Genetic determinants of venous thrombosis

Haan, H.G. de

CHAPTER 5

Genetic variants in Cell Adhesion Molecule 1 (CADM1): a validation study of a novel endothelial cell venous thrombosis risk factor

de Haan HG, Bezemer ID, Vossen CY, van Hylckama Vlieg A, Böehringer S, Hasstedt SJ, Levy S, Rosendaal FR, Bovill EG.

Thromb Res. 2014;134(6):1186-92.

ABSTRACT

Introduction

In a protein C deficient family, we recently identified a candidate gene, *CADM1*, which interacted with protein C deficiency in increasing the risk of venous thrombosis (VT). This study aimed to determine whether *CADM1* variants also interact with protein C pathway abnormalities in increasing VT risk outside this family.

Materials and methods

We genotyped over 300 *CADM1* variants in the population-based MEGA case-control study. We compared VT risks between cases with low protein C activity (N=194), low protein S levels (N=23), high factor VIII activity (N=165) or factor V Leiden carriers (N=580), and all 4004 controls. Positive associations were repeated in all 3496 cases and 4004 controls.

Results

We found 22 variants which were associated with VT in one of the protein C pathway risk groups. After mutual adjustment, six variants remained associated with VT. The strongest evidence was found for rs220842 and rs11608105. For rs220842, the odds ratio (OR) for VT was 3.2 (95% CI 1.2-9.0) for cases with high factor VIII activity compared with controls. In addition, this variant was associated with an increased risk of VT in the overall study population (OR 1.5, 95% CI 1.0-2.2). The other variant, rs11608105, was not associated with VT in the overall study population (OR 1.0, 95% CI 0.8-1.1), but showed a strong effect on VT risk (OR 21, 95% CI 5.1-88) when combined with low protein C or S levels.

Conclusions

In a population-based association study, we confirm a role for *CADM1* variants in increasing the risk of VT by interaction with protein C pathway abnormalities.

INTRODUCTION

We have identified a candidate gene, cell adhesion molecule 1 (*CADM1*), which appears to interact with protein C deficiency to increase the risk of venous thrombosis in an extended French Canadian family with type I protein C deficiency due to a *PROC* 3363C insertion ("Vermont family").[1] The 300kb *CADM1* gene is also known as nectin-like protein 2 (*NECL2*), tumor suppressor in lung cancer 1 (*TSLC1*), synapse cell adhesion molecule (*SynCAM1*), spermatogenic immunoglobulin super family (*SgIGSF*), and immunoglobulin super family 4 (*IGSF4*).[2-6] CADM1, an immunoglobulin cell adhesion molecule involved in binding interactions supporting intercellular adhesion, has been best characterized as a constitutive cell-cell adhesion molecule in epithelial cells and at neuronal synapses.[4,5]

In the Vermont family study, several single nucleotide variants (SNVs) in *CADM1* showed a strong association with venous thrombosis in interaction with protein C deficiency.[1] For example, among protein C deficient family members, carriers of the rs6589488 minor allele had a 17-fold increased risk of venous thrombosis (OR 17, 95% CI 13.5-21.4) compared with homozygous major allele carriers. Subsequent *CADM1* gene expression assays, using blood outgrowth endothelial cells cultured from family members, showed a decreased expression compared with controls, lending phenotypic support to the SNV associations. We also demonstrated *CADM1* in endothelial cells, where it appears to be selectively involved in endothelial cell migration, suggesting a role in maintenance of endothelial barrier function.[1,7]

Activated Protein C, bound to the endothelial protein C receptor (APC-EPCR) on the endothelial membrane, mediates endothelial barrier enhancement through activation of protease activated receptor 1 (PAR-1) and the sphingosine-1-phosphate-receptor-1 (S_1P_1) pathways.[8-12] This APC-EPCR mediated activation of PAR-1 and S_1P_1 leads to activation of endothelial Rac1 and the cytoskeletal rearrangements associated with endothelial barrier enhancement.[10,11,13] The *CADM1* pathway,[14] which is associated with migration and adhesion in epithelial cells, appears to mediate this epithelial cell behavior, in part, through regulating small Rho-GTPases including Rac1.[15,16] This suggests that our observation of a strong interaction between the *CADM1* and protein C genes in increasing thrombosis risk in the Vermont family may be related to a shared common signalling pathway involving the small Rho-GTPases. Thus, the *CADM1* pathway interaction with the protein C system may represent a novel biological pathway

conferring increased risk for venous thrombosis at the level of the vessel wall due to impaired maintenance of endothelial barrier function.

In order to validate the association between *CADM1* and thrombosis observed in the Vermont family study, we investigated *CADM1* gene variants in the Multiple Environmental and Genetic Assessment of risk factors for venous thrombosis (MEGA study), a case-control study on venous thrombosis including over 4000 patients and 4000 controls. To study the effect of *CADM1* variants on thrombosis risk, we primarily focused on subsets of thrombosis patients with protein C pathway abnormalities (i.e. low levels of protein C or S, high factor VIII levels, and the factor V Leiden variant) as *CADM1* variants were found to interact with protein C deficiency in the Vermont family study.[1] Protein S interacts closely with protein C in the inactivation of the procoagulant factors Va and VIIIa,[17] and synergistic effects of *CADM1* with protein C deficiency might therefore also occur with protein S deficiency, high levels of factor VIII, or activated protein C resistance due to factor V Leiden (*F5*, rs6025).

MATERIAL AND METHODS

Study population

The MEGA study is a population-based case-control study.[18,19] Consecutive patients aged 18 to 70 years with a first venous thrombosis of the leg or arm, or with a pulmonary embolism were recruited from 6 anticoagulation clinics in the western part of the Netherlands between 1999 and 2004. Partners of patients, as well as additional individuals recruited by random digit dialling and frequency-matched on age and sex, were invited as control subjects. All participants received a standardized questionnaire about risk factors for venous thrombosis. A blood sample was taken approximately 3 months after discontinuation of anticoagulant therapy (usually 3-12 months after the diagnosis of venous thrombosis), or after a year when patients continued their anticoagulant therapy, and from control subjects. Participants who refused to or were unable to provide a blood sample and patients and their partners included after June 1, 2002 were offered the option of providing a buccal swab sample for DNA. Exclusion criteria were previous venous thrombosis (patients and controls), no venous thrombosis (patients, after checking hospital records), age younger than 18 or older than 70, severe psychiatric problems, inability to speak Dutch and, for genetic and

blood sample analysis, poor sample quality. For the present analysis, we only included individuals from North- or Western European origin (90%), which was assessed by self-reported country of birth of the parents, in order to avoid population stratification. This left 1970 patients and 2490 control subjects (N=4460) with a plasma and DNA sample and another 1526 patients and 1514 control subjects (N=3040) with only a DNA sample eligible for analysis.

