# Genetic variation and natural anticoagulants

Yuekai Ji

A thesis

submitted in partial fulfillment of

requirements for the degree of

Master of Science

University of Washington2020

Committee:

Nicholas L. Smith

Sara Lindstroem

Maria Sabater Lleal

Program Authorized to Offer Degree:

Epidemiology

**ABSTRACT**

Protein C (PC), protein S (PS) and antithrombin (AT) belong to the main anticoagulant pathways that regulate hemostasis and further affect the risk of thrombotic events. Previous epidemiological studies of circulating levels of PC, PS and AT, genetic variation, and risk of venous thromboembolism (VTE) have been limited by small sample sizes and restriction to specific candidate genes only. Utilizing summary results from nine genome-wide association studies (GWAS), we conducted transethnic meta-analyses on AT, PC, and free and total PS. We identified potential novel genome-wide significant signals for AT close to the *GCKR*, *SNX17* and *BAZ1B* genes in both European ancestry based analysis and transethnic analysis (Transethnic P-value = 4.41e-16, 1.45e-15, 4.01e-09, respectively), but these associations were not observed in African population in our cohorts (P-value = 0.35, 0.55, 0.36, respectively). Signals at *GCKR* presented significant heterogeneity between ancestries (P-value = 0.017). Also, a genome-wide signal on AT in African American population was found close to *TXNL4B-HP* (Transethnic P-value = 4.37e-26) but was not significant in the European ancestry population (P-value = 0.02). Another novel genome-wide association signal was found close to the *ORM1 gene* in relation with PS (PS free P-value = 1.16e-19, PS total P-value = 1.03e-15), suggesting possible regulatory mechanisms for PS. We also found a potential genome-wide signal (Transethnic p-value =3.73e-24) at *SNX17* for PC, which is in high linkage disequilibrium (LD) with variants in *GCKR*. Confirmation of these potential novel associations require further analysis in independent populations.

**INTRODUCTION**

Proteins C (PC) and S (PS) are both vitamin-K-dependent natural anticoagulants whose association with venous thromboembolism (VTE) and bleeding disorders have been characterized previously [1-5]. PC is activated to produce activated protein C (APC) in the presence of thrombomodulin (TM), which then binds with PS to create the APC-PS complex [6]. The anticoagulation property of the complex mainly relies on inactivating and inhibiting coagulation factors Va and VIIIa, which attenuates blood clotting ability[7]. Approximately 60% of PS circulates in plasma bounded noncovalently with C4b binding protein (C4bBP), which is of biologically active

function. The rest remains unbounded PS free is also in an active form different from the bonded PS[8]. PS participates in the process as a co-factor, stimulating and enhancing the inactivation function of APC to Factor Va and Factor VIIIa by approximately 20-fold[9]. In addition, PS expresses some ability to independently inhibit the thrombin activatable fibrinolysis inhibitor (TAFI), suggesting that protein S may be a regulator in the early process of blood clotting.[10] Free PS has also been associated with risk of VTE for its role in the process of thrombin generation[11]. Antithrombin (AT) is another natural anticoagulant independent of the PC pathway and is also a crucial regulatory molecule in the mammalian circulation systems[12]. By binding with heparin and other heparin like glycosaminoglycans, AT's ability of inhibiting crucial coagulation proteases such as coagulation factor Xa is highly accelerated, thus inhibiting and regulating the clot formation.

Protein Level and Risk of VTE

All three natural anticoagulant proteins have been examined for association with risk of VTE previously. In 1981, Griffin et al found that plasma levels of PC antigen were significantly lower than normal levels in a family with recurring thrombosis history[2]. Similarly, PS deficiency was found to be an inherited thromboembolic disease in a study by Schwarz et al [3]. The association between low levels of AT and familial thromboembolic diseases was first reported in 1965[4], and it was hypothesized that decreased AT levels could lead to higher risk of thrombosis through heparin resistance. Bucciarelli et al [13]examined the incidence of VTE in families with AT, PS and PC deficiencies and found that the incidence of VTE was higher among those who showed deficiency for at least one of these factors compared with healthy populations. A similar association between natural anticoagulants deficiency and VTE has been reported by others [5]. Sanson et al found that the annual incidence of VTE for people with deficiency in PC or AT is relatively higher than the general population: 1.9 % for AT deficiency, 1.0% for PC deficiency and 0.8% for overall participants.

