Published in final edited form as:

J Thromb Haemost. 2013 April ; 11(4): 715–723. doi:10.1111/jth.12157.

Genetic variants associated with protein C levels

C.Y. Vossen^{*,†}, B.P. Koeleman^{*}, S.J. Hasstedt[‡], I.J. Nijman^{*}, I.J. Renkens^{*}, P.W. Callas[§],
F.R. Rosendaal^{†,¶,**}, and E.G. Bovill^{††}

^{*}Medical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands [†]Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands [‡]Human Genetics, University of Utah, Salt Lake City, UT, USA [§]Biostatistics, University of Vermont, Burlington, VT, USA [¶]Thrombosis and Hemostasis, Leiden University Medical Center, Leiden, The Netherlands ^{**}Eindhoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden, The Netherlands ^{††}Pathology, University of Vermont, Burlington, VT, USA

Summary

Background—Normal protein C (PC) plasma levels range widely in the general population. Factors influencing normal PC levels are thought to influence the risk of venous thrombosis. Little is known about the underlying genetic variants.

Objectives—We performed a genome scan of normal PC levels to identify genes that regulate normal PC levels.

Patients/Methods—We performed a variance components linkage analysis for normal PC levels in 275 individuals from a single, large family. We then sequenced candidate genes under the identified linkage peak in eight family members: four with high and four with low, but normal PC levels. For variants showing a difference in carriers between those with high and low PC levels, we re-evaluated linkage in the 275 family members conditional on the measured genotype effect. Genotype-specific mean PC levels were determined using likelihood analysis. Findings were replicated in the Leiden Thrombophilia Study (LETS).

Results—We identified a quantitative trait locus at chromosome 5q14.1 affecting normal PC plasma level variability. Next-generation sequencing of 113 candidate genes under the linkage peak revealed four SNPs in *BHMT2*, *ACOT12*, *SSBP2* and *XRCC4*, which significantly increased PC levels in our thrombophilic family, but not in LETS.

Conclusions—We identified four genes at chromosome 5q14.1 that might influence normal PC levels. *BHMT2* seems the most likely candidate to regulate PC levels via homocysteine, a competitive inhibitor to thrombin. Failure to replicate our findings in LETS might be due to differences between the studies in genetic background and linkage disequilibrium patterns.

Keywords

Genetic factors; linkage; next-generation sequencing; normal variation; protein C

Introduction

Protein C (PC) is a vitamin-K dependent glycoprotein which after activation by the thrombin-thrombomodulin complex on the surface of endothelial cells inhibits thrombin

generation by degrading activated factor V and factor VIII[1;2]. Individuals with inherited PC deficiency show an increased risk of developing venous thrombosis[3;4]. Normal PC levels range widely in the general population resulting in a considerable overlap in levels between individuals with confirmed hereditary PC deficiency and their PC normal relatives[5]. Therefore, variants influencing normal PC levels could play a role in increasing venous thrombosis risk. This has been confirmed by Spek *et al.* who found a 1.6-fold increased risk for venous thrombosis for a genotype consisting of three polymorphisms in the PC coding gene (*PROC*), which was associated with lower PC levels in non-deficient individuals[6].

The heritability of normal PC levels has been estimated to exceed 40%[7,8]. The GAIT study found evidence for a quantitative trait locus affecting normal PC levels on chromosome 16q[9]. The GAIT study sample consists of 398 subjects from 21 extended Spanish pedigrees, of which 12 families were selected through an affected proband and 9 families were randomly selected without regard to phenotype[9]. The 12 families that were selected in this study through an affected proband were responsible for 100% of the linkage signal[9], which suggests that the linkage signal at chromosome 16q might be family-specific. Other genetic factors might play a role in regulating PC levels in other families or in the general population. A recent genome-wide association scan for PC antigen levels with about 2.5 million single-nucleotide polymorphisms (SNPs) in 8048 individuals of European ancestry revealed several loci that might influence PC levels at 2p23 (*GCKR*), 2q13-14 (*PROC*), 20q11 (*PROCR* and *EDEM2*) and 7q11.23 (*BAZ1B*)[10]. A functional variant in *PROCR* (rs867186) explained 10% of the variation in protein C levels and supported the findings from another study, which found an association with protein C levels in 336 European-Americans, explaining 13% of the phenotypic variation[11].