Protein C pathway abnormalities

We selected individuals with protein C pathway abnormalities, i.e., low protein C activity, low protein S levels, high factor VIII activity levels, or factor V Leiden carriership. The protein C, protein S and factor VIII abnormalities were not individually diagnosed, but instead we used clinically relevant cut-off levels to categorize individuals as abnormal. Low protein C activity was defined by taking the lower limit of normal (67% of normal in our laboratory) as cut-off point. When individuals were on oral anticoagulant therapy at time of blood draw, we calculated the expected protein C activity relative to factor VII activity by linear regression according to a method described by O'Brien et al.[20] The observed levels were classified as "low" when the observed/expected ratio was below the geometric mean minus 2 standard deviations as calculated among control subjects. Of 1959 patients and 2471 control subjects with protein C (and factor VII) measurements, 194 patients (10%; mean protein C activity 43% of normal; range 19-66) and 28 control subjects (1%; mean protein C activity 42% of normal; range 30-62) had low protein C activity. Of these 194 patients and 28 control subjects, 178 patients and 21 controls were on oral anticoagulant therapy at the time of the blood draw.

Similarly to the selection of individuals with low protein C, we selected low protein S individuals by selecting total protein S levels below the lower limit of normal (67% of normal) for individuals not on oral anticoagulant therapy at the time of the blood draw and calculated protein S levels relative to factor II for patients using oral anticoagulant therapy at the time of the blood draw. Of the 1828 patients and 2252 control subjects with protein S (and factor II) measurements, 23 patients (1%; mean protein S level 58% of normal; range 32-66) and 26 controls (1%; mean protein S level 60% of normal; range 45-67) had low protein S levels. Of these 33 patients and 28 controls, 3 patients and none of the controls were on oral anticoagulant therapy at the time of the blood draw.

High factor VIII was defined as activity levels higher than the geometric mean plus 2 standard deviations as calculated among control subjects, which was 204 IU/ml. In total, 165 (8%) of 1969 patients and 51 (2%) of 2488 control subjects with factor VIII levels available had high factor VIII activity levels.

For the factor V Leiden subgroup analysis, we selected 580 (17%) patients and 219 (5%) control subjects who carried the variant from among 3493 patients and 4000 control subjects with factor V Leiden genotypes available.

Laboratory analysis

Collection and processing of blood and buccal swab samples, subsequent DNA isolation and genotyping of factor V Leiden variant have been described previously.[18] Measurements of protein C activity were performed with a chromogenic assay and factor II, VII and VIII activity measurements were based on clotting time assays using immune-depleted plasma, deficient for the factor under study. 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 in our laboratory were 1.4% and 3.5%, respectively, for protein C, 2.7% and 4.2% for factor II, 3.4% and 4.0% for factor VII, 3.6% and 8.9% for factor VIII and 5.0% and 3.5% for protein S. All measurements were performed on a single blood draw.

SNV Selection

We selected 364 SNVs throughout *CADM1* and 2kb downstream and 10kb upstream of the gene in order to include conserved elements which may play a regulatory role (chr11:114,543,000-114,893,000, NCBI B36 assembly). From the *CADM1* SNVs that were genotyped in the European HapMap population, we chose 86 tagging SNVs with minor allele frequency (MAF)>0.01 by pairwise tagging (r^2>0.8) as implemented in Haploview.[21] From the HapMap list we added 42 SNVs from blocks with multiple SNVs for redundancy and 29 SNVs in regions where the distance between adjacent SNVs was largest. In addition, we selected 99 SNVs that had not been genotyped by HapMap but were validated in dbSNP and 108 SNVs that we identified by resequencing the region in the Vermont family. Of 364 SNVs selected for genotyping, 47 were excluded because of poor assay performance, 3 SNV assays were excluded because of atypical clustering

and 30 were not polymorphic in the MEGA study population, which left 284 SNVs for statistical analysis. Genotyping was performed at the Johns Hopkins University through the NHLBI Genotyping and Resequencing Service. Genotyping quality was assessed by establishing the call rate (>99%) and the Hardy-Weinberg equilibrium of each SNV.

Statistical analysis

The primary analysis was to compare allele frequencies between patients with specific abnormalities in the protein C pathway (i.e. low protein C, low protein S, high factor VIII or factor V Leiden) and all control subjects. The choice for taking all control subjects as a reference group was made because few control subjects had low protein C activity or low protein S levels.

Odds ratios (OR) and 95% confidence intervals (95% CI) were computed using logistic regression for an additive genetic model. The reference allele was the most prevalent (major) allele in the total study population and the OR was calculated per additional minor allele copy. Variants that were associated with venous thrombosis in the primary analysis (one of the subgroups of protein C pathway abnormalities versus all controls) with p-value <0.05 were further studied. Next, linkage disequilibrium (LD) between SNVs of interest was studied in Haploview.[21] Of the variants that were in strong linkage disequilibrium, defined as r^2 of 0.7 or higher, we selected the variant with the highest allele frequency in controls for follow-up. To assess the causal effects of the SNVs, we mutually adjusted the associations by entering all positive variants into a conditional logistic regression model. Positive associations were repeated in the overall MEGA study (3496 cases and 4004 controls) and studied for the joint effects of the variants and the protein C pathway abnormality under study.

With more than 250 variants tested for association with venous thrombosis in each subgroup, the chance of false positive findings is substantial. In order to decrease the chance of false-positive reporting, we calculated an FDR-adjusted q-value.[22]

RESULTS

Characteristics of the study population are presented in Table 1. We studied 284 variants in four subgroups of venous thrombosis patients with a protein C pathway abnormality,

i.e., patients with low protein C activity (N=194), patients with low protein S levels (N=23), patients with high FVIII activity (N=165) and patients carrying the FV Leiden polymorphism (N=580), and all controls (N=4004). The subgroups were not mutually exclusive, i.e., 72 patients (12%) had multiple abnormalities in the protein C pathway.

Table 1. Characteristics of the MEGA study population.

	Patients (N=3496)	**Controls (N=4004)**
Men (%)	1633 (46.7)	1892 (47.3)
Mean age (SD)	49.18 (12.81)	48.40 (12.36)
FVL carrier (%)	580 (16.60)	219 (5.48)
Plasma available	1970	2490
Low protein C (%)	194 (9.90)	28 (1.13)
Low protein S (%)	23 (1.26)	26 (1.15)
High factor VIII (%)	165 (8.38)	51 (2.05)

SD standard deviation; FVL Factor V Leiden
Low protein C was defined as activity levels below 67% of normal or when on anticoagulant treatment relative to factor VII (see Methods). Similarly, low protein S was defined as activity levels below 67% of normal or when on anticoagulant treatment relative to factor II (see Methods). High factor VIII was defined as activity levels higher than the geometric mean plus two standard deviations among controls (see Methods).

Associations between CADM1 variants and VT within protein C pathway subgroups

For all 284 variants, allele frequencies among all MEGA study patients and all MEGA study controls are listed in Supplemental Table 1. Twelve of the 284 variants were monomorphic among control subjects and eight were monomorphic among patients of the overall MEGA study. In addition, several variants were monomorphic in one of the subgroups of patients with a protein C pathway abnormality: 16 variants among patients with low protein C activity, 46 variants among patients with low protein S levels, 14 variants among patients with high factor VIII activity and 17 variants among patients carrying factor V Leiden. These variants could not be studied.