Genetic Variation in the 3 Structural Genes and Levels of Protein and VTE Risk

*PROC*, *PROS1*, and *SERPINC1* encode PC, PS and AT, respectively [14-16]. Variation in these protein-coding genes has been known for decades to be associated with circulating levels of the

respective coded proteins and risk of VTE in family studies. In early 1990s, a series of studies were conducted by the Charter Molecular Genetics Lab[17, 18] to investigate the association between mutations in *PROC* and risk of VTE in thrombophilia families. *PROC* variants causing decreased activity or antigen levels of PC was found in familial deep vein thrombosis and pulmonary embolism patients. In 1999, a case-control study found another variant associated with higher PC concentrations and lower risk of VTE[19]. Similar associations between genetic variation and protein levels as well as VTE risk were found in subsequent candidate gene studies. A recent family-based study found[20] several variants located in *PROC* and *PROS1* genes related with risk of DVT, and a majority of those variants had minor allele frequency (MAF) larger than 0.05. VTE risk associated with variation in the *SERPINC1* gene was previously investigated in a Danish case-cohort study where a variant (rs2227589, G /A, MAF=0.013) was associated with lower AT activity level and higher risk of VTE[21]. However, this association was not highly significant and was not detected by other studies likely due to the low frequency of the risk allele in the general population. Other studies have examined genetic variation in these three natural anticoagulant structural genes and the risk of VTE; most of these studies were family-based studies where rare variation was identified [2-4, 20-22].

GWASs on PC, PS and AT

Genome-wide association study (GWAS) refers to studies testing agnostically genotype-phenotype associations through genome-wide data from participating individuals. Since the first published GWAS in 2005[23], it has been used to investigate genome-wide associations with multiple diseases and traits[24]. There have been some GWASs of AT, PC and PS in the past decades. Several GWASs on different ancestral data from ARIC cohorts were conducted and further identified several loci associated with PC levels[5, 25-27]. Oudot-mellakh et al conducted a GWAS on PC in 951 participants from MARTHA [28], confirming the effect of variants in *PROCR* on PC levels. There has also been two GWAS in the GAIT cohort with genome-wide findings for AT, PC PS levels on *SERPINC1*, *DANJ6* and *PROCR*[29, 30]. But these GWAS have been either limited with small sample sizes or restricted to single ancestries, motivating the need for larger, transethnic studies for further identification of genetic associations for these three phenotypes.

Other Genes Regulating Levels of Natural Anticoagulants

Besides variants located in *PROC*, *PROS1*, and *SERPINC1*, variants within other loci have been found to be associated with levels of these three natural coagulants. For plasma levels of PC antigen, variants as rs1260326 at *GCKR*, rs6120849 at *BAZ1B* and rs17145713 at *EDEM2* were found to be genome-wide significant in the ARIC cohort[26]. These novel variants found were common in the population (MAF respectively: 0.40, 0.22 and 0.19) and were replicated in independent populations. A variant located in an exon of the protein C receptor (*PROCR*) has been associated with changes in PC antigen levels in this analysis. De la Morena-Barrio et al[31] found that three variants (rs713703, rs762057, rs240082) located in the glycosyltransferase coding gene *LARGE* were associated with secretion of AT independently from expression of *SERPINC1*. Among those three variants, rs240082 had a relatively lower MAF (0.026) while the other two were more common (MAF>0.05). However, in their validation analysis, only rs713703 showed statistical significance and this was further confirmed through functional studies of *LARGE*. For PS, Athanasiadis et al found potential signals at *DANJC6* affecting PS levels within the GAIT2 cohorts[29]. Two common intronic variants in *DANJC6* were associated with free PS levels.

