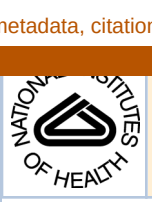

Published in final edited form as:
J Thromb Haemost. 2008 October ; 6(10): 1625–1632. doi:10.1111/j.1538-7836.2008.03118.x.

PROC, *PROCR*, and *PROS1* polymorphisms, plasma anticoagulant phenotypes, and risk of cardiovascular disease and mortality in older adults: the Cardiovascular Health Study

Alexander P. Reiner[*], **Cara L. Carty**[*], **Nancy S. Jenny**[†], **Caroline Nievergelt**[‡], **Mary Cushman**[†], **Deborah J. Stearns-Kurosawa**[§], **Shinichiro Kurosawa**[§], **Lewis H. Kuller**[¶], and **Leslie A. Lange**[**]

[*] Department of Epidemiology University of Washington

[†] Departments of Pathology and Medicine, University of Vermont College of Medicine

[‡] Department of Psychiatry University of California, San Diego

[§] Department of Pathology and Laboratory Medicine, Boston University

[¶] Department of Epidemiology, University of Pittsburgh

[**] Department of Genetics University of North Carolina

Abstract

Background & Objectives—Genes encoding protein C anticoagulant pathways are candidates for athero-thrombotic and other aging-related disorders.

Methods—Using a tagSNP approach, and data from the Cardiovascular Health Study (CHS), we assessed associations of common polymorphisms of *PROC*, *PROS1*, and *PROCR* with (1) plasma protein C, soluble protein C receptor (sEPCR), and protein S levels measured in a sub-sample of 336 participants at study entry; (2) risk of incident clinical outcomes (coronary heart disease or CHD, stroke, and mortality) in 4,547 participants during follow-up. Secondarily, we explored associations between plasma protein C, S, and sEPCR levels and other candidate genes involved in thrombosis, inflammation, and aging.

Results—The *PROCR* Ser219Gly polymorphism (*rs*867186) was strongly associated with higher sEPCR levels, explaining 75% of the phenotypic variation. The Ser219Gly variant was also associated with higher levels of circulating protein C antigen. An *IL10* polymorphism was associated with higher free protein S levels. The minor alleles of *PROC rs*2069901 and *PROS1 rs*4857343 were weakly associated with lower protein C and free protein S levels, respectively. There was no association between *PROCR* Ser219Gly and risk of CHD, stroke, or mortality. The minor allele of another common *PROCR* tagSNP, *rs*2069948, was associated with lymphoid *PROCR* mRNA expression and with increased risk of incident stroke and all-cause mortality, and decreased healthy survival during follow-up.

Conclusions—A common *PROCR* variant may be associated with decreased healthy survival in older adults. Additional studies are warranted to establish the role of *PROCR* variants in ischemic and aging-related disorders.

Keywords

protein C; protein C receptor; protein S; stroke; aging

Send Correspondence To: Alex Reiner, Department of Epidemiology, Box 357236, University of Washington, Seattle, Washington 98195, Phone Number: 206-685-9062, FAX Number: 206-543-8525, apreiner@u.washington.edu.

INTRODUCTION

Protein C is activated on endothelium by the thrombin-thrombomodulin-endothelial protein C receptor (EPCR) complex [1]. In the presence of cofactor protein S, activated protein C (APC) proteolytically inactivates coagulation factors VIIIa and Va, thereby inhibiting clot formation. By binding to EPCR and protease-activated receptor-1, APC also exerts anti-inflammatory and cytoprotective effects on a variety of cell types [2].

The Ser219Gly variant (*rs*867186) tags a common *PROCR* haplotype, and is associated with higher soluble EPCR (sEPCR) levels, explaining ~85% of the phenotypic variance [3,4]. The heritabilities of circulating protein C, free protein S, and total protein S range between 30–50% [5,6], but the specific genetic factors responsible are not as well characterized [7,8]. Promoter polymorphisms of the protein C gene (*PROC*) have been reported to account for ~5% of the phenotypic variability [9,10].

Mutations and polymorphisms of *PROC*, *PROCR*, and *PROS1* (coding for protein S) are associated with risk of familial venous thrombotic disease. Recently, the *PROCR* Ser219Gly variant was associated with increased thrombin generation and increased risk of coronary heart disease (CHD) in men from the Northwick Park Heart Study [11]. The role of Ser219Gly, or other common polymorphisms of the *PROC*, *PROCR*, and *PROS1* genes, in risk of arterial thrombotic disease or mortality have not been carefully examined.

Here, using a tagSNP approach, and data from the population-based Cardiovascular Health Study (CHS), we assessed associations between common polymorphisms and haplotypes of the *PROC*, *PROS1*, and *PROCR* genes and (a) plasma protein C, sEPCR and protein S levels measured in a cross-sectional sub-sample of 336 participants at study entry, and (b) risk of incident clinical outcomes (MI, stroke, and mortality) in 4,547 participants during follow-up. Secondarily, we explored associations of other candidate genes involved in thrombosis, inflammation, and aging with plasma protein C, sEPCR, and protein S levels.

METHODS

The CHS population, inclusion criteria for the current study, and follow-up for clinical events, including years of healthy life [12], are described in detail under Supplemental Methods. A total of 336 men and women with baseline protein C, S and sEPCR measurements not taking warfarin were included in the baseline cross-sectional analysis. The number of participants eligible for analysis of clinical events during follow-up was 4,547. All study participants provided written informed consent for use of their DNA for genetic testing.