During the first stage of the analysis, we identified 22 *CADM1* variants that were associated with venous thrombosis (p-value<0.05) in one of the subgroups of patients with a protein C pathway abnormality and all controls (Table 2). One variant was associated with venous thrombosis in the low protein C subgroup, nine variants in the low protein S subgroup, six variants in the high factor VIII subgroup, and seven variants in the factor V Leiden subgroup (Table 2). Only one variant (rs11608105) was associated

with venous thrombosis in multiple subgroups, i.e. the low protein C subgroup (OR 1.57, 95% CI 1.05-2.34) and the low protein S subgroup (OR 2.98, 95% CI 1.27-7.02). To correct for multiple testing, we calculated FDR-adjusted q-values after which none of the variants remained associated with venous thrombosis (Table 2).

Table 2. Associations with venous thrombosis in the different subgroups.

	Risk allele frequency, %					FDR
	Patients	Controls	OR	95% CI	p-value	q-value
Low protein C patients						
rs11608105	7.22	4.72	1.57	1.05-2.34	0.026	1
Low protein S patients						
rs4938182	32.6	19.8	1.95	1.05-3.63	0.034	0.756
rs4450197	8.70	2.04	4.95	1.67-14.7	0.004	0.333
rs10128746	13.0	3.63	4.40	1.75-11.1	0.002	0.333
rs11215418	10.9	3.62	3.37	1.29-8.83	0.013	0.371
rs45595941	4.35	0.70	6.71	1.54-29.3	0.011	0.371
rs45616036	4.35	0.84	5.03	1.25-20.3	0.023	0.575
rs11608105	13.0	4.72	2.98	1.27-7.02	0.013	0.371
rs45520832	2.17	0.11	20.1	2.45-166	0.005	0.333
rs45583332	4.35	0.70	6.71	1.54-29.3	0.011	0.371
High factor VIII patients						
rs10891823	9.47	6.48	1.48	1.02-2.16	0.040	0.999
rs11215504	7.58	4.35	1.79	1.18-2.73	0.006	0.750
rs11215515	7.10	4.26	1.75	1.14-2.68	0.010	0.833
rs11215458	5.62	3.65	1.61	1.00-2.60	0.050	0.999
rs220842	1.52	0.51	3.02	1.18-7.74	0.022	0.999
rs10891856	9.47	5.75	1.75	1.21-2.55	0.003	0.750
Factor V Leiden patients						
rs12577709	15.9	13.6	1.19	1.01-1.42	0.041	0.988
rs45545346	1.73	3.42	0.50	0.32-0.79	0.003	0.741
rs45608938	3.89	5.42	0.71	0.52-0.97	0.032	0.988
rs17443832	3.97	5.38	0.73	0.54-0.99	0.045	0.988
rs45578937	5.10	7.07	0.71	0.54-0.93	0.014	0.865
rs45458294	4.84	6.92	0.68	0.52-0.91	0.008	0.741
rs314497	7.84	5.86	1.37	1.08-1.73	0.009	0.741

OR odds ratio; CI confidence interval; FDR false discovery rate
In the univariable analysis, 22 variants were associated with venous thrombosis. The risk allele frequency was calculated in the subgroup of cases in which the variant was identified and in the overall controls.

Next, we studied linkage disequilibrium between the positive variants. Of the 22 variants, four pairs of variants were in strong linkage disequilibrium (Figure 1; $r^2 \geq 0.7$). Of each pair of variants, the variant having the highest risk allele frequency among controls was selected for the remaining analyses. To study the causal effects of the positive variants on venous thrombosis, we entered all positive variants within each subgroup in a logistic regression model. In the subgroup of protein S, two variants remained associated with venous thrombosis, i.e., rs11608105 and rs45520832 (Table 3; OR 3.54, 95% CI 1.46-8.60 and OR 22.1, 95% CI 2.35-208, respectively). In addition, two variants, i.e., rs11215504 and rs220842, remained associated with venous thrombosis in patients with high factor VIII activity (OR 1.89, 95% CI 1.24-2.88 and OR 3.23, 95% CI 1.17-8.97, respectively). In the patients that carried FV Leiden, another two variants, i.e., rs45608938 and rs45545346, remained associated with a decreased risk of venous thrombosis (Table 3; OR 0.71, 95% CI 0.52-0.97 and OR 0.53, 95% CI 0.30-0.93 respectively).

Table 3. Mutually adjusted associations with venous thrombosis in the different subgroups.

	Risk allele frequency, %			
	Patients	**Controls**	**OR**	**95% CI**
Low protein C patients				
rs11608105	7.22	4.72	1.57	1.05-2.34
Low protein S patients				
rs4938182	32.6	19.8	1.60	0.79-3.22
rs4450197	8.70	2.04	1.22	0.21-7.25
rs10128746	13.0	3.63	3.04	0.72-13.0
rs45616036	4.35	0.84	1.92	0.07-51.9
rs11608105	13.0	4.72	3.54	1.46-8.60
rs45520832	2.17	0.11	22.1	2.35-208
rs45583332	4.35	0.70	4.27	0.15-124
High factor VIII patients				
rs10891823	9.47	6.48	1.13	0.59-2.14
rs11215504	7.58	4.35	1.89	1.24-2.88
rs11215515	7.10	4.26	1.16	0.56-2.40
rs11215458	5.62	3.65	1.03	0.42-2.50
rs220842	1.52	0.51	3.23	1.17-8.97
rs10891856	9.47	5.75	1.60	0.89-2.85

Table 3. Continued

	Risk allele frequency, %			
	Patients	**Controls**	**OR**	**95% CI**
Factor V Leiden patients				
rs12577709	15.9	13.6	1.09	0.90-1.32
rs45545346	1.73	3.42	0.53	0.30-0.93
rs45608938	3.89	5.42	0.71	0.52-0.97
rs45578937	5.10	7.07	0.93	0.66-1.30
rs314497	7.84	5.86	1.27	0.98-1.65

OR odds ratio; CI confidence interval
When including the positive associations per subgroup together in a logistic regression model, six variants remained associated with venous thrombosis.

Associations to venous thrombosis in overall MEGA study

We further investigated the six variants, which remained associated with venous thrombosis after mutual adjustment, in the overall MEGA study population in order to study the effect on venous thrombosis independently of the protein C pathway abnormalities. We observed a weak association between rs220842 and venous thrombosis (OR 1.49, 95% CI 0.99-2.24) and between rs11215504 and venous thrombosis (OR 1.14, 95% CI 0.98-1.33). The other four variants were not associated with venous thrombosis in the overall MEGA study population (Table 4).

Table 4. Associations with venous thrombosis in MEGA overall study population.