In this study, we conduct a GWAS meta-analysis to assess the association between low frequency and common variants (as defined by specific MAF criteria) and these three natural anticoagulants, which serve as intermediate phenotypes for VTE risk. We include participants of African American and of European ancestry from nine cohorts[25, 28, 32-34]. We used TOPMed imputation panel which has guaranteed our capability of characterizing low frequency variants.

## METHODS

Study Design

The setting for the analysis is the Hemostasis Working Group within the Cohorts of Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium[35]. The Working Group has organized and conducted multiple GWAS of hemostasis phenotypes over the past decade[35-52].We

conducted new meta-analyses of existing GWAS summary data on three hemostasis phenotypes (PC, PS, and AT).

## Studies and Participants

We included summary GWAS data from nine studies with one or more of the three phenotypes[25, 28, 32, 33, 53-57]. Table 1 provides information on 9 studies and which phenotypes are available. Briefly, LURIC, MARTHA and RETROVE were hospital-based studies and GAIT was family-based while the remaining seven studies were all population-based studies. Except for family-based GAIT study, which recruited participants aged from <1 to 88, all participants in the other studies were adults aged over 18. Two studies, ARIC and CHS, included African-ancestry participants while the rest included European-ancestry individuals. Due to the small number of African participants, we only used African ancestral data from ARIC. Detailed information of all studies can be found in the Supplementary Table S2.

## Natural Anticoagulant Phenotypes

Phenotype data were previously measured in each study and include plasma coagulation activity or antigen level of PC, PS (total PS and free PS) and AT [25, 28, 32, 33, 53-56]. Measurement of AT were performed as activity levels. Details of phenotype data (means, counts, standard error and type) can be found in Table 1.

## Genotyping and Study-specific Analyses

Participating studies had genotyping performed on various commercial platforms. Details about genotyping platforms can be found in Supplemental Table S2. Each study imputed genotypes to the Trans-Omics for Precision Medicine (TOPMed) imputation[58] panel before association analysis. Study specific quality control was implemented before the analysis. For each phenotype in each study, inverse normal transformation was performed and regressed with adjustment variables such as age, sex, principle components and study-specific variables, and residuals were captured for analysis. Each study then performed ancestry-stratified association analysis using linear regression on the inverse

normal transformed residuals. Women and men were coded as 0,1,2 and 0, 2 for the X chromosome, respectively.

Quality Control and Meta-analysis

We conducted extra quality control on participating cohorts to discover and resolve problems across studies. Briefly, meta-level quality control focuses on problems within studies, including issues with alleles, strand and population stratification as implemented in the EasyQC pipeline[59].

We meta-analyzed study-level GWAS results across 22 autosomal chromosomes plus the X chromosome when available. Activity and antigen measurements for PC, PS were analyzed separately first, and then combined. Details of the meta-analysis can be found in the supplementary methods. All meta-analyses were conducted in METAL[60]. Meta-analyses in PS and AT utilized fixed-effects inverse variants models. Models weighted by sample size were incorporated in the analyses of PC since we combined two different phenotype units in the analysis. Genomic-control coefficients ($\lambda_{GC}$) were checked and all analyses were adjusted for $\lambda_{GC}$ of individual cohorts.

Meta-analysis level quality control

A comparison between two parallel meta-analysis was also conducted in order to uncover and resolve problems happening in the process of meta-analysis. Log files covering the detailed process of meta-analysis and descriptive statistics were used to check consistency in analysis coding.

**RESULTS**

Overall, nine cohorts participated in our analysis and a total of 30,055 ( 27,367 European-ancestry participants and 2,688 African American-ancestry participants) were included in the analyses. Detailed information about genotyping, phenotype measure and cohorts' characteristics has been published elsewhere before [25, 28, 32, 33, 53-57].

