)	8 (5 to 11)	1.39 (1.11 to 1.74)	1.24 (0.99 to 1.56)

The 3 SNPs with the strongest evidence for association with DVT were in the genes *CYP4V2*, *SERPINC1*, and *GP6*. The *CYP4V2* gene encodes a member of the CYP450 family 4 that is not known to be related to thrombosis [174,175]. The *CYP4V2* gene is located on chromosome 4 in a region containing genes encoding coagulation proteins prekallikrein (*KLKB1*) and factor XI (*F11*). We also found 4 other SNPs in the *CYPV42/KLKB1/F11* locus that were associated with DVT. No previous reports exist of genetic variants in *CYP4V2* and *KLKB1* and their association with DVT. There exists no evidence for an association between prekallikrein levels and DVT [176], while there is evidence for elevated factor XI levels [160,171]. It remains unclear whether only one of these SNPs, or all of them affect DVT risk.

The *SERPINC1* gene encodes antithrombin, a serine protease inhibitor located on chromosome 1 that plays a central role in natural anticoagulation. Deficiencies of antithrombin are rare but result in a strong thrombotic tendency [177]. The SNP in *SERPINC1* (rs2227589) had a minor allele frequency of about 10% in the controls and was associated with a modest thrombotic tendency. The *GP6* gene encodes glycoprotein VI, a 58-kDa platelet membrane glycoprotein that plays a crucial role in the collagen-induced activation and aggregation of platelets [178] and may play a role in DVT [179].

The SNPs in the genes *F9*, *NR1I2*, *RGS7*, and *NAT8B* are of interest for further validation. The *F9* gene encodes factor IX, a vitamin K–dependent coagulation factor, of which high levels have been shown to increase the risk of DVT [180]. The SNP rs6048, also known as *F9* Malmö, is a common polymorphism at the third amino acid residue of the activation peptide of factor IX [181].

The SNP in *CYP4V2* (rs13146272) is located close to the gene encoding coagulation factor XI. Factor XI levels have been reported to be associated with DVT in LETS [171] and in a large analysis of pedigrees [160]. We confirmed the association between DVT and factor XI levels in MEGA. Interestingly, the 5 SNPs in the *CYP4V2* region that were associated with DVT in both LETS and MEGA-1 were also associated with factor XI levels. However, the association between these 5 SNPs and DVT does not seem to be completely explained by variation in factor XI levels because adjusting for factor XI level did not remove the excess DVT risk of these 5 SNPs. Thus, if only part of the risk associated with these genetic variants is mediated through levels of factor XI, some of the risk might also be due to effects on protein function.

Several variants in the *F11* gene (rs5974, rs5970, rs5971, rs5966, rs5976, and rs5973) were previously tested for association with factor XI levels in patients with DVT and atherosclerosis, but no relationship was observed [182]. In the present study, rs5974 (r^2 = 1.0 with 5970) and rs5971 (r^2 = 1.0 with 5966 and rs5976) were not associated with DVT in LETS. Rs5973 was not genotyped because its minor allele frequency was lower than 2% (HapMap CEPH population). In a study of West African volunteers [129], rs3822056 and rs3733403 were associated with transcription factor binding affinity and slightly increased factor XI levels, but neither SNP was associated with DVT in LETS. In a study among white postmenopausal women [183], rs3822057 and rs2289252 were associated with DVT. Both of these associations were indirectly confirmed in the present study because 2 of the 5 SNPs in the *CYPV42* region that were consistently associated with DVT and factor XI levels are in linkage disequilibrium with rs3822057 (r^2 = 0.9 with rs2036914) and rs2289252 (r^2 = 0.8 with rs3756008).

The association between genetic variants and DVT may depend on clinical variables or other risk factors for DVT, such as surgery or the use of oral contraceptives. Because we aimed to identify variants that are associated with DVT in general and from a large set of SNPs, we did not study subgroups. Clinical utility, however, may well depend on interaction with these clinical variables and should form a focus of subsequent studies.

The associations between SNPs and DVT were modest, for instance homozygous carriership of the AA genotype of rs13146272 in *CYP4V2* increased risk 1.49-fold. However, because the variants are common, they might be useful risk indicators especially when combined with other risk factors. Moreover, the associations found might represent a diluted effect of an unmeasured SNP in linkage disequilibrium or indicate a region with several variants involved in DVT susceptibility. The results from the *CYP4V2* region illustrate the need for further study, because some ORs found in that region were higher than initially found for rs13146272.

Only 9 of 18 stage 3 SNPs were indeed tested in MEGA-2 DNA in stage 4. The reason for this was that in order to save MEGA-2 DNA, stage 4 SNPs were genotyped using multiplexed oligoligation assays, and assays for only 9 stage 3 SNPs were available at the time of this study. Therefore, a future extension of this study may yield additional SNPs associated with DVT.

The replication criteria that we used to identify SNPs associated with DVT may have caused us to miss some truly associated positive variants. Although the statistical power to detect associations between DVT and uncommon genetic variation was high, a rare variant with a modest association with DVT may have been missed.

Our analysis was limited to a northwestern European population. Confounding in a genetic study may arise from population stratification, ie, the presence of ethnic groups with different allele and disease frequencies within a study. In LETS, no information on ethnicity was collected. However,