	Risk allele frequency, %			
***CADM1* variants**	**Patients**	**Controls**	**OR**	**95% CI**
rs11608105	4.57	4.72	0.97	0.83-1.13
rs45520832	0.14	0.11	1.27	0.52-3.13
rs11215504	4.94	4.35	1.14	0.98-1.33
rs220842	0.76	0.51	1.49	0.99-2.24
rs45608938	5.33	5.42	0.98	0.85-1.13
rs45545346	3.13	3.42	0.92	0.77-1.09

OR odds ratio; CI confidence interval

Joint effect of CADM1 variants and protein C pathway abnormalities

We studied the joint effect of the thrombosis associated variants and the protein C pathway abnormalities by using homozygous major allele carriers without the protein C pathway abnormality under study as a reference for the odds ratio (Table 5). The combination of carrying variant rs11608105 and having low protein C or protein S levels was associated with a 21-fold increased risk (95% CI 5.08-88.8) of venous thrombosis. Compared with non-carriers having low protein C or S levels, the risk of venous thrombosis was a 4-fold increased (95% CI 1.00-18.7) in carriers of the risk allele with low protein C or S levels.

Similar to findings in the overall MEGA study population, variant rs220842 was associated with an increased risk of venous thrombosis (OR 1.88, 95% CI 1.07-3.31; Table 5) in individuals without high factor VIII activity. The joint effect of the variant and high factor VIII activity could not be studied as only patients and no controls with high factor VIII activity carried the variant (N=5; Table 5). Furthermore, having high factor VIII activity and carrying the risk allele of variant rs11215504 was associated with a 6.5-fold increased risk of venous thrombosis. This exceeded the risk for rs11215504 or the defect alone (Table 5) albeit with a wide confidence interval due to the small number of carriers with also a defect (95% CI 2.48-17.1). For the other positive variants, no clear joint effect with a protein C pathway abnormality could be calculated (rs45520832) or was observed (rs45608938, rs45545346) (Table 5).

Table 5. Combined associations for *CADM1* SNVs with protein C pathway abnormalities and venous thrombosis.

***CADM1* variants**		**Pathway defect**		**Patients, N**	**Controls, N**	**OR**	**95% CI**
rs11608105	No	PC/PS	No	1503	1996	1(REF)	
rs11608105	Yes	PC/PS	No	123	206	0.79	0.63-1.00
rs11608105	No	PC/PS	Yes	181	49	4.91	3.55-6.77
rs11608105	Yes	PC/PS	Yes	32	2	21.3	5.08-88.8
rs45520832	No	PS	No	1798	2219	1(REF)	
rs45520832	Yes	PS	No	5	3	2.06	0.49-8.62
rs45520832	No	PS	Yes	22	26	1.04	0.59-1.85
rs45520832	Yes	PS	Yes	1	0	NA	NA
rs220842	No	FVIII	No	1775	2416	1(REF)	
rs220842	Yes	FVIII	No	29	21	1.88	1.07-3.31
rs220842	No	FVIII	Yes	160	51	4.27	3.10-5.89
rs220842	Yes	FVIII	Yes	5	0	NA	NA
rs11215504	No	FVIII	No	1638	2222	1(REF)	
rs11215504	Yes	FVIII	No	166	215	1.05	0.85-1.30
rs11215504	No	FVIII	Yes	141	46	4.16	2.96-5.84
rs11215504	Yes	FVIII	Yes	24	5	6.51	2.48-17.1
rs45608938	No	FVL	No	2591	3379	1(REF)	
rs45608938	Yes	FVL	No	318	397	1.04	0.89-1.22
rs45608938	No	FVL	Yes	533	198	3.51	2.96-4.17
rs45608938	Yes	FVL	Yes	45	19	3.09	1.80-5.29
rs45545346	No	FVL	No	2721	3524	1(REF)	
rs45545346	Yes	FVL	No	192	255	0.98	0.80-1.18
rs45545346	No	FVL	Yes	559	210	3.45	2.92-4.07
rs45545346	Yes	FVL	Yes	20	9	2.88	1.31-6.33

OR odds ratio; CI confidence interval; PC protein C; PS protein S; FVL Factor V Leiden; REF reference.

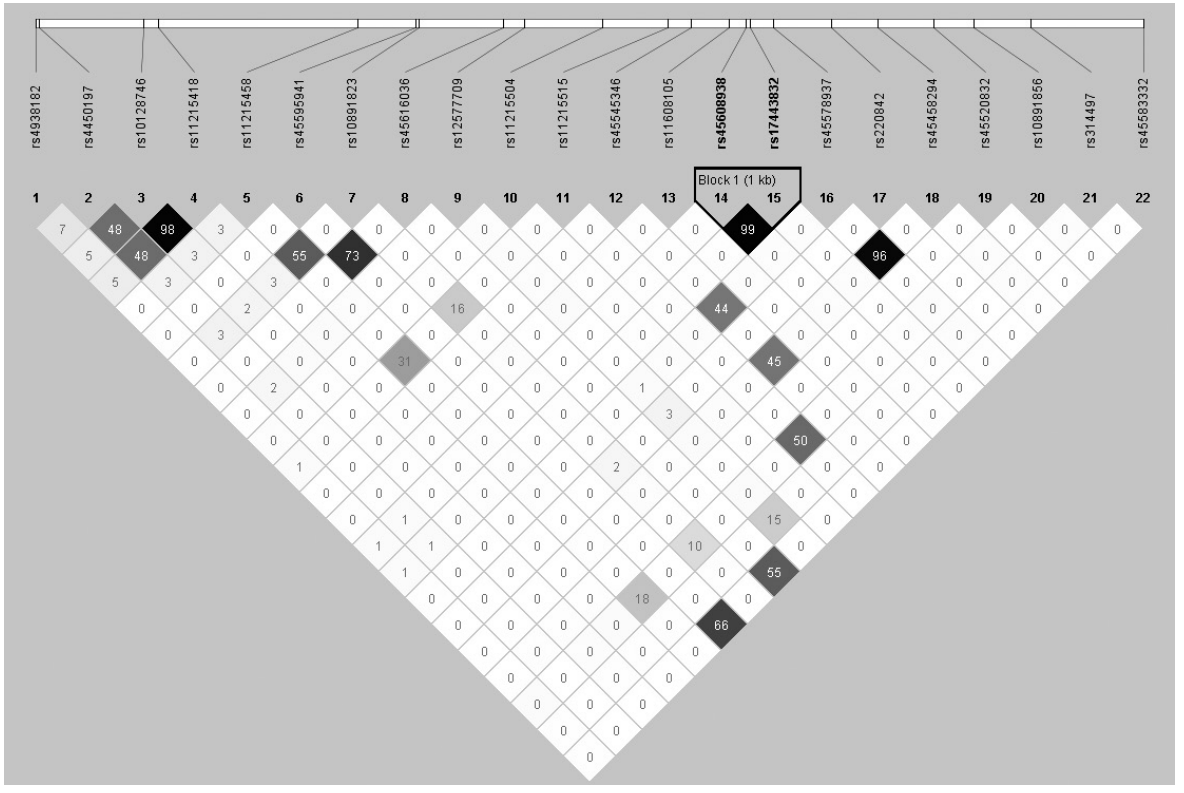

Figure 1. Linkage disequilibrium between *CADM1* variants which were associated with venous thrombosis in the univariable analysis of specific protein C pathway abnormality groups. Linkage equilibrium between the variants, as calculated in the controls, is expressed as r squared.