To further identify genes influencing PC levels, we performed a genome scan of normal PC levels in a large, genetically homogeneous French-Canadian thrombophilic family and sought validation for our findings in a population-based case-control study.

Material and methods

Study populations

The Vermont family is a large kindred with type I PC deficiency of which the ascertainment and evaluation of the family members has been previously described[3]. All participating subjects or their legal guardians gave informed consent. The Human Experimentation Committees of the University of Vermont, College of Medicine, Burlington (VT) and the Beth Israel Hospital, Boston (MA) have approved the study. The pedigree of this family spans 7 generations with data available on the most recent 5. For the genome scan of normal PC levels, we included only family members without the 3363C insertion in the protein C gene, i.e. non-deficient family members, with levels of PC above 67% of normal (n=294). We excluded 19 family members who were on oral anticoagulant treatment at the time of the blood draw (n=2), those with a confirmed history of venous thrombosis (n=6), and those without age (n=3) or genome scan information (n=8). Subsequently, we sequenced candidate genes under the identified linkage peak in samples from eight family members who were informative for the linkage peak on 5q14.1: four members with levels at the low extreme end of the distribution of the 275 individuals included in the linkage analysis (range 68–75% of normal) and four members with levels at the high end of the PC level distribution (range 187–207% of normal; see Figure 1 for a work flow diagram). An individual was considered informative for the peak when excluding this individual from the linkage analysis lead to a reduction in the LOD score. We started selecting individuals from the end of the distribution, i.e. either 68% or 207% of normal, and preferred individuals being informative for the peak over individuals having the most extreme levels. So, an

individual with a level of 187% of normal and a larger reduction in LOD score when removing this person from the linkage analysis was preferred over an individual with a level of 204% and a lower reduction in LOD score after removing this person from the linkage analysis. The selected individuals within the high or low level groups were not very closely related (two were 2nd degree relatives, the rest was not closer related than 3rd degree). Interesting variants were genotyped in the 275 family members who were originally included in the linkage analysis to determine their effect on PC levels in the thrombophilic family (see Figure 1). Lastly, variants that were associated with normal PC levels in the 275 Vermont family members were genotyped in 461 control individuals with normal PC levels (>68% of normal) from a case-control study, The Leiden Thrombophilia Study (LETS; see Figure 1). In LETS, 474 patients (age 18–70 years) have been included with a first objectively confirmed deep vein thrombosis from three anticoagulation clinics in the Netherlands (Leiden, Amsterdam and Rotterdam) between January 1988 and December 1992[12]. As control subjects, 474 friends or partners of patients were included in the study[12]. Each participant filled out a questionnaire, provided a blood sample and provided written informed consent. The study was approved by the Medical Ethics Committee of the Leiden University Medical Center, Leiden, The Netherlands.

PC activity assays

Blood from members of the Vermont family was collected into siliconized glass Vacutainer tubes containing 3.8% buffered citrate solution (Becton Dickinson, Franklin, NJ). Platelet poor plasma was produced within 1 hour by centrifugation at 3,000g for 10 minutes at room temperature. Samples were stored at –70°C and thawed at 37 °C just before assay performance. PC activity levels in the Vermont family were measured by a clot-based functional assay[13] following manufacturer's instructions. In LETS, blood was collected in tubes containing 0.106 M trisodium citrate. Plasma was prepared by centrifugation at 2,000g for 10 minutes at room temperature and stored at –70°C. PC activity levels were determined using a chromogenic assay (Chromogenix, Mölndal, Sweden).