Blood collection and analysis

Baseline blood was collected in a fasting state, and a special tube designed to prevent in vitro clotting activation (SCAT-1, Haematologic Technologies, Inc., Essex Junction, VT) was used [13]. Blood samples were analyzed at the Central CHS Laboratory at the University of Vermont. Protein C antigen, free protein S (unbound to C4b-binding protein) total protein S (free + C4b-binding protein-bound), and sEPCR were measured by enzyme linked immunosorbent assays (ELISAs) as previously described [14,15]. The assay CVs were 2.5%, 9.9%, 6.7%, and 3%, respectively.

TagSNP selection

Single nucleotide polymorphisms (SNPs), pair-wise linkage disequilibrium (LD) patterns, and haplotypes for our candidate genes were identified from the SeattleSNPs candidate gene

SNP discovery resource and database (http://pga.mbt.washington.edu/). Polymorphic sites were identified by direct re-sequencing of genomic sequence from 23 European-Americans encompassing all exons, introns, untranslated regions and ~ 2 kb of additional flanking sequence on either 5′ or 3′ end. For *PROC*, *PROCR*, and *PROS1*, a total of 39, 13, and 44 polymorphic sites were identified by SeattleSNPs, of which 28, 8, and 17 had minor allele frequency (MAF)≥10%. TagSNPs were identified using the pair-wise LD binning procedure implemented in the LDSelect algorithm of Carlson et al [16], at a ≥10% MAF threshold and a linkage disequilibrium (LD) threshold of $r^2 \geq 0.64$ to create bins. Using this procedure, we identified 4 tagSNP bins for *PROC*, 2 for *PROCR*, and 4 for *PROS1* (see Table 1). One *PROS1* tagSNP bin could not be assayed due to assay failure. Therefore linkage disequilibrium coverage in *PROC, PROS1*, and *PROCR* for untyped SNPs with minor allele frequencies of 10% or greater was 100%, 75%, and 100%, respectively at an r-squared threshold of ≥0.64. TagSNPs were similarly selected for the remaining 127 candidate genes using SeattleSNPs or HapMap databases at an LD threshold of $r^2 \geq 0.64$ and MAF ≥10%.

SNP genotyping

PROC, PROS1, and *PROCR* genotyping was performed in all consenting CHS participants at the Laboratory for Clinical Biochemistry Research (University of Vermont) with the ABI TaqMan platform using Assays By Design on an ABI 7900 real time thermal cycler under standard conditions (Applied Biosystems, Foster City, CA). Overall genotype missing rate was <0.1%, and blind duplicate concordance rates were >99%. Genotype distributions were in Hardy-Weinberg equilibrium for all *PROC, PROCR*, and *PROS1* tagSNPs, with the exception of *PROS1 rs*4857343 (uncorrected p=0.002). Additional genotyping was performed using the Illumina GoldenGate platform by the Center for Inherited Disease Research (CIDR; Johns Hopkins University, Baltimore, MD) for 876 tagSNPs covering common linkage disequilibrium patterns across 127 thrombosis, inflammation, and aging-related genes, as previously described in detail [17].

Statistical analysis

Our primary hypotheses were to test whether common polymorphisms of the *PROC, PROCR*, and *PROS1* genes were associated with (a) protein C, sEPCR, and protein S levels; (b) risk of incident clinical CVD and aging-related outcomes. Secondarily, we also explored associations of other candidate genes involved in thrombosis, inflammation, and aging on plasma protein C, sEPCR, and protein S levels. Cross-sectional relationships between baseline protein C, S and sEPCR measurements were assessed using Spearman's rank correlation coefficient (ρ). Associations between individual SNP genotypes and baseline quantitative phenotypes were assessed using multiple linear regression models. All regression models were minimally adjusted for age and sex. To reduce the influence of environmental variation, protein C and protein S regression models were additionally adjusted for cholesterol and triglycerides (which were strong cross-sectional correlates of protein C and S levels). Covariate-adjusted mean adjusted plasma anticoagulant levels (and 95% confidence intervals) were estimated for each genotype group from the regression coefficients (β) and standard errors. Associations between individual tagSNP genotypes and risk of incident CVD events or mortality during follow-up were assessed using Cox regression, adjusted for major clinical risk factors (age, sex, diabetes, hypertension, clinical CVD status, smoking, and serum creatinine). Covariate-adjusted hazard ratios were estimated from the regression coefficient, assuming a constant effect size for each additional copy of the minor allele. In analyses of incident MI and stroke during follow-up, we excluded participants who had experienced an MI (n=447) or stroke (n=165) prior to baseline. Assessment and adjustment of potential within-Europe population structure during our primary candidate gene analyses was performed as described under Supplemental Methods.