DISCUSSION

In this study we aimed to validate the *CADM1* gene, encoding cell adhesion molecule 1, as a gene involved in the etiology of venous thrombosis. We identified this gene as a candidate risk gene in the Vermont family.[1] The thrombosis association was most pronounced among individuals in this family with both variation in *CADM1* and protein C deficiency. To confirm the interaction of protein C deficiency and *CADM1* variants in increasing the risk of thrombosis, we studied 284 variants in *CADM1* in the population-based MEGA study. We performed analyses mainly by comparing thrombosis cases with protein C pathway abnormalities, i.e. low protein C or S levels, high factor VIII activity or factor V Leiden, with all controls.

For six variants in the *CADM1* gene, a consistent association with venous thrombosis was observed in one of the subgroups of protein C pathway abnormalities. Within individuals with low protein C or S levels, rs11608105 showed a 21-fold increased risk of venous thrombosis. Another variant (rs220842) was associated with venous thrombosis in the overall MEGA population and was only present in patients, and not in control subjects, with high factor VIII activity. Whether the variants are causal or are in linkage disequilibrium with unmeasured causal variants is not known. Our results suggest independent effects for the two variants. Both variants lie in intron 1, which comprises 240 kB of the 300 kB *CADM1* gene. There are a number of transcription factor binding sites and regulatory elements in intron 1. Examination of the 500 bp sequence flanking the variants revealed the occurrence of conserved elements (across 37 mammals) and open chromatin regions (DNase I hypersensitivity assay).[23] This suggests that epigenetic control may be the underlying functional mechanism by which these variants exert their effect on venous thrombosis.

One of the drawbacks of our study is the relatively low number of individuals per protein C pathway abnormality subgroup, which decreased our power to detect effects for *CADM1* variants. In addition, testing multiple SNVs for association with venous thrombosis increases the chance of false-positive associations. We therefore calculated FDR-adjusted q-values, after which we were no longer able to detect an association between the *CADM1* variants and venous thrombosis. We sought support for our hypothesis through addressing the association between venous thrombosis and *CADM1* variants in not only patients with low protein C levels, but also in other subgroups of patients with protein C pathway abnormalities. Although there was some overlap in

patients within the protein C pathway subgroups, we observed almost no overlap in the thrombosis-associated *CADM1* variants across the subgroups of protein C pathway abnormalities. Only one variant (rs11608105) was found to be associated with venous thrombosis in multiple subgroups, in this case in patients with low protein C and S levels. In some cases, the direction of the odds ratio for venous thrombosis risk of the positive *CADM1* variant differed across the protein C pathway abnormalities (Supplemental Table 2). Taken together, this may suggest that genetic variation in *CADM1* interacts only with single or specific factors within the protein C pathway.

Another drawback of our study is that the protein C and protein S deficiencies were not individually diagnosed, but we determined levels below clinical cut-offs using a single test. Therefore, the prevalence of the protein C pathway abnormalities may vary and some misclassification may have occurred. It is unlikely though to have affected the comparisons on a group level. In addition, as in all case-control studies, we cannot rule out that the thrombotic event itself influenced the coagulation factor levels, in particular the levels of the acute phase reactant factor VIII. However, the median time between blood draw and thrombotic event was 10 months and we did not observe any difference between the mean FVIII levels of blood samples drawn less than 6 months after the thrombotic event and blood samples drawn 6 or more months after the thrombotic event (mean levels of 134.9 and 132.7 IU/ml, respectively).

We identified several variants of which the risk allele was carried by patients or control subjects only. These might be involved as risk or protective alleles for venous thrombosis when co-occurring with a protein C pathway abnormality. However, since these variants were rare and the number of individuals was low, we are not able to draw conclusions about these variants.

The variant that was most strongly associated in the French Canadian family study, rs6589488, was not associated in the overall MEGA study (OR 1.07, 95% CI 0.98-1.17) nor in one of the subgroups of protein C pathway abnormalities (Supplemental Table 3). Linkage disequilibrium, as determined by r^2, with the variants consistently associated with venous thrombosis in our analysis (listed in Table 2) was low (<0.15). One explanation for the lack of a clear effect of rs6589488 in the current study is that the variants in the family study are rare mutations, private to this family or the French Canadian population. The results found in the current case-control study for

a joint thrombophilic effect of *CADM1* variants with protein C deficiency, protein S deficiency, or high factor VIII levels does suggest though that the *CADM1* pathway might play a role in the biology of hemostasis in the general population as well. The *CADM1* pathway links to the actin cytoskeleton and in the cancer literature its oncogenic effect is due to variants in *CADM1* as well as downstream proteins.[24-27] Analysis of genes of downstream members of the *CADM1* pathway might identify additional novel risk factors for venous thrombosis. Another possibility is that mutations in the gene for protein C (*PROC*) itself affect the interaction between CADM1 and protein C pathway. However, this would involve an indirect interaction between the downstream pathways associated respectively with the Endothelial Cell Protein C receptor and CADM1, as there is no evidence for a direct interaction of protein C with CADM1.

In conclusion, this study found some evidence of a joint effect of genetic variation in *CADM1* and protein C pathway abnormalities on the risk of venous thrombosis. This study aimed to validate a previous genetic study in a large thrombophilic family study, but could not replicate the specific associations observed in the family. Therefore, further study of the *CADM1* pathway is needed to determine whether abnormalities of the *CADM1* pathway link the risk for venous thrombosis to the vessel wall.