Genome scan and candidate-gene sequencing—Genotyping with 375 autosomal markers was performed as described by Hasstedt *et al.*[14]. After identifying the linkage peak at chromosome 5q14.1 (at 86 cM of the Marshfield map, 95% CI 81–94 cM), we sequenced the exonic sequences (including 5 bp of the intron boundaries, 1 kb of the promoter region and the 3'UTR) of 113 unique genes (see Supplementary Table 1) between markers D5S1351 (78.3 cM Marshfield map) and D5S1725 (97.8 cM Marshfield map) after genomic enrichment of the target sequences using capturing probes. We designed sixty nucleotide long capturing probes with an average tiling of 10 nucleotides according to the Ensembl database (www.ensembl.org) using the GRCh37/hg19 build of the human reference genome[15]. The design was spotted twice on a 1M custom Agilent array with randomized positions. In total, 211344 forward probes and 218180 reverse probes were included, i.e. 97% of all possible probes covering 1603231 bp (97%) of the design area excluding 333270 repeat-masked positions. Three percent of the probes have strong GC bias or simple sequences or did not uniquely map to the genome and were therefore not included in the capture sequences. As described earlier[15], 1 µg of genomic DNA from each patient was fragmented, adapter-ligated, barcoded, hybridized to the designed capturing probes and printed on microarray slides. The eight selected samples from the Vermont family study were deposited on one sequencing slide and sequenced using a SOLiD (Sequencing by Oligonucleotide Ligation and Detection) analyzer version 4.

Mapping of sequencing data, variant calling and variant validation

Sequencing reads were mapped against the Ensembl reference genome (GRCh37/hg19) using Burrows-Wheeler Aligner[16]. Variants were filtered out when they were not

supported by 3 unique reads, when they did not appear at least once in the first 25 bases of a read (i.e. the region of a read that binds most accurately the target of interest) and when they did not have a coverage 10x, base quality 25 and variant percentage 15%. All eight family members with normal PC levels at the extreme low or high end of the distribution of the 275 included in the original linkage analysis who were selected for sequencing were informative for the linkage peak at chromosome 5q14.1. We therefore only considered a variant of interest when it was called 3 times in the four individuals in one group and called

1 in the four individuals in the other group with a difference of at least 3 carriers (i.e. four carriers versus 1 carrier, or three carriers versus no carriers in the two groups). For variants in the 3' or 5' untranslated region, we were more stringent; a variant was considered of interest only when the variant did not show up in any read in non-carriers. Variants that were regarded interesting based on the above criteria were Sanger sequenced in the same eight samples to confirm that they were true positive findings. To this end, we amplified the locations of interest with specific primers designed using Primer3 (v. 0.4.0; <http://fokker.wi.mit.edu/primer3/put.htm>) and sequenced the PCR products using Big Dye terminator and AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA) and an Applied Biosystems 3730 DNA analyzer. Supplement 1 includes the Sanger sequencing protocol used. We used the PolyPhred program to identify variants[17]. Confirmed interesting variants were genotyped in the 275 family members that were included in the original linkage analysis by performing KBiosciences competitive allele-specific PCR SNP genotyping assays according to the manufacturer's instructions (KASPar, KBioscience Ltd., Hoddesdon, UK). In addition, variants that showed an effect on PC levels in the Vermont family were genotyped in 461 LETS control individuals with PC levels above 68% of normal. Supplement 2 includes a Table with the sequences that formed the basis of the KASPar assays designed by KBioscience. Fluorescence endpoint reading was done on an Applied Biosystems 7900 HT analyzer.