We used various procedures to control for multiple hypotheses testing during our primary and secondary analyses. For each of our primary candidate genes, we conducted a "gene-wide" test of significance by performing 100,000 permutations of the data sampled under the null hypothesis. Empirical *p*-values were determined by counting the proportion of times the observed test statistic was greater than the maximum value of the permuted null test statistics across all SNPs in each gene, thereby accounting for the number of SNPs tested and the variable correlation structure, while controlling the experiment wise type 1 error rate at 5% for each gene [18]. We also used statistical methods to infer haplotypes from the genotype data, using the probability-weighted haplotypes as the unit of analysis [19]. To further correct our primary analyses of *PROC*, *PROCR*, and *PROS1* genotypes for testing of multiple, correlated traits, we conducted an "experiment-wise" test of significance using a procedure that computes empirical p-values while retaining the original correlation structure among both genotypes and traits under the null hypothesis by simulation from a multivariate normal distribution [20]. To correct for multiple comparisons in our secondary analysis of 876 tagSNPs from 130 candidate genes with protein C, S, and sEPCR levels, we assessed overall statistical significance using the false discovery rate (FDR) *q* value, which is an estimate of the proportion of features called significant that are truly null [21].

mRNA expression analysis

To assess associations between *PROCR* genotype and gene expression, we utilized the Sanger GENe Expression VARiation project (GENEVAR), a web-accessible collection of genome-wide microarray-based gene expression measurements obtained from Epstein-Barr virus-transformed lymphoblastoid cell lines from 60 unrelated European-American HapMap (CEPH) samples [22]. Differences in mRNA expression levels for genes of interest were tested between genotype groups using ANOVA.

RESULTS

Participant characteristics among those eligible for each analysis (cross-sectional plasma phenotype and longitudinal time to clinical event) are summarized in Table 1. The distributions of protein C, free protein S, and total protein S levels were approximately normal. As previously reported in other studies [3,4], sEPCR had a multi-modal distribution. Mean levels of protein C, sEPCR, free protein S, and total protein S were 3.40 ±5.95 mg/L, 222.55 ±90.10 μg/L, 5.01 ±0.94 mg/L, and 22.38 ±3.14 mg/L, respectively. Total protein S levels were strongly correlated with free protein S (ρ=0.36; p<0.0001) and protein C (ρ=0.31; p<0.0001). Protein C levels were correlated with free protein S (ρ=0.18; p=0.002) and sEPCR (ρ=0.21; p=0.0001). Protein C, free protein S and total protein S were strongly correlated with baseline plasma cholesterol and lipid levels, as previously reported [14]. sEPCR levels initially showed no significant correlation with any traditional cardiovascular risk factors. However, when adjusted for the strong effect of the Ser219Gly polymorphism, male sex (p=0.02) and age (p<0.0001) were associated with higher sEPCR levels.

PROC, PROCR, and *PROS1* genotype and protein C, S, sEPCR levels

Among the entire CHS sample, the r^2 measure of linkage disequilibrium between pairs of SNPs in each gene were all below 0.65. The associations between *PROC*, *PROCR*, *PROS1* tagSNPs, haplotypes and their respective plasma phenotypes are shown in Table 2. Under an additive genetic model, each additional copy of the *PROCR* Ser219Gly polymorphism was associated 182 ±6 μg/L higher sEPCR levels, explaining 75% of the phenotypic variation. In unadjusted models, the minor allele of *PROCR rs*2069948 appeared to be associated with lower sEPCR levels (β =−34 ±7 μg/L); however, adjustment for Ser219Gly greatly weakened the association (β=−3.7 ± 3.8 μg/L; p=0.33). Analysis of the 3 common *PROCR* haplotypes confirmed that the haplotype tagged by Ser219Pro (*PROCR*-3) was strongly

associated with higher sEPCR levels (p<0.0001) relative to *PROCR*-1 containing both major tagSNP alleles, while there was no difference in sEPCR concentration between the haplotype tagged by *rs*2069948 (*PROCR*-2) and *PROCR*-1 (supplemental Table 1). Each additional copy of the *PROC rs*2069901 variant allele was associated with −0.13 ± 0.04 lower mg/L protein C levels (gene-wide *p*=0.01). *PROC* haplotypes containing the *rs*2069901 variant allele were also associated with lower protein C (supplemental Table 1). Each additional copy of the *PROS1 rs*4857343 minor allele (β=−0.28 ± 0.10 mg/L; *p*=0.005), or alternatively, the two haplotypes containing the *rs*4857343 minor allele (supplemental Table 3), was associated with lower free protein S levels. Upon experiment-wide correction for testing multiple genes and phenotypes, the association between Ser219Gly and sEPCR levels was still strongly significant (*p*<0.0001); however, neither the protein C nor the free protein S genotype-phenotype associations were significant (*p*=0.11 and 0.14, respectively).

Additional thrombosis/inflammation gene – protein C, S, EPCR phenotype associations

We screened 876 additional tagSNPs across 130 candidate genes for association with plasma protein C, S, and sEPCR levels. The tagSNP identifiers, location, regression beta coefficients, standard errors, p-values and false discovery rate *q*-values are shown in Supplemental Table 4. At a false discovery rate threshold of *q*<0.05, the *PROCR* Ser219Gly polymorphism was associated with higher protein C levels (β = 0.63 ± 0.09 mg/L; *p*-value 9×10^{-11}; *q*-value <1 $\times10^{-8}$), explaining 13% of the variation in protein C. The minor allele of an interleukin-10 gene (*IL10*) intronic polymorphism on chromosome 1 (*rs*1878672) was associated with higher free protein S levels (β = 0.34 ± 0.08 mg/L; *p*-value 3×10^{-5}; *q*-value = 0.03), explaining 5% of the variation in free protein S. The same *IL10 rs*1878672 variant was associated with 0.16 ± 0.06 mg/L higher protein C levels (*p*-value 0.006), though the protein C association did not meet the same significance threshold (*q*-value =0.50) as the free protein S association.