REFERENCES

1. Hasstedt SJ, Bezemer ID, Callas PW, Vossen CY, Trotman W, Hebbel RP, et al. Cell adhesion molecule 1: a novel risk factor for venous thrombosis. *Blood*. 2009;114:3084-91.
2. Giangreco A, Jensen KB, Takai Y, Miyoshi J, Watt FM. Necl2 regulates epidermal adhesion and wound repair. *Development*. 2009;136:3505-14.
3. Gomyo H, Arai Y, Tanigami A, Murakami Y, Hattori M, Hosoda F, et al. A 2-Mb sequence-ready contig map and a novel immunoglobulin superfamily gene IGSF4 in the LOH region of chromosome 11q23.2. *Genomics*. 1999;62:139-46.
4. Ito A, Okada M, Uchino K, Wakayama T, Koma Y, Iseki S, et al. Expression of the TSLC1 adhesion molecule in pulmonary epithelium and its down-regulation in pulmonary adenocarcinoma other than bronchioloalveolar carcinoma. *Lab Invest*. 2003;83:1175-83.
5. Biederer T, Sara Y, Mozhayeva M, Atasoy D, Liu X, Kavalali ET, et al. SynCAM, a synaptic adhesion molecule that drives synapse assembly. *Science*. 2002;297:1525-31.
6. Van der Weyden L, Arends MJ, Chausiaux OE, Ellis PJ, Lange UC, Surani MA, et al. Loss of TSLC1 causes male infertility due to a defect at the spermatid stage of spermatogenesis. *Mol Cell Biol*. 2006;26:3595-609.
7. Tatsumi K, Taatjes DJ, Wadsworth MP, Bouchard BA, Bovill EG. Cell adhesion molecule 1 (CADM1) is ubiquitously present in the endothelium and smooth muscle cells of the human macro- and micro-vasculature. *Histochem Cell Biol*. 2012;138:815-20.
8. Feistritzer C, Riewald M. Endothelial barrier protection by activated protein C through PAR1-dependent sphingosine 1-phosphate receptor-1 crossactivation. *Blood*. 2005;105:3178-84.
9. Niessen F, Furlan-Freguia C, Fernandez JA, Mosnier LO, Castellino FJ, Weiler H, et al. Endogenous EPCR/aPC-PAR1 signaling prevents inflammation-induced vascular leakage and lethality. *Blood*. 2009;113: 2859-66.
10. Finigan JH, Dudek SM, Singleton PA, Chiang ET, Jacobson JR, Camp SM, et al. Activated protein C mediates novel lung endothelial barrier enhancement: role of sphingosine 1-phosphate receptor transactivation. *J Biol Chem*. 2005;280:17286-93.
11. Bae JS, Rezaie AR. Thrombin inhibits nuclear factor kappaB and RhoA pathways in cytokine-stimulated vascular endothelial cells when EPCR is occupied by protein C. *Thromb Haemost*. 2009;101:513-20.
12. McVerry BJ, Garcia JG. In vitro and in vivo modulation of vascular barrier integrity by sphingosine 1-phosphate: mechanistic insights. *Cell Signal*. 2005;17: 131-9.
13. Wojciak-Stothard B, Ridley AJ. Shear stress-induced endothelial cell polarization is mediated by Rho and Rac but not Cdc42 or PI 3-kinases. *J Cell Biol*. 2003;161:429-39.
14. Murakami Y. Involvement of a cell adhesion molecule, TSLC1/IGSF4, in human oncogenesis. *Cancer Sci*. 2005;96:543-52.

15. Masuda M, Kikuchi S, Maruyama T, Sakurai-Yageta M, Williams YN, Ghosh HP, et al. Tumor suppressor in lung cancer (TSLC)1 suppresses epithelial cell scattering and tubulogenesis. *J Biol Chem*. 2005;280:42164-71.
16. Kawano S, Ikeda W, Kishimoto M, Ogita H, Takai Y. Silencing of ErbB3/ErbB2 signaling by immunoglobulin-like Necl-2. *J Biol Chem*. 2009;284:23793-805.
17. Martinelli I, De Stefano V, Mannucci PM. Inherited risk factors for venous thromboembolism. *Nat Rev Cardiol*. 2014;11:140-56.
18. Blom JW, Doggen CJ, Osanto S, Rosendaal FR. Malignancies, prothrombotic mutations, and the risk of venous thrombosis. *JAMA*. 2005;293:715-22.
19. Van Stralen KJ, Rosendaal FR, Doggen CJ. Minor injuries as a risk factor for venous thrombosis. *Arch Intern Med*. 2008;168:21-6.
20. O'Brien AE, Tate GM, Shiach C. Evaluation of protein C and protein S levels during oral anticoagulant therapy. *Clin Lab Haematol*. 1998;20:245-52.
21. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21:263-5.
22. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 1995;57:289-300.
23. University of California SC: Genome browser version NCBI 36/hg18
24. Sakurai-Yageta M, Masuda M, Tsuboi Y, Ito A, Murakami Y. Tumor suppressor CADM1 is involved in epithelial cell structure. *Biochem Biophys Res Commun.* 2009;390:977-82.
25. Yageta M, Kuramochi M, Masuda M, Fukami T, Fukuhara H, Maruyama T, et al. Direct association of TSLC1 and DAL-1, two distinct tumor suppressor proteins in lung cancer. *Cancer Res*. 2002;62:5129-33.
26. Fukuhara H, Masuda M, Yageta M, Fukami T, Kuramochi M, Maruyama T, et al. Association of a lung tumor suppressor TSLC1 with MPP3, a human homologue of Drosophila tumor suppressor Dlg. *Oncogene*. 2003;22:6160-5.
27. Shingai T, Ikeda W, Kakunaga S, Morimoto K, Takekuni K, Itoh S, et al. Implications of nectin-like molecule 2/IGSF4/RA175/SgIGSF/TSLC1/SynCAM1 in cell-cell adhesion and transmembrane protein localization in epithelial cells. *J Biol Chem*. 2003;278:35421-7.

SUPPLEMENTAL TABLES

Table S1. Minor allele frequencies of *CADM1* variants in overall MEGA study population

		Minor allele frequency, %	
***CADM1* variant**	**Position**	**Patients**	**Controls**
rs11215392	114543618	2.65	2.75
rs34157656	114544511	44.3	45.0
rs10444329	114544893	18.1	17.0
rs17118020	114545350	1.50	1.33
rs17118023	114546173	18.1	16.9
rs17649730	114546639	15.0	14.1
rs4936321	114546799	47.6	46.2
rs11606837	114548047	49.4	48.4
rs4938182	114548246	21.0	19.8
rs45460594	114548330	3.09	2.75
rs45486791	114548565	0.53	0.44
rs45539744	114548882	0.01	0.02
rs4450197	114549421	2.41	2.04
rs1048932	114550060	43.9	43.0
rs45483591	114551963	0.04	0.03
rs45445298	114554121	0.36	0.45
rs17304149	114554390	48.3	49.1
rs17118046	114554937	3.84	3.80
rs45508098	114555249	16.6	15.8
rs7928746	114556120	2.03	2.36
rs4938183	114556779	4.10	4.09
rs45479795	114557630	4.41	4.27
rs11215400	114557845	27.3	27.4
rs45483594	114558449	16.6	15.9
rs12807135	114558718	49.6	50.4
rs45594631	114559767	0.00	0.02
rs11215403	114563795	25.9	25.8
rs45604639	114565259	0.33	0.28
rs7937380	114565377	26.0	26.0
rs45614835	114565529	16.3	15.6
rs4936322	114566743	45.1	43.8
rs45605138	114567521	1.70	1.78
rs4245160	114567760	0.01	0.00
rs45625839	114568381	0.01	0.01
rs7101437	114568851	49.9	50.4
rs45628237	114569486	0.33	0.26
rs11215406	114570292	27.3	27.3
rs11215407	114570503	6.13	6.38