Statistical analysis

Using the variance component method in SOLAR (Sequential Oligogenic Linkage Analysis Routines)[18], we estimated the proportion of the phenotypic variance attributed to polygenes (heritability) and the proportion of the variance attributed to environmental factors shared within a household (common household effect) for normal PC levels. The parameters for heritability (h^2), household effect (c^2), as well as the effects of covariates (age, sex, age*sex and age²) were estimated simultaneously using maximum likelihood analysis. Plasma concentrations of PC activity were assumed to distribute as a multivariate normal density (kurtosis below 0.8). We subsequently performed variance components linkage analysis for levels of PC in SOLAR[18], which uses maximum likelihood analysis to estimate the heritability attributed to a specific genomic location (q^2) and the residual heritability. Age- and sex-adjustments were made simultaneously at each genomic location. LOD scores were computed as the $\log_{10}$ likelihood for q^2 estimated to $q^2=0$. The 95% confidence interval was estimated by the points on the curve defined by dropping the LOD score by one unit. To assess the impact of interesting variants identified by next-generation sequencing on the linkage signal found on 5q14 in 275 family members, we re-evaluated linkage conditional on the measured genotype effect [19]. Variants were coded as an additive model (1=major allele homozygotes, 2=heterozygotes, 3=minor allele homozygotes). Likelihood analysis in jPAP[20] was used to estimate mean PC levels per genotype for variants that affected the linkage signal on 5q14, including polygenic inheritance to account for family structure and with adjustment for age and sex. Genotype-specific means were estimated one variant at a time; a common standard deviation was assumed. In the LETS control individuals, we determined and compared genotype-specific means adjusted for age and sex by running a general linear model procedure with PC levels

as dependent variable, the variant of interest as fixed factor, and age and sex as covariates using SPSS statistical software (version 20.0; SPSS, Chicago, IL, USA).

Results

A total of 275 family members (112 men, 163 women) from 175 households (range individuals per household 1–5) without PC deficiency, a definite history of venous thrombosis or anticoagulant treatment at the time of the blood draw were included in the analysis. Their mean age at the blood draw was 31 years (range 0–81 years). PC levels ranged from 68% to 207% of normal.

Variance components analysis revealed a high heritability for PC levels of 80.5% (standard error (SE) 11.2; $p < 0.000$) and revealed no household effect (0.0%). Subsequent linkage analysis showed significant linkage evidence at 5q14.1 with a LOD score of 3.5 (Figure 2 and 3; 86 cM Marshfield map, 95% CI 81–94 cM, nearest marker D5S1501).

Next-generation sequencing identified 871 variants in the candidate genes under the linkage peak at 5q14.1 in eight family members who were informative for the linkage peak: four with high (187–207% of normal) and four with low, but normal PC levels (68–75% of normal). Only nine variants showed a sufficient difference in the number of carriers (3) between the four individuals with high and the four individuals with low PC levels (Table 1). Of the two *SSBP2* 3' UTR variants, which were identified in the same three individuals, rs16878630 was conserved across mammals in Ensembl (www.Ensembl.org), whereas rs78668688 was not. Therefore, we only genotyped SNP rs16878630. The remaining eight variants were genotyped in the 275 family members that were originally included in the linkage analysis. For two individuals, all genotyping assays failed. Linkage analyses performed conditional on each of the variants separately in the remaining 273 family members revealed that four variants in *BHMT2*, *ACOT12*, *SSBP2* and *XRCC4* significantly reduced the linkage signal (Table 2). Performing a linkage analysis conditional on all four SNPs reduced the LOD score from 2.73 to 0.87 ($N=263$ with genotyping information for the four SNPs, $p=0.0033$). Table 3 shows for each of the four variants an increase in PC levels with each addition of a minor allele. These four variants were genotyped in 461 population-based LETS control individuals with normal PC levels. Genotyping failed for all variants for four samples. We found no association with PC levels for the genotyped variants in the remaining 457 LETS control individuals (Table 4).

Discussion

We performed a genome scan of normal PC activity levels in members of a large thrombophilic family and identified a linkage peak at chromosome 5q14.1. Further investigation of the candidate genes under the linkage peak by next-generation sequencing identified four SNPs in *BHMT2*, *ACOT12*, *SSBP2* and *XRCC4* that significantly increased PC levels in our thrombophilic family. Given the reduction of the LOD score at the linkage peak to a LOD score below 1 when re-evaluating linkage conditional on all four SNPs indicates that these SNPs are functional themselves or in strong linkage disequilibrium (LD) with functional variants explaining the linkage peak. The association with an increase in PC levels could not be replicated in LETS, a population-based case-control study. In contrast, we found in LETS a trend towards a decrease in PC levels. This “flip-flop” association of the genotyped SNPs in LETS compared to the Vermont family might be explained by a difference between the studies in the correlation of the genotyped SNPs with other loci influencing PC levels[21]. These other loci might harbor the “true” functional variants that are in strong LD with the genotyped non-causal SNPs, or might harbor other causal variants that interact with the genotyped SNPs in influencing PC levels. The Vermont family is