PROC, PROCR, and *PROS1* genotype and risk of incident CVD and mortality

There was no evidence for association between *PROC* or *PROS1* genotype and risk of incident CVD or death. On the other hand, each additional copy of the minor allele of *PROCR rs*2069948 was associated with 13% ± 6% (p=0.02) increased risk of stroke and 7% ± 3% increased risk of mortality (p=0.01), even after adjusting for major risk factors (Table 3). Moreover, each additional copy of the *PROCR rs*2069948 variant was associated 0.32 ± 0.09 fewer years of healthy life (p=0.0003). There was no association between Ser219Gly and risk of incident CVD or mortality. The *PROCR* genotype-clinical phenotype results did not differ among subgroups defined according to age, gender, diabetes, or baseline subclinical CVD status. Upon experiment-wide correction for multiple testing, only the *PROCR rs*2069948 association with years of healthy life remained statistically significant (*p*=0.01).

Association between *PROCR rs*2069948 and cellular gene expression

Examination of linkage disequilibrium patterns surrounding the *PROCR* gene on chromosome 20 revealed that *PROCR rs*2069948 was strongly correlated with several other SNPs located in *PROCR* as well as with additional SNPs located outside *PROCR*. Supplemental Table 5 shows the location, function, and pair-wise correlation coefficient for all other SNPs showing strong linkage disequilibrium (r^2>0.5) in a 200 kb window surrounding *PROCR rs*2069948. We assessed the relationship between *PROCR* genotype and lymphoid mRNA expression levels of *PROCR* and neighboring chromosome 20 genes in 60 European-American HapMap individuals (Table 4). *rs*2069948 was associated with lower *PROCR* mRNA expression and also with higher and lower mRNA expression levels, respectively, of *EDEM2* (ER degradation-enhancing-mannosidase-like protein 2) and *GSS*

(glutathione or GSH synthase). *EDEM2* and *GSS* are both involved in responses to cellular damage that accumulate during aging. *PROCR* Ser219Gly polymorphism was neither associated with *PROCR* mRNA expression, nor with mRNA expression of any other genes in the region (Table 4).

DISCUSSION

In older European-American men and women from CHS, the *PROCR* Ser219Gly variant was strongly associated with higher sEPCR levels. There were much weaker associations between common variants of the *PROC* and *PROS1* genes and protein C and free protein S levels, respectively. In an exploratory analysis that included a larger set of thrombosis- and inflammation-related genes, an *IL10* polymorphism was associated with higher free protein S and protein C levels. In addition, the *PROCR* Ser219Gly variant was associated with higher levels of protein C. There was no evidence of association between Ser219Gly and risk of clinical CVD. On the other hand, another common *PROCR* variant *rs*2069948 was associated with increased risk of stroke and mortality and decreased healthy survival during follow-up and with reduced *PROCR* mRNA expression in lymphoid cells.

Our results confirm that the *PROCR* Ser219Gly polymorphism is the major determinant of phenotypic variation in sEPCR [3,4]. In contrast, the majority of inter-individual variation in protein C and protein S levels does not appear to be explained by common polymorphisms of *PROC* or *PROS1*, respectively. The *PROC rs*2069901 variant allele weakly associated with lower protein C levels in CHS tags the same common “CG” haplotype as the *PROC* -1654 C/A and -1641 G/T promoter polymorphisms previously reported to be associated with lower plasma levels [9,10] and lower transcriptional efficiency of protein C [23]. The *PROS1* variant associated with lower free protein S levels in CHS is located in intron 2 and is in strong linkage disequilibrium with 2 other polymorphisms *rs*8178583 and *rs*7650230 located in the 5′ flanking region, but outside of the minimal *PROS1* promoter [24,25]. While our tagSNP approach efficiently captures LD patterns among common SNPs across our candidate genes, we were not able to capture one common *PROS1* LD bin (containing 4 SNPs with MAF ≥10%) due to assay failure. Moreover, the common tagSNP approach design does not address the possibility that rare variants within these same genes, such as the previously reported protein S Heerlen Ser460Pro variant [26] (estimated population prevalence ~ 0.5%), may contribute significantly to phenotypic variance.

Since protein C and sEPCR and levels were moderately correlated, it is interesting to note that the *PROCR* Ser219Gly variant was a fairly strong determinant of protein C, explaining a greater amount of the phenotypic variation than the *PROC* gene itself. The mechanism of association between *PROCR* Ser219Gly and higher sEPCR levels appears to involve increased EPCR shedding from the endothelium [27], which may result in reduced protein C activation [28]. Therefore, it is possible that higher sEPCR levels relative to membrane-bound EPCR associated with Ser219Gly might effectively increase the amount of circulating protein C by stabilizing it.