Table S1. Continued

		Minor allele frequency, %	
***CADM1* variant**	**Position**	**Patients**	**Controls**
rs10891805	114571691	3.61	3.72
rs45456599	114571885	16.3	15.5
rs45617644	114571999	0.01	0.02
rs45574838	114572238	0.04	0.00
rs6589484	114576024	3.69	3.76
rs45529533	114576096	11.0	11.1
rs45479100	114577512	16.6	15.9
rs12226198	114579444	5.74	5.86
rs10128746	114580646	3.71	3.63
rs11215415	114580742	2.18	2.44
rs45505693	114583362	0.64	0.89
rs3802858	114583702	45.0	44.2
rs3802857	114583828	35.1	35.4
rs11215418	114585104	3.70	3.62
rs7125361	114585252	44.8	43.7
rs9645660	114586773	49.2	48.1
rs11215419	114587020	49.4	50.6
rs45516099	114587093	16.8	15.8
rs7482812	114588382	3.00	2.81
rs6589486	114589507	45.8	46.8
rs12281523	114589876	5.39	5.26
rs45525440	114590677	5.37	5.18
rs45489793	114592265	18.2	17.2
rs11215424	114592631	28.5	28.9
rs4938190	114592960	47.9	47.0
rs7106961	114593510	1.57	1.66
rs7947402	114593630	49.3	48.0
rs45593334	114594650	28.5	28.7
rs45583736	114595117	0.03	0.04
rs4245161	114595636	0.03	0.00
rs7479259	114595925	45.1	45.9
rs45614535	114596076	44.9	45.7
rs11825649	114597503	1.63	1.51
rs45460202	114597825	1.63	1.76
rs1938736	114598207	18.2	17.2
rs11215427	114598648	28.5	29.0
rs12575340	114600534	17.2	16.4
rs11215430	114601206	5.30	5.14
rs10891812	114601641	46.8	46.1
rs6589488	114602166	15.2	14.3

Table S1. Continued

		Minor allele frequency, %	
***CADM1* variant**	**Position**	**Patients**	**Controls**
rs12284489	114602367	5.32	5.20
rs12280033	114603084	7.04	6.83
rs12417740	114603646	45.2	46.0
rs11215431	114604893	0.03	0.00
rs11602686	114605848	45.3	46.2
rs11215433	114606504	7.01	6.89
rs10458967	114608081	5.32	5.18
rs10458969	114608403	16.6	15.5
rs11215437	114609382	24.8	25.0
rs10891814	114609820	38.6	37.8
rs10502200	114610942	3.34	3.52
rs45593037	114612214	4.64	4.36
rs947802	114613194	38.9	38.2
rs12283904	114614312	0.00	0.03
rs2269737	114616515	19.2	19.1
rs11215439	114617425	19.1	18.2
rs12421121	114617518	19.1	19.7
rs17118125	114619942	19.0	19.0
rs11215445	114620383	22.6	22.5
rs9633941	114621837	19.5	18.8
rs12225639	114622453	16.1	15.2
rs45624531	114622551	19.0	19.1
rs10502199	114625825	15.7	15.2
rs1892773	114627836	20.5	20.7
rs7127390	114627937	20.4	20.4
rs4936325	114630329	15.4	15.1
rs17118149	114630440	0.09	0.09
rs45604331	114632418	0.00	0.02
rs45538440	114634182	0.31	0.35
rs45577334	114634631	1.40	1.34
rs6589490	114637110	37.5	37.2
rs11215455	114639795	20.3	20.8
rs2154690	114640754	38.0	37.7
rs11215456	114640983	17.6	16.9
rs4938193	114641217	20.0	20.7
rs4597099	114641818	37.4	37.2
rs10891818	114642013	35.7	35.5
rs10891819	114642457	18.9	18.0
rs11215458	114645061	3.71	3.65
rs7950069	114645763	15.6	14.8

Table S1. Continued

		Minor allele frequency, %	
***CADM1* variant**	**Position**	**Patients**	**Controls**
rs11215459	114646718	1.62	1.26
rs45539832	114648118	5.51	5.41
rs4938194	114648551	38.1	37.9
rs10891820	114649664	12.6	13.0
rs12577839	114649744	0.01	0.00
rs45451094	114649852	0.01	0.03
rs17118172	114650309	5.27	5.02
rs12788053	114652701	20.3	21.0
rs10502203	114655447	1.40	1.16
rs17519855	114656695	0.23	0.24
rs7944529	114657017	11.6	11.2
rs7944955	114657247	31.8	31.8
rs7931895	114657509	31.7	31.8
rs11215462	114658528	0.03	0.00
rs17118198	114660163	0.09	0.09
rs45595941	114662359	0.79	0.70
rs11215466	114663198	18.2	17.5
rs10891823	114663444	6.74	6.48
rs2014270	114664443	12.6	13.0
rs17441594	114664964	11.6	11.0
rs7936399	114665469	38.6	38.3
rs17441610	114667144	11.6	11.0
rs4938195	114668875	12.7	13.0
rs7104872	114670321	19.5	19.1
rs7928044	114670523	6.53	6.41
rs11215470	114671854	0.03	0.00
rs45581535	114674341	3.69	3.70
rs45488901	114674457	11.6	11.0
rs11215474	114674839	19.0	18.3
rs7104113	114675467	38.8	38.6
rs45505692	114676989	0.01	0.02
rs10891825	114678381	34.8	33.8
rs2040456	114683727	0.03	0.00
rs2157612	114684281	10.9	10.4
rs7949084	114685949	46.5	47.0
rs12290790	114688338	10.8	10.3
rs45616036	114688640	0.82	0.84
rs17442145	114688855	1.26	1.15
rs17442179	114689108	3.28	3.43
rs45626034	114693674	0.01	0.03

Table S1. Continued

		Minor allele frequency, %	
***CADM1* variant**	**Position**	**Patients**	**Controls**
rs988873	114694438	21.0	21.3
rs11607436	114694623	2.18	2.42
rs2366904	114695046	46.7	46.1
rs12577709	114695169	14.0	13.6
rs17118264	114695758	13.6	13.3
rs4396320	114696475	46.8	46.3
rs12284145	114696918	0.03	0.00
rs45467696	114699386	0.07	0.03
rs45474291	114701226	11.0	10.6
rs45508698	114701763	0.69	0.96
rs10891829	114703906	46.6	46.0
rs45543336	114707093	10.9	10.4
rs45469396	114707350	0.53	0.55
rs17118279	114707907	0.10	0.12
rs10891832	114710833	43.7	42.6
rs10488710	114712386	33.1	33.3
rs10891833	114712918	38.5	38.5
rs7952231	114713208	38.5	38.6
rs9888216	114714603	44.3	43.4
rs2105976	114715710	44.2	43.2
rs7105871	114717935	20.7	21.3
rs45465296	114718461	12.2	11.8
rs11215504	114718584	4.94	4.35
rs4938201	114723923	40.4	40.7
rs12575143	114726812	2.51	2.39
rs45599536	114727833	0.39	0.49
rs10891836	114728167	44.3	43.1
rs2105982	114729014	44.3	43.1
rs7120311	114729924	22.7	23.4
rs11215512	114732381	44.2	43.1
rs10891839	114733207	33.0	33.3
rs10891840	114734721	44.3	43.1
rs17521934	114735633	12.0	11.7
rs11215515	114738087	4.41	4.26
rs45559239	114738583	0.04	0.03
rs45455497	114740709	0.40	0.49
rs11215517	114742555	10.4	10.1
rs10891842	114744233	39.1	39.1
rs10160742	114744607	7.06	7.06
rs45545346	114745259	3.13	3.42