genetically homogeneous[22] compared to the LETS control group, which could lead to differences in LD patterns and haplotype frequencies between the studies and thereby in the correlation between the genotyped SNPs and other loci. Although our findings could be family-specific, it could also be that the genetic homogeneity of the Vermont family might have enabled us to observe the association between the four variants and protein C levels. The heterogeneity of the population-based LETS study might have precluded subsequent replication of our findings. Therefore, instead of genotyping these variants in other population-based studies, we believe that further *in vitro* studies are necessary to determine the role of the four identified candidate genes in influencing protein C levels and the possibility of extrapolation to the general population.

It is unclear how *ACOT12*, involved in pyruvate metabolism, and the DNA repair protein *XRCC4* and single-stranded DNA-binding protein 2 (*SSBP2*) might be involved in regulating PC levels. For the betaine-homocysteine methyltransferase 2 protein (*BHMT2*), we did find evidence in the literature for a possible effect on regulating PC levels[23]. *BHMT2* is involved in the regulation of homocysteine metabolism. Homocysteine can act as a competitive inhibitor to thrombin and thereby reduce PC activation[23]. Further basic research is needed to understand how variants in *BHMT2* can affect the regulation of PC levels via homocysteine. *BHMT2* rs682985, which increased PC levels in our study, has previously been found to be more frequently present among ischemic stroke patients compared to control individuals in a study by Giusti *et al.* [24]. Haplotype analysis, however, did not confirm an independent association between rs682985 and ischaemic stroke[24].

No linkage evidence was found at other locations, including the location at chromosome 2 of the PC structural gene and the region on chromosome 16 identified as a candidate quantitative trait locus by the GAIT study[9] and the regions on chromosomes 2, 7 and 20 identified by a recent genome-wide association scan for PC antigen levels [10]. The discrepancy between findings of the GAIT study and our study can be explained by differences in the genetic make-up of the study subjects; the GAIT study included 398 individuals of 21 extended Spanish pedigrees[9], whereas our study included a very homogeneous large family of which all heterozygotes for the 3363 PC mutation living in North America are related to a couple of French settlers within 10 generations[22].

Next-generation sequencing platforms have reduced the cost and time for sequencing candidate genes, but also are more likely to produce errors than for example Sanger sequencing[25]. Although the variants presented in our paper were confirmed to be positive by Sanger sequencing, we cannot exclude the possibility that variants were missed when sequencing genes under the linkage peak. For example *BHMT2* exons 1 and 7 (of a total of 8 exons) were mostly not covered well, which affected the overall mean coverage for *BHMT2* (Table 5). For *ACOT12*, exons 1,3,9 and 10 (of a total of 15 exons) and for *SSBP2* exons 3, 6, 10, 12 and 13 (of a total of 17 exons) were less well covered affecting overall coverage (Table 5). For *XRCC4*, all 7 exons were well covered (Table 5). In addition, protein C levels were measured only once in the 8 family members of which samples were selected for sequencing the candidate genes under the linkage peak at 5q14.1. Although this will induce uncertainty about the actual levels of the selected individuals, the genetic variants identified in these individuals seemed to explain the linkage peak and were clearly associated with protein C levels in the 275 family members included in the initial linkage analysis.

In conclusion, we found a quantitative trait locus at chromosome 5q14.1 affecting normal PC plasma level variability. Next-generation sequencing revealed four SNPs in *BHMT2*, *ACOT12*, *SSBP2* and *XRCC4* that significantly increased PC levels in our thrombophilic

family. Our findings could not be replicated in a population-based case-control study, which might suggest that our findings are family-specific and of limited importance. However, it might also be that the association between the identified variants and protein C levels could only be found in a genetically homogeneous family such as the Vermont family. Further mechanistic studies will be necessary to determine the role of the four identified candidate genes in influencing protein C levels and the possibility of extrapolation to the general population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank all participating Vermont family members and all LETS participants.