Antithrombin Discovery

We included eight cohorts in the meta-analysis of AT for African and European ancestry populations (ARIC, CHRIS, CHS, GABC, GAIT, LURIC, MARTHA and RETROVE). Among these cohorts, ARIC and CHS provided data from both ancestries. 27,783 individuals (25,095 European and 2,688 African American participants) and 33,842,284 variants contributed to the analysis. GABC expressed slightly high level of coefficients' inflation($\lambda_{GC}$=1.04), while all the other studies had $\lambda_{GC}$ estimates smaller than 1.03, indicating that inflation of coefficients was fine. P-values of all variants were plotted by chromosome and region order in Figure 1 (Manhattan plot), quantile to quantile plot of P-values of these variants (Q-Q plot) was presented in Figure 2. In total, 426 SNPs met the genome-wide significant level in transethnic analysis. A total of four loci contained genome-wide significant signals: 1q25.1, 2p23.3, 7q11.3 and 16q22.2. Variants with the smallest P-value in each region are listed in Table 2 along with the meta-analysis P-value, ancestry-specific MAF, beta estimates, closest gene, and whether this association is novel. Forest plots of those genome-wide signals are presented in Supplementary Figure 1-10.

Genome-wide significant signal found at chromosomal position 1q25.1 spanned six genes (*SDCCAG3P2 - ENST00000326105*, *SERPINC1*, *KLHL20*, *ANKRD45*, *TEX50*, and *RABGP1L*). Rs557150901 in *SDCCAG3P2* had the smallest P-value in this locus (transethnic P-value 9.50e-16, MAF EA <0.01, MAF AA <0.01) But this variant was in high LD (D' = 1, $r^2$ = 0.120, Genome1000 European population) with missense variant rs2227624 (A/T) in *SERPINC1* (transethnic P-value: 9.49e-15, MAF EA <0.01, MAF AA missing). We considered the signal at *SDCCAG3P2* to be driven by the known association at *SERPINC1*. However, rs2227624 was only observed in our European ancestry population, and was monomorphic in our African American population. Five other genes within 1,132kb of rs2227624 (*SDCCAG3P2 - ENST00000326105*, *RABGPIL*, *KLHL20*, *ANKRD45*, and *TEX50*) likely harbor variants in LD with unmeasured *SERPINC1* variants.

At chromosome 2, genome-wide signals were found at 2p23.3 and spanned two genes: *SNX17* and *GCKR*. Top SNPs rs4665972 in *SNX17* and rs11127048 in *GCKR* were not significant in our African American population and (P-value respectively: 0.55, 0.35). R*s*11127048 (G/A) had the smallest P-value in regions close to *GCKR* (transethnic P-value = 1.45e-15, MAF EA = 0.45, MAF

AA = 0.29). This SNP was 6kb upstream of *GCKR* and has not been reported to be associated with AT levels previously. rs4665972 (T/C) had the smallest P-value in this region and it is located in *SNX17*(Transethnic P-value = 4.41e-16, MAF EA =0.451, MAF AA = 0.09), and was intronic to *SNX17*. *SNX17* has not been reported to be associated with AT levels previously. These two SNPs were in high linkage disequilibrium with each other (D' = 0.8851, $r^2$ = 0.7802, Genome1000 European population).

Genome-wide associations at chromosome 11 were located at q11.23 in gene *BAZ1B*. rs13244268 (T/C) had the smallest P-value in this region (transethnic P-value=4.01e-9, MAF EA = 0.14, MAF AA =0.04). This association was only observed in European ancestral participants in our analysis (EA P-value = 6.00e-9, AA P-value =0.36). *Rs*13244268 was located in an intron in *BAZ1B*, a gene coding for bromodomain protein[61], and was also found to be related with levels of PC by previous GWAS in the ARIC cohort. [26]

At chromosome 16, genome-wide signals were found at 16q22.2. rs5471 (A/C) had the smallest P-value in this region (transethnic P-value = 4.37e-26, MAF EA = 0.0084, MAF AA =0.1341), which was located in an intronic region of *TXNL4B* and 5 prime UTR variant of *HP*. This association was genome-wide significant in African ancestral data only (EA P-value = 0.02, AA P-value =7.76e-25). In a previous study, rs5471 was found to regulate haptoglobin level in Ghanaian population[62], *TXNL4B* and *HP* were also reported to be related with lipid levels[63, 64].