We also observed a moderately strong correlation between protein C, protein S, and plasma lipid levels. Protein C and protein S are functionally- and structurally-related vitamin K-dependent proteins. It is possible that common polymorphisms of shared regulatory genes account for at least some of the correlation in their plasma concentration. In this regard, it is interesting to note that *IL10* genotype was associated with both free protein S and protein C levels. *IL10* is located in a region on chromosome 1q32, previously identified as a QTL for protein S in a genome-wide linkage analysis [7]. The chromosome 1q32 region contains two genes encoding subunits of C4 binding protein (*C4BPA* and *C4BPB*), an inflammation-sensitive protein. The beta subunit of C4 binding protein binds strongly to protein S and

effectively determines free protein S levels. *C4BPA* and *C4BPB* are located ~300 kb from the *IL10* gene. Examination of HapMap data showed only weak linkage disequilibrium between the associated *IL10* variants and the *C4BP* gene region, with all pair-wise SNP r^2 with IL10 *rs*4857343 < 0.12. Therefore additional fine-mapping in this extended region on chromosome 1q32, which also contains several other complement and interleukin-related genes, will be required to further characterize the variant(s) responsible for the protein S and protein C phenotypic associations.

A recent report suggested that *PROCR* Ser219Gly was associated with increased thrombin generation and increased risk of CHD in men and in diabetics or those with metabolic syndrome [11]. Our results do not support this finding, either overall or in subgroups defined according to age, gender, or diabetes. Failure to replicate the Ser219Gly CVD association in CHS could be related to differences in sample characteristics between studies. Another potential issue is that the risk associated with *PROCR* Ser219Gly in the report by Ireland et al appeared to be limited to homozygotes for the minor Gly219 allele [11]. Gly219 homozygotes comprised only 0.8% of the CHS cohort; therefore even larger sample sizes may be required to detect a recessive effect.

Based on data from the HapMap, the *PROCR rs*2069948 variant associated with healthy aging in CHS is part of an extended haplotype on chromosome 20 that contains 45 other polymorphisms, including 9 *PROCR* SNPs in complete linkage disequilibrium, as well as several variants located in neighboring genes (Supplemental Table 5). One of the SNPs comprising the risk haplotype, *rs*6088747, is located within an enhancer region 5.5 kb upstream of the *PROCR* translation start site that is essential for *PROCR* mRNA expression inendothelium and hematopoietic cells [29]. Another SNP, *rs*9574, is located within the proximal 3′ untranslated region of *PROCR*. It is possible that one or more of these polymorphism influences *PROCR* mRNA stability or processing, as suggested by the observed association between *rs*2069948 and decreased lymphoid *PROCR* mRNA expression. The precise molecular mechanism of this association requires further study, including assessment of the relative effects of *PROCR* genotype on membrane versus soluble forms as well as cell type-specific mRNA expression. It should also be noted that the same *PROCR* haplotype tagged by *rs*2069948 has been associated with higher circulating activated protein C levels and decreased risk of venous thrombosis [30], though other groups have reported that the Ser219Gly haplotype, but not the *rs*2069948 haplotype, is associated with increased risk of venous thrombosis [3,4].

Because of the anticoagulant, anti-inflammatory, cytoprotective, and immuno-modulatory properties of the protein C system [1,2], genes encoding protein C pathway components are potential candidates not only for athero-thrombotic and ischemic conditions [31], but also for inflammatory and other aging-related disorders. Endothelial protein C receptor is expressed by various blood cells, and binding of protein C arrests cellular migration [32,33]. The effects of the protein C/EPCR pathway on endothelial or immune cell function may not only be related to the apparent influence of the *PROCR rs*2069948 haplotype on healthy aging, but also to the beneficial effects of activated protein C in the treatment of severe sepsis [34,35]. Finally, there was some evidence of association between *rs*2069948 and expression of other genes in the region, *EDEM2* and *GSS*, which are involved in protein unfolding response and oxidative damage responses, respectively. These data additionally suggest the possibility that other cellular processes may play a role in the apparent effect of the *PROCR* risk haplotype on thrombosis, mortality or healthy aging.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by contracts N01-HC-85079through N01-HC-85086, N01-HC-35129, N01 HC-55222, N01 HC-15103, grants R01 HL-071862 and U01 HL080295 from the NHLBI and U19 AG023122 from the National Institute on Aging Longevity Consortium, with additional contribution from the National Institute of Neurological Disorders and Stroke. A full list of participating CHS investigators and institutions can be found at http://www.chs-nhlbi.org. Genotyping services for CHS were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, Contract Number N01-HG-65403. Genotyping services for CHS were also provided by the Johns Hopkins University under federal contract number (N01-HV-48195) from the National Heart, Lung, and Blood Institute. We utilized public re-sequencing data from the SeattleSNPs program, supported by NIH U01 HL66682 (http://www.pga.gs.washington.edu/).