We would also like to thank E. Strengman, A. Minten, R. van 't Slot and P.J. Noordijk for performing laboratory analyses, and S. van Lieshout for performing next-generation sequencing bio-informatics analyses.

Funding source

The Vermont family study was supported by NHS grant 2009B084, NHLBI grant PHS HL46703 and a grant from the Leducq Foundation, Paris, France for the development of Transatlantic Networks of Excellence in Cardiovascular Research (grant 04 CVD 02). The Leiden Thrombophilia Study was supported by The Netherlands Heart Foundation (grant 89.063). The funders had no role in data gathering, data analysis and presentation of the results.

References

1. Walker FJ, Sexton PW, Esmon CT. The inhibition of blood coagulation by activated Protein C through the selective inactivation of activated Factor V. *Biochim Biophys Acta*. 1979; 571:333–42. [PubMed: 508770]
2. Fulcher CA, Gardiner JE, Griffin JH, Zimmerman TS. Proteolytic inactivation of human factor VIII procoagulant protein by activated human protein C and its analogy with factor V. *Blood*. 1984; 63:486–9. [PubMed: 6419799]
3. Bovill EG, Bauer KA, Dickerman JD, Callas P, West B. The clinical spectrum of heterozygous protein C deficiency in a large New England kindred. *Blood*. 1989; 73:712–7. [PubMed: 2521802]
4. Vossen CY, Conard J, Fontcuberta J, Makris M, van der Meer FJ, Pabinger I, Palareti G, Preston FE, Scharrer I, Souto JC, Svensson P, Walker ID, Rosendaal FR. Risk of a first venous thrombotic event in carriers of a familial thrombophilic defect. The European Prospective Cohort on Thrombophilia (EPCOT). *J Thromb Haemost*. 2005; 3:459–64. [PubMed: 15748234]
5. Allaart CF, Poort SR, Rosendaal FR, Reitsma PH, Bertina RM, Briët E. Increased risk of venous thrombosis in carriers of hereditary protein C deficiency defect. *Lancet*. 1993; 341:134–8. [PubMed: 8093743]
6. Spek CA, Koster T, Rosendaal FR, Bertina RM, Reitsma PH. Genotypic variation in the promoter region of the protein C gene is associated with plasma protein C levels and thrombotic risk. *Arterioscler Thromb Vasc Biol*. 1995; 15:214–8. [PubMed: 7749828]
7. Souto JC, Almasy L, Borrell M, Garí M, Martínez E, Mateo J, Stone WH, Blangero J, Fontcuberta J. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation*. 2000; 101:1546–51. [PubMed: 10747348]
8. Vossen CY, Hasstedt SJ, Rosendaal FR, Callas PW, Bauer KA, Broze GJ, Hoogendoorn H, Long GL, Scott BT, Bovill EG. Heritability of plasma concentrations of clotting factors and measures of a prethrombotic state in a protein C deficient family. *J Thromb Haemost*. 2004; 2:1–6.
9. Buil A, Soria JM, Souto JC, Almasy L, Lathrop M, Blangero J, Fontcuberta J. Protein C levels are regulated by a quantitative trait locus on chromosome 16: results from the Genetic Analysis of