Genome-wide signals found at chromosomes 1 and 2 were only significant in European ancestry populations and the signal at chromosome 16 was detected only in our African American population. Heterogeneity of direction and magnitude for those SNPs were not significant across participating cohorts. Variants at *GCKR* presented a significant heterogeneity in direction across European and African American ancestries (heterogeneity P-value = 0.017, positive effect relative to AT levels in European), while rest of the signals were not found significantly different by ancestry. In our genome-wide findings in AT analysis, CHS was the most influential cohorts. Except for rs11127048 in *GCKR*, CHS has expressed weight larger than 40% in the analysis.

Protein C Discovery

Eight cohorts (ARIC, CHS, GABC, GAIT, LURIC, MARTHA, TSS, RETROVE) were included in the transethnic analysis of PC. A total of 16,597 individuals and 33,342,224 variants contributed to the analysis. Similar to AT analysis, only GABC had $\lambda_{GC}$ estimate larger than 1.04, so there were no significant signs for inflation of test statistics. A Manhattan plot of these results is shown Figure 3 and the Q-Q plot is presented in Figure 4. We found 2,513 SNPs meeting the genome-wide significant level, identifying five regions related to level change of PC at 1p13.3, 2q14.3, 2p23.3, 7q11.23 and 20q11.22. Coefficients, P-values, ancestry stratified MAF and closest genes of SNPs with strongest association in each region are listed in Table 2.

The genome-wide association signal at 1p13.3 and was located in the 3'UTR of *CELSR2*; rs12740374 (T/G) had the smallest P-value in the region (transethnic P-value = 7.15e -14, MAF EA = 0.26, MAF AA = 0.22). This association was only found in the African ancestral data (P-value = 4.67e-11); the variant was similarly common in the European-ancestry participants but the association did not achieve genome-wide significance (P-value =2.1e-04). This SNP was reported to be related with levels of low-density lipoprotein cholesterol (LDL-C)[65] and lower PC plasma levels and may be further related to cardiovascular disease[25].

Two loci reaching genome-wide significance were found at chromosome 2 at 2q13.3 and 2p23.3. The 2q13.3 locus included *PROC* and the common variant rs1799809 had the smallest P-value in this region (transethnic P-value = 1.40e-87, MAF EA = 0.43, MAF AA = 0.32). This variant showed significant association in both ancestries. *Rs*1799809 was reported to be associated with PC levels in previous GWAS on ARIC cohort[26] At 2p23.3, the smallest p-value was found at rs4665972 (T/C) (transethnic P-value =3.73e-24, MAF EA = 0.42, MAF AA = 0.09) a common variant in EA participant intronic to *SNX17*. In the African-ancestry participants, we did not detect an association (P-value = 0.14). Note that this intron variant of *SNX17* was not reported for the association with PC before. There was no direct evidence of the association between *SNX17* variants and PC levels, but previous study had identified its association with blood lipid levels[66]. As described in AT findings, rs4665972 is in high LD with upstream variant rs11127048 of *GCKR*. In previous analyses on ARIC cohort, the association between *GCKR* and PC levels were found significant[25, 26].

Genome-wide association found at chromosomal position 7q11.23 was close to *BAZ1B;* the

rs34594435 (C/T) variant had the smallest P-value in this region (transethnic P-value = 4.72e-10, MAF EA = 0.19, MAF AA = 0.08). This variant was 40kb downstream of BAZ1B, and was in high LD with rs6976930 which is an intron of *BAZ1B*; as such, we consider this signal affected by the known association of variants in *BAZ1B*.This association was only found significant in European-ancestry participants (P-value = 2.05e-10); no association was found in African American-ancestry individuals (P-value = 0.89). In European ancestry, rs6976930 also had a significant signal (P-value = 1.04e-09)

At chromosome 20, genome-wide association was at q11.22, the location of the protein C receptor structural gene, *PROCR*. *Rs*11907011 had the smallest P-value in this region (transethnic P-value = 2.57e-390, MAF EA =0.09, MAF AA = 0.06). Analyses of both ancestries showed significant signals (EA P-value =1.40e-33, AA P-value =1.51e-358). *Rs*11907011 was intronic to *PROCR* and was reported in previous meta-analysis to be related with PC levels[25-27].