References

1. Dahlback B, Villoutreix BO. The anticoagulant protein C pathway. FEBS Lett. 2005; 579:3310–6. [PubMed: 15943976]
2. Mosnier LO, Zlokovic BV, Griffin JH. The cytoprotective protein C pathway. Blood. 2007; 109:3161–72. [PubMed: 17110453]
3. Uitte de Willige S, Van Marion V, Rosendaal FR, Vos HL, de Visser MC, Bertina RM. Haplotypes of the EPCR gene, plasma sEPCR levels and the risk of deep venous thrombosis. J Thromb Haemost. 2004; 2:1305–10. [PubMed: 15304035]
4. Saposnik B, Reny JL, Gaussem P, Emmerich J, Aiach M, Gandrille S. A haplotype of the EPCR gene is associated with increased plasma levels of sEPCR and is a candidate risk factor for thrombosis. Blood. 2004; 103:1311–8. [PubMed: 14576048]
5. Souto JC, Almasy L, Borrell M, Gari M, Martinez E, Mateo J, Stone WH, Blangero J, Fontcuberta J. Genetic determinants of hemostasis phenotypes in Spanish families. Circulation. 2000; 101:1546–51. [PubMed: 10747348]
6. 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; 2:242–7. [PubMed: 14995985]
7. Almasy L, Soria JM, Souto JC, Coll I, Bacq D, Faure A, Mateo J, Borrell M, Munoz X, Sala N, Stone WH, Lathrop M, Fontcuberta J, Blangero J. Genetic Analysis of Idiopathic Thrombophilia project. A quantitative trait locus influencing free plasma protein S levels on human chromosome 1q: results from the Genetic Analysis of Idiopathic Thrombophilia (GAIT) project. Arterioscler Thromb Vasc Biol. 2003; 23:508–11. [PubMed: 12615682]
8. 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]
9. 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]
10. Aiach M, Nicaud V, Alhenc-Gelas M, Gandrille S, Arnaud E, Amiral J, Guize L, Fiessinger JN, Emmerich J. Complex association of protein C gene promoter polymorphism with circulating protein C levels and thrombotic risk. Arterioscler Thromb Vasc Biol. 1999; 19:1573–6. [PubMed: 10364092]
11. Ireland H, Konstantoulas CJ, Cooper JA, Hawe E, Humphries SE, Mather H, Goodall AH, Hogwood J, Juhan-Vague I, Yudkin JS, di Minno G, Margaglione M, Hamsten A, Miller GJ, Bauer KA, Kim YT, Stearns-Kurosawa DJ, Kurosawa S. EPCR Ser219Gly: elevated sEPCR, prothrombin F1+2, risk for coronary heart disease, and increased sEPCR shedding in vitro. Atherosclerosis. 2005; 183:283–92. [PubMed: 15921688]
12. Diehr P, Patrick DL, Bild DE, Burke GL, Williamson JD. Predicting future years of healthy life for older adults. J Clin Epidemiol. 1998; 51:343–53. [PubMed: 9539891]

13. Cushman M, Cornell ES, Howard PR, Bovill EG, Tracy RP. Laboratory methods and quality assurance in the Cardiovascular Health Study. Clin Chem. 1995; 41:264–270. [PubMed: 7874780]
14. Sakkinen PA, Cushman M, Psaty BM, Kuller LH, Bajaj SP, Sabharwal AK, Boineau R, Macy E, Tracy RP. Correlates of antithrombin, protein C, protein S, and TFPI in a healthy elderly cohort. Thromb Haemost. 1998; 80:134–9. [PubMed: 9684799]
15. Stearns-Kurosawa DJ, Swindle K, D'Angelo A, Della Valle P, Fattorini A, Caron N, Grimaux M, Woodhams B, Kurosawa S. Plasma levels of endothelial protein C receptor respond to anticoagulant treatment. Blood. 2002 Jan 15; 99(2):526–30. [PubMed: 11781234]
16. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet. 2004; 74:106–120. [PubMed: 14681826]
17. Lange LA, Reiner AP, Carty CA, Jenny NS, Cushman M, Lange EM. Common genetic determinants of plasma fibrin D-dimer concentration in older European- and African-American adults. J Thromb Haemostas. 2008; 6:654–659.
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–75. [PubMed: 17701901]
19. Lin DY. Haplotype-based association analysis in cohort studies of unrelated individuals. Genet Epidemiol. 2004; 26:255–264. [PubMed: 15095385]
20. Conneely KN, Boehnke M. So many correlated tests, so little time! Rapid adjustment of p-values for multiple correlated tests. Am J Hum Genet. 2007; 81:1158–68. [PubMed: 17966093]
21. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A. 2003; 100:9440–5. [PubMed: 12883005]
22. Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, Hunt S, Kahl B, Antonarakis SE, Tavare S, Deloukas P, Dermitzakis ET. Genome-wide associations of gene expression variation in humans. PLoS Genet. 2005; 1:e78. [PubMed: 16362079]
23. Scopes D, Berg LP, Krawczak M, Kakkar VV, Cooper DN. Polymorphic variation in the human protein C (PROC) gene promoter can influence transcriptional efficiency in vitro. Blood Coagul Fibrinolysis. 1995; 6:317–21. [PubMed: 7548679]
24. de Wolf CJ, Cupers RM, Bertina RM, Vos HL. The constitutive expression of anticoagulant protein S is regulated through multiple binding sites for Sp1 and Sp3 transcription factors in the protein S gene promoter. J Biol Chem. 2006; 281:17635–43. [PubMed: 16672217]
25. Hall AJ, Peake IR, Winship PR. Regulation of the human protein S gene promoter by liver enriched transcription factors. Br J Haematol. 2006:135538–46.
26. Duchemin J, Gandrille S, Borgel D, Feurgard P, Alhenc-Gelas M, Matheron C, Dreyfus M, Dupuy E, Juhan-Vague I, Aiach M. The Ser 460 to Pro substitution of the protein S alpha (PROS1) gene is a frequent mutation associated with free protein S (type IIa) deficiency. Blood. 1995; 86:3436–43. [PubMed: 7579448]
27. Qu D, Wang Y, Song Y, Esmon NL, Esmon CT. The Ser219-->Gly dimorphism of the endothelial protein C receptor contributes to the higher soluble protein levels observed in individuals with the A3 haplotype. J Thromb Haemost. 2006; 4:229–35. [PubMed: 16409473]
28. Kurosawa S, Stearns-Kurosawa DJ, Hidari N, Esmon CT. Identification of functional endothelial protein C receptor in human plasma. J Clin Invest. 1997; 100:411–418. [PubMed: 9218519]
29. Mollica LR, Crawley JT, Liu K, Rance JB, Cockerill PN, Follows GA, Landry JR, Wells DJ, Lane DA. Role of a 5′-enhancer in the transcriptional regulation of the human endothelial cell protein C receptor gene. Blood. 2006; 108:1251–9. [PubMed: 16627757]
30. Medina P, Navarro S, Estelles A, Vaya A, Woodhams B, Mira Y, Villa P, Migaud-Fressart M, Ferrando F, Aznar J, Bertina RM, Espana F. Contribution of polymorphisms in the endothelial protein C receptor gene to soluble endothelial protein C receptor and circulating activated protein C levels, and thrombotic risk. Thromb Haemost. 2004; 91:905–11. [PubMed: 15116250]
31. Cheng T, Liu D, Griffin JH, Fernandez JA, Castellino F, Rosen ED, Fukudome K, Zlokovic BV. Activated protein C blocks p53-mediated apoptosis in ischemic human brain endothelium and is neuroprotective. Nat Med. 2003; 9:338–42. [PubMed: 12563316]