- Idiopathic Thrombophilia (GAIT) Project. *Arterioscler Thromb Vasc Biol.* 2004; 24:1321–5. [PubMed: 15142863]
10. Tang W, Basu S, Kong X, Pankow JS, Aleksic N, Tan A, Cushman M, Boerwinkle E, Folsom AR. Genome-wide association study identifies novel loci for plasma levels of protein C: the ARIC study. *Blood.* 2010; 116:5032–6. [PubMed: 20802025]
 11. Reiner AP, Carty CL, Jenny NS, Nievergelt C, Cushman M, Stearns-Kurosawa DJ, Kurosawa S, Kuller LH, Lange LA. PROC, PROCr and PROS1 polymorphisms, plasma anticoagulant phenotypes, and risk of cardiovascular disease and mortality in older adults: the Cardiovascular Health Study. *J Thromb Haemost.* 2008; 6:1625–32. [PubMed: 18680534]
 12. van der Meer FJ, Koster T, Vandenbroucke JP, Briët E, Rosendaal FR. The Leiden Thrombophilia Study (LETS). *Thromb Haemost.* 1997; 78:631–5. [PubMed: 9198229]
 13. Sala N, Owen WG, Collen D. A functional assay of protein C in human plasma. *Blood.* 1984; 63:671–5. [PubMed: 6199055]
 14. Hasstedt SJ, Scott BT, Callas PW, Vossen CY, Rosendaal FR, Long GL, Bovill EG. Genome scan of venous thrombosis in a pedigree with protein C deficiency. *J Thromb Haemost.* 2004; 2:868–73. [PubMed: 15140118]
 15. Harakalova M, Mokry M, Hrdlickova B, Renkens I, Duran K, van Roekel H, Lansu N, van Roosmalen M, de Bruijn E, Nijman IJ, Kloosterman WP, Cuppen E. Multiplexed array-based and in-solution genomic enrichment for flexible and cost-effective targeted next-generation sequencing. *Nat Protoc.* 2011; 6:1870–86. [PubMed: 22051800]
 16. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009; 25:1754–60. [PubMed: 19451168]
 17. Nickerson DA, Tobe VO, Taylor SL. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res.* 1997; 25:2745–51. [PubMed: 9207020]
 18. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet.* 1998; 62:1198–211. [PubMed: 9545414]
 19. Almasy L, Blangero J. Exploring positional candidate genes: linkage conditional on measured genotype. *Behav Genet.* 2004; 34:173–7. [PubMed: 14755182]
 20. Hasstedt SJ, Thomas A. Detecting pleiotropy and epistasis using variance components linkage analysis in jPAP. *Hum Hered.* 2011; 72:258–63. [PubMed: 22189468]
 21. Lin PI, Vance JM, Pericak-Vance MA, Martin ER. No gene is an island: the flip-flop phenomenon. *Am J Hum Genet.* 2007; 80:531–8. [PubMed: 17273975]
 22. Couture P, Bovill EG, Demers C, Simard J, Delage R, Scott BT, Valliere JE, Callas PW, Jomphe M, Rosendaal FR, Aiach M, Long GL. Evidence of a founder effect for the protein C gene 3363 inserted C mutation in thrombophilic pedigrees of French origin. *Thromb Haemost.* 2001; 86:1000–6. [PubMed: 11686315]
 23. Rodgers GM, Conn MT. Homocysteine, an atherogenic stimulus, reduces protein C activation by arterial and venous endothelial cells. *Blood.* 1990; 75:895–901. [PubMed: 2154269]
 24. Giusti B, Saracini C, Bolli P, Magi A, Martinelli I, Peyvandi F, Rasura M, Volpe M, Lotta LA, Rubattu S, Mannucci PM, Abbate R. Early-onset ischaemic stroke: analysis of 58 polymorphisms in 17 genes involved in methionine metabolism. *Thromb Haemost.* 2010; 104:231–42. [PubMed: 20458436]
 25. Kircher M, Kelso J. High-throughput DNA sequencing--concepts and limitations. *Bioessays.* 2010; 32:524–36. [PubMed: 20486139]

Figure 1.
Work flow diagram

Figure 2.
Results from the genome scan for normal PC plasma levels including 275 members from the Vermont family for all autosomal chromosomes

Figure 3.
Results from the genome scan for normal PC plasma levels including 275 members from the Vermont family for chromosome 5

Table 1

Candidate gene variants identified in 8 family members

Gene	Variant	Location	Carriers, N	
			low PC (68–75% of normal) N=4	high PC (187–207% of N=4
<i>BHMT2</i>	rs682985	Coding region	4 (4 homozygous)	3
<i>RASGRF2</i>	rs34193571	Coding region	3	0
<i>DMGDH</i>	rs2253262	5' UTR	4	1
<i>ACOT12</i>	rs10371	Coding region	2	4 (3 homozygous)
<i>CMYA5</i>	rs10942901	Coding region	0	3
	rs62621912	Coding region	0	3
<i>SSBP2</i>	rs16878630	3' UTR	0	3
	rs78668688	3' UTR	0	3
<i>XRCC4</i>	rs2035990	3' UTR	0	3