No significant heterogeneity was found within associations described above in direction and magnitude either across different cohorts or ancestries. However, some variants located in *PROCR* and *PROC* exhibited significant heterogeneity across different ancestries, this probably due to smaller sample size of African population.

Protein S Discovery

For PS, 6 cohorts (CHS, GABC, GAIT, LURIC, RETROVE and TSS) were included in the analysis of PS total, and 5 cohorts (CHS, MARTHA, LURIC RETROVE, GAIT) were included for the analysis of PS free. We included 6,257 and 4,006 European ancestral individuals in the analyses of PS total and free, respectively. 11,643,066 variants were investigated in the analysis of PS free and 13,040,841 variants were investigated in the analysis of PS total. Inflation of test statistics was acceptable: GAIT2 was the only cohort having $\lambda_{GC}$ value larger than 1.03 for both sub-phenotype of PS; while rest cohort showed smaller $\lambda_{GC}$ estimates. P-values of variants investigated in these analyses were plotted in Manhattan plots by chromosome and region order were listed in Figure 5 and 6 for PS free and total, respectively. Q-Q plots are presented by Figure 7 and 8. Sixteen and four

variants exceeded the genome-wide significant level for the analyses of PS free and PS total, respectively. Forest plots of those genome-wide signals are presented in Supplementary Figure 11-14.

We identified genome-wide significant variants for PS free at 3q11.1 and 9q32, and genome-wide significant variants for PS total at 9q32. We listed top variants in these two regions along with their summarized P-value, effect, and possible related genes. Genome-wide association identified at chromosomal position 3q11.1 was located at *PROS1*. rs121918472 (A/C) has the smallest P-value (PS free P-value = 2.30e-11, MAF = 0.02) and has been reported to be associated with PS deficiency, PS Heerlen, and risk of thrombosis[67, 68]. In our analysis, this variant was positively associated with PS free level, significance was not observed in PS total, and this is probably related with low MAF in European population (<0.05) of this variant. We also found a missense variant rs566931451(C/A) located at chromosomal position 3q11.2 in the analysis of PS free (P-value = 2.67e-12, PS free MAF = 0.01). This variant is located between *EPHA3* and *PROS1*. Genome-wide significant association found at chromosomal position 9q32 was located 2kb upstream of *ORM1*, and rs150611042 (C/A) was the SNP with smallest P-value (PS free P-value =1.16e-19, PS total P-value = 1.03e-15, PS free MAF = 0.09, PS total MAF = 0.02 ). rs150611042 has not been related with the levels change of PS before. *ORM1* has been related with thrombosis generation in previous meta-analysis[69].

Heterogeneity of the direction and magnitude of these 4 SNPs were not significant across different cohorts. There was no significant sign of imbalanced contribution from different cohorts in the analysis of PS free and PS total.

**DISCUSSION**

This study is conducted within CHARGE hemostasis working group. We collected summary data from 9 GWASs and meta-analyzed 3 phenotypes (AT, PC, PS) through these data. Transethnic meta-analyses were utilized in AT and PC analyses. In general, we found novel genome-wide signals of AT level at *GCKR*, *SNX17*, *BAZ1B* and *TXNL4B-HP* region. (*GCKR* and *SNX17* may derive from same signal at 2p23.3) *GCKR* and *BAZ1B* were identified to be related with PC levels in previous analysis and/or our sub-analysis of PC. In PC analysis, except for genome-wide signal at rs4665972 (*SNX17)*, which is in high LD with variant rs11127048 in *GCKR*, our results were close to previous analysis on

ARIC cohort. We also presented novel genome-wide association found at *ORM1* with PS total and PS free, which located at 9q32 region.