32. Sturn DH, Kaneider NC, Feistritzer C, Djanani A, Fukudome K, Wiedermann CJ. Expression and function of the endothelial protein C receptor in human neutrophils. Blood. 2003; 102:1499–505. [PubMed: 12714492]

33. Feistritzer C, Mosheimer BA, Sturn DH, Riewald M, Patsch JR, Wiedermann CJ. Endothelial protein C receptor-dependent inhibition of migration of human lymphocytes by protein C involves epidermal growth factor receptor. J Immunol. 2006; 176:1019–25. [PubMed: 16393989]

34. Kerschen EJ, Fernandez JA, Cooley BC, Yang XV, Sood R, Mosnier LO, Castellino FJ, Mackman N, Griffin JH, Weiler H. Endotoxemia and sepsis mortality reduction by non-anticoagulant activated protein C. J Exp Med. 2007; 204:2439–48. [PubMed: 17893198]

35. Zheng X, Li W, Song Y, Hu Y, Ferrell GL, Esmon NL, Esmon CT. Non-hematopoietic EPCR regulates the coagulation and inflammatory responses during endotoxemia. J Thromb Haemost. 2007; 5:1394–400. [PubMed: 17445091]

Table 1

CHS participant characteristics, by analysis

Characteristic	Cross-sectional plasma phenotype (n =336)	Longitudinal time to clinical event (n=4,745)
Mean age, years (range)	77 ± 7.4 (65–98)	73 ± 5.6 (64–98)
Female sex, %	47	57
Current smokers, %	7	11
Body mass index, kg/m^2, (range)	25.8 ± 4.7 (14.8–46.4)	26.4 ± 4.5 (14.6–48.3)
Total cholesterol, mg/dL, (range)	204 ± 38 (87.8–329.8)	212 ± 39 (58.8–429.8)
LDL cholesterol, mg/dL, (range)	124 ± 34 (32.8–252.8)	130 ± 36 (24.8–336.8)
HDL cholesterol, mg/dL, (range)	55 ± 16 (24–128)	54 ± 16 (15 – 149)
Triglycerides, mg/dL, (range)	129 ± 58 (52 – 410)	144 ± 79 (37 – 1323)
Diagnosis of hypertension, %	56	56
Diabetes, %	11	15
Baseline history of MI or stroke, %	0	22
Protein C level, μg/dL, (range)	3.46 ± 0.73 (1.67–6.31)	--
Free protein S level, μg/dL, (range)	5.01 ± 0.94 (1.12–8.03)	--
Total protein S level, μg/dL, (range)	22.38 ± 3.14 (13.86–30.22)	--
Soluble protein C receptor level, μg/dL, (range)	222.55 ± 90.10 (74.99–700.99)	--

Data are presented as % or mean ± standard deviation, as indicated.

Table 2

Associations of *PROC*, *PROCR*, and *PROS1* genotype with plasma protein C, soluble protein C receptor, and protein S (n=336)