Abbreviations: PC=protein C, UTR=untranslated

Table 2

Influence of gene variants on the LOD score at 5q14.1 in 273 family members

Gene	Variant	N	MAF	Original LOD score	LOD score conditional on the variant	p-value*
<i>BHMT2</i>	rs682985	273	0.42	3.5	2.5	0.004
<i>RASGRF2</i>	rs34193571	269	0.12	3.5	3.5	0.55
<i>DMGDH</i>	rs2253262	255	0.42	2.5	2.1	0.10
<i>ACOT12</i>	rs10371	273	0.29	3.5	2.8	0.02
<i>CMYA5</i>	rs10942901	273	0.18	3.5	3.2	0.08
	rs62621912	272	0.17	3.1	2.9	0.30
<i>SSBP2</i>	rs16878630	271	0.16	3.1	1.9	0.003
<i>XRCC4</i>	rs2035990	264	0.15	3.1	1.8	0.006

Abbreviations: MAF=minor allele frequency, LOD= logarithm (base 10) of odds

* p-values for the χ^2 -statistic

Table 3

Genotype-specific PC activity levels in 273 family members

Gene	Variant	Total N	Major allele homozygotes		Heterozygotes		Minor allele homozygotes		p-value [#]
			N	Mean PC level (SE) [*]	N	Mean PC level (SE) [*]	N	Mean PC level (SE) [*]	
<i>BHMT2</i>	rs682985	273	89	112.95 (3.89)	137	118.06 (3.30)	47	128.53 (4.42)	0.003
<i>ACOT12</i>	rs10371	273	139	115.60 (3.25)	107	122.47 (3.48)	27	133.32 (6.55)	0.005
<i>SSBP2</i>	rs16878630	271	192	116.07 (3.14)	73	125.94 (4.10)	6	141.67 (10.88)	0.002
<i>XRCC4</i>	rs2035990	264	187	115.98 (3.15)	74	127.16 (4.04)	3	140.10 (14.19)	0.002

Abbreviations: PC=protein C, SE=standard error

^{*} Levels are in % of normal and adjusted for family structure, age and sex[#] 1 versus 3 means

Table 4

Genotype-specific PC activity levels in 457 LETS control individuals

Gene	Variant	Total N	Major allele homozygotes		Heterozygotes		Minor allele homozygotes		p-value #
			N	Mean PC level (SE) *	N	Mean PC level (SE) *	N	Mean PC level (SE) *	
<i>BHMT2</i>	rs682985	457	187	103.0 (1.2)	212	104.6 (1.2)	58	100.2 (2.2)	0.191
<i>ACOT12</i>	rs10371	457	338	103.8 (0.9)	113	102.0 (1.6)	6	108.3 (6.9)	0.471
<i>SSBP2</i>	rs16878630	456	408	103.7 (0.8)	48	100.8 (2.4)	0	--	0.260
<i>XRCC4</i>	rs2035990	453	404	103.7 (0.8)	49	101.2 (2.4)	0	--	0.332

Abbreviations: LETS=Leiden Thrombophilia Study, PC=protein C, SE=standard error

* Levels are in % of normal and adjusted for age and sex

1 mean versus 3 means for *BHMT2* rs682985 and *ACOT12* rs10371, and 1 mean versus 2 means for *SSBP2* rs16878630 and *XRCC4* rs2035990

Table 5

Sequencing coverage for the four candidate genes of interest

Gene	Mean % coverage				
	Low PC (68–75% of normal)		high PC (187–207% of normal)		
	N=4		N=4		
	1x	10x	20x	1x	10x 20x
BHMT2	91	85	69	90	81 61
ACOT12	92	84	71	92	82 69
SSBP2	83	62	44	86	67 41
XRCC4*	86	82	71	86	85 80

Abbreviations: PC=protein C

* Coverage described of transcript enst00000396027. Results were similar for transcript enst0000038635.