Table S1. Continued

		Minor allele frequency, %	
***CADM1* variant**	**Position**	**Patients**	**Controls**
rs45580634	114746210	1.66	1.73
rs17118309	114746787	0.97	1.10
rs220850	114753565	49.4	48.8
rs4938202	114754917	39.2	39.1
rs11608105	114756400	4.57	4.72
rs45585234	114761471	0.00	0.03
rs45608938	114761668	5.33	5.42
rs17443832	114762977	5.31	5.38
rs220869	114767246	0.04	0.09
rs45578937	114769761	6.58	7.07
rs45514899	114771373	0.00	0.03
rs220872	114771575	50.5	49.9
rs7114341	114774371	44.5	43.5
rs45555732	114775788	1.02	1.14
rs11215532	114776409	44.7	43.8
rs4938203	114780571	44.6	43.7
rs220828	114782015	42.9	41.6
rs2366914	114784746	36.6	36.8
rs45559131	114786337	0.00	0.03
rs220842	114787382	0.76	0.51
rs17118328	114787872	1.95	1.82
rs220843	114788745	16.0	16.8
rs220847	114791327	49.3	48.7
rs11215545	114791960	42.9	41.9
rs12273801	114795200	0.01	0.01
rs7106275	114797011	0.26	0.21
rs220860	114799274	16.1	16.7
rs220861	114799402	6.26	6.55
rs45455306	114799791	1.37	1.54
rs220862	114801129	14.1	15.0
rs45458294	114801307	6.51	6.92
rs220864	114801841	14.1	14.6
rs220865	114802160	22.2	23.3
rs10891854	114804638	38.8	39.0
rs220836	114807081	20.5	21.3
rs45522132	114807342	1.37	1.61
rs7122693	114809573	43.4	42.7
rs45587938	114810013	10.8	10.1
rs17444623	114812143	18.8	19.8
rs17451032	114813684	1.02	1.16

Table S1. Continued

		Minor allele frequency, %	
***CADM1* variant**	**Position**	**Patients**	**Controls**
rs45509898	114813930	2.09	2.26
rs45473492	114816541	1.29	1.44
rs45520832	114818047	0.14	0.11
rs45625135	114818415	1.55	1.71
rs220838	114819312	18.9	19.7
rs12801130	114820321	36.7	35.8
rs17118342	114820680	0.07	0.25
rs160604	114823801	0.01	0.00
rs544083	114825691	17.3	18.0
rs220840	114826173	17.3	18.1
rs314474	114826343	17.2	17.9
rs314476	114827516	18.8	19.8
rs10502202	114829700	21.5	22.1
rs10891856	114830116	6.17	5.75
rs1155756	114830467	37.4	36.6
rs7927390	114831701	18.8	19.8
rs10047420	114834362	38.0	37.5
rs45490692	114835734	0.62	0.54
rs314491	114840421	20.2	20.8
rs10891859	114840831	35.8	35.4
rs314494	114841812	20.3	20.9
rs314495	114842583	20.3	20.8
rs314496	114842787	20.2	20.9
rs45474398	114844445	3.52	3.88
rs17451771	114845558	6.55	7.03
rs314497	114847142	6.27	5.86
rs11827474	114848809	0.01	0.00
rs17118360	114849006	0.07	0.16
rs1460909	114851977	0.43	0.36
rs314503	114852071	6.52	7.01
rs314507	114854460	0.01	0.00
rs314512	114858104	6.46	7.03
rs314513	114858508	6.49	7.02
rs314514	114861898	1.74	1.62
rs7924765	114862746	0.00	0.01
rs12281277	114866132	0.00	0.01
rs11215574	114868653	25.9	25.8
rs17524208	114871498	3.36	3.89
rs973550	114872351	0.27	0.28
rs17524278	114875616	6.49	7.13

Table S1. Continued

		Minor allele frequency, %	
***CADM1* variant**	**Position**	**Patients**	**Controls**
rs314464	114878567	0.01	0.01
rs45583332	114880825	0.80	0.70
rs11215581	114884622	0.49	0.36
rs314469	114885900	7.07	7.54
rs314468	114887234	6.58	7.20
rs7101558	114892659	6.94	7.44

Table S2. Associations of positive variants identified in different protein C pathway subgroups with venous thrombosis.

CADM1	Overall MEGA		Low protein C activity		Low protein S levels		High factor VIII activity		FV Leiden carriers	
variants	cases	controls	cases	OR (95% CI)	cases	OR (95% CI)	cases	OR (95% CI)	cases	OR (95% CI)
rs11608105	4.57	4.72	7.22	1.57 (1.05-2.34)	13.0	2.98 (1.27-7.02)	2.73	0.56 (0.29-1.11)	3.36	0.93 (0.69-1.26)
rs45520832	0.14	0.11	0	NA	2.17	20.1 (2.45-166)	0	NA	0.09	0.77 (0.10-6.05)
rs11215504	4.94	4.35	5.67	1.32 (0.85-2.06)	2.17	0.49 (0.07-3.56)	7.58	1.79 (1.18-2.73)	5.44	1.26 (0.96-1.67)
rs220842	0.76	0.51	1.03	2.03 (0.72-5.73)	0	NA	1.52	3.02 (1.18-7.74)	0.69	1.35 (0.63-2.90)
rs45608938	5.33	5.42	5.93	1.10 (0.72-1.69)	6.52	1.21 (0.38-3.87)	4.55	0.84 (0.50-1.41)	3.89	0.71 (0.52-0.97)
rs45545346	3.13	3.42	2.84	0.83 (0.45-1.51)	4.35	1.27 (0.32-5.13)	2.44	0.71 (0.35-1.44)	1.73	0.50 (0.32-0.79)

The risk allele frequencies were calculated (shown in percentages) in the overall MEGA study population and in subgroups of patients with protein C pathway abnormalities.

Table S3. The *CADM1* variant found in the Vermont family assessed for associations with venous thrombosis in overall MEGA population and subgroups.

rs6589488	Risk allele frequency, %	OR (95% CI)
Overall controls	14.3	REF
Overall patients	15.2	1.07 (0.98-1.17)
Low protein C	14.2	0.99 (0.74-1.32)
Low protein S	17.4	1.26 (0.59-2.69)
High factor VIII	15.2	1.07 (0.79-1.45)
FVL carriers	14.5	1.02 (0.85-1.21)