Gene	dbSNP ID	Seattle SNPs ID*	Location	MAF	Estimated mean (95%CI)	P-value
			Protein C (mg/L)			
PROC	*rs*2069901	1386 T/C	5′ flanking	0.45	3.77 (3.66–3.88)	0.01
					3.58 (3.49–3.68)	
					3.53 (3.39–3.68)	
	*rs*2069914	4813 G/A	Intron 2	0.25	3.64 (3.55–3.73)	0.97
					3.60 (3.48–3.71)	
					3.78 (3.54–3.97)	
	*rs*2069915	4919 G/A	Intron 2	0.40	3.58 (3.47–3.69)	0.38
					3.65 (3.55–3.74)	
					3.71 (3.56–3.86)	
	*rs*2069928	10454 G/T	Intron 7	0.19	3.63 (3.55–3.72)	0.94
					3.62 (3.50–3.73)	
					3.82 (3.51–4.12)	
			Soluble EPCR (μg/L)			
PROCR	rs2069948	4054 T/C	Intron 1	0.44	248 (230–266)	$<10^{-5}$†
					215 (200–229)	
					180 (155–205)	
	rs867186	6118 A/G	Ser219Gly	0.09	176 (170–182)	$<10^{-5}$
					364 (353–375)	
					532 (490–573)	
			Total Protein S (mg/L)			
PROS1	rs4857343	49612 G/A	Intron 2	0.20	22.7 (22.2–23.1)	0.39
					22.6 (21.9–23.3)	
					21.4 (19.9–22.9)	
	rs8178626	66205 A/T	Intron 4	0.45	22.6 (21.9–23.3)	0.36
					22.8 (22.3–23.3)	
					21.9 (21.1–22.6)	
	rs4857037	66847 T/C	Intron 4	0.10	22.6 (22.1–23.0)	0.81

Gene	dbSNP ID	Seattle SNPs ID*	Location	MAF	Estimated mean (95%CI)	P-value
					22.4 (21.5–23.3)	
					22.9 (19.4–26.4)	
			Free Protein S (mg/L)			
PROS1	rs4857343	49612 G/A	Intron 2	0.20	5.12 (4.98–5.26)	0.01
					5.07 (4.87–5.27)	
					4.26 (3.82–4.70)	
	rs8178626	66205 A/T	Intron 4	0.45	5.19 (4.98–5.41)	0.10
					5.05 (4.89–5.20)	
					4.98 (4.74–5.22)	
	rs4857037	66847 T/C	Intron 4	0.10	5.09 (4.97–5.22)	0.35
					4.97 (4.70–5.24)	
					4.68 (3.63–5.73)	

CI = confidence interval; MAF = minor allele frequency

Mean phenotype levels for each genotype group were estimated from linear regression coefficients, as described under Methods. P-values are corrected for testing multiple SNPs in each gene, as described under Methods.

*SeattleSNPs. NHLBI Program for Genomic Applications, SeattleSNPs, Seattle, WA (URL: http://pga.gs.washington.edu) [accessed March 2008].

†upon adjustment for the effect of Ser219Gly, the estimated mean sEPCR for each rs2069948 genotype group was 211, 212, and 211 μg/L (p for trend=0.33)

Table 3

Associations of *PROCR* genotype and risk of incident CVD events and mortality during follow-up (n=4,547)

dbSNP	Genotype	Myocardial Infarction Hazard Ratio (95% CI)	Stroke Hazard Ratio (95% CI)	Mortality Hazard Ratio (95% CI)	Years of Healthy Life $\beta \pm S.E.$
rs2069948	TT	1	1	1	
	TC	1.01 (0.88–1.14)	1.17 (0.98–1.38)	1.04 (0.96–1.14)	−0.26 ± 0.14
	CC	1.08 (0.91–1.28)	1.27 (1.02–1.57)	1.16 (1.04–1.29)	−0.65 ± 0.18
		p=0.60	*p*=0.02	*p*=0.01	*p*=0. 0003
rs867186 (Ser219Gly)	AA	1	1	1	
	AG	0.89 (0.76–1.04)	0.93 (0.76–1.13)	0.91 (0.82–1.01)	0.34 ± 0.16
	GG	1.27 (0.75–2.15)	1.14 (0.51–2.55)	1.16 (0.77–1.75)	0.0 ± 0.70
		p=0.21	*p*=0.73	*p*=0.15	*p*=0.11

Adjusted for age, sex, diabetes, hypertension, clinical CVD status, smoking, and serum creatinine. P-values are corrected for testing multiple SNPs in each gene, as described under Methods.

Table 4

Gene expression levels associated with *PROCR* polymorphism genotypes

Gene Symbol	Position on chromosome 20 (bp)	Gene Name	Function	rs2069948 (β ± S.E.)	rs867186 (Ser219Gly) (β ± S.E.)
GSS	32,979,897– 33,007,262	glutathione synthase	cytoprotection from oxidative damage	−0.10 ± 0.05 **(p=0.03)**	0.06 ± 0.09 (p=0.52)
MYH7B	33,026,867– 33,053,897	myosin, heavy chain 7B, cardiac muscle, beta	Muscle contraction	0.02 ± 0.01 (p=0.15)	0.02 ± 0.03 (p=0.45)
TRPC4AP	33,053,868– 33,144,279	tumor necrosis factor receptor-associated ubiquitous scaffolding and signaling protein	TNF receptor-interacting protein	0.02 ± 0.03 (p=0.44)	−0.03 ± 0.05 (p=0.51)
EDEM2	33,166,831– 33,198,804	ER degradation-enhancing-mannosidase-like protein 2	Unfolded ER protein response	0.09 ± 0.04 **(p=0.04)**	0.11 ± 0.09 (p=0.20)
PROCR	33,223,435– 33,228,826	Protein C receptor	Coagulation, inflammation	−0.20 ± 0.05 **(p<0.001)**	−0.10 ± 0.09 (p=0.26)

Genes are listed in order of their position on chromosome 20q11.