Genetic Findings for Antithrombin

At chromosome 1, 6 genes were identified through genome-wide association signals. These signals showed significance only in European specific and transethnic analysis. rs557150901 had the smallest P-value around this region but was in high LD with rs2227624 which is a missense variant in *SERPINC1*. *SERPINC1* has been recognized as the coding gene for antithrombin for decades[4], and variants at *SERPINC1* were found to be related with antithrombin levels[70]. Moreover, rs2227624 was in relation with risk of VTE and DVT[71, 72]We hypothesize that the signal at rs557150901 was caused by its high LD with signals at *SERPINC1*. The other four signals in this region were either in linkage equilibrium with rs2227624 or in weak correlation with it, and their genome distance with rs2227624 are smaller than the distance to rs557150901 (1,132kb). There was no report of these genes in relation with AT levels before, and we conceive it plausible that these signals derived from rare variants located at 1q25.1, especially at known effective gene *SERPINC1*. Further analysis conditional on *SERPINC1* variant is needed to properly illustrate the actual effects of these signals.

Genome-wide association at 2p23.3 yielded findings of two genes in this region: *SNX17* and *GCKR*. Intron variant of *SNX17* rs4665972 was in high LD with rs11127048 located near *GCKR*. Note that these 2 associations were not significant in African ancestry. *SNX17* codes for sorting nexin 17 (Snx17). No previous evidence was found on its effect to AT level, but recent study have introduced a discovery on how Snx17 could affect cardiovascular disease[73]. Snx17 was found to be a regulator for LDL receptor such as LDL-receptor protein 1 (LPR-1)[74]. LPR-1 was found to be a influential factor in regulating factor V. There was also evidence about the interaction between Snx17 and P-selectin[75], and the latter is involved in thrombin generation process. *GCKR* variants was highly pleotropic, its association with Factor VII[76, 77], Factor XI[45], and c-reactive protein[78, 79] has been found in previous meta-analyses and GWASs. We also observed its association with PC in our meta-analyses. Some theories have been raised for possible mechanisms of its association with thrombosis related phenotypes, mainly depending on its regulation effect on glycosylated protein in liver[80]. Evidence on association between

other hemostasis related proteins or factors and *GCKR* variants could be a strong support for its determinant role in the genome-wide association found at 2p23.3.

At chromosomal position 7q11.23, top signal of rs13244268 was intronic to *BAZ1B*, which encodes the bromodomain protein, and its absence could be related with Williams-Beuren syndrome[81]. This SNP was found significant in bivariate and univariate GWAS of c-reactive protein and high-density lipoprotein (HDL) in the study of Ligthart et al[82]. Along with previous findings of association between *BAZ1B* and PC level and our similar findings in the analysis of PC, it was plausible to assume the existence of pleiotropic effects of *BAZ1B* variants within thrombosis and hemostasis traits. Note that this association was not observed in our African cohort.

The last genome-wide significant findings of AT level were located at chromosomal position 16q22.2. Top signal rs5471 was intron to *TXNL4B*and 5' UTR of *HP*. *HP* is the coding gene for pre-protein of haptoglobin (Hp). In the study of Soejima et al, association between rs5471 and serum Hp level was confirmed in Ghanaian population[62]. Hp serves as the binding protein of hemoglobin, further affecting the release of hemoglobin from red blood cells[83] Its phenotype Hp2-2 was investigated to be potential regulator for inflammation and reverse cholesterol transportation[84, 85]. Moreover, in the study of Vormittag et al[86], Hp2-2 was examined to have higher prevalence in VTE patients. The binding process of Hp to hemoglobin could prevent the generation of oxidized LDL[87] from the activation pathway of free hemoglobin. Oxidized LDL could increase prothrombinase activity in vivo[88], and affect the level of Factor VII, IX and XI[89]. *TXNL4B* codes for the thioredoxin-like 4B protein. Variants at *TXNL4B* was also found to be related with haptoglobin levels[90] and lipid levels[63, 91]. Thus, it is also plausible to be the determinant gene in this region. In general, these evidences suggested potential role of Hp in the inflammation-induced thrombosis, and further build plausible pleiotropic effect for the association found in our analysis.

Genetic Findings for Protein C

A total of 5 genome-wide associations were found significant in the analysis of PC. Only 2 associations at *PROC* and *PROCR|MMP4-AS21-EDEM2* showed significant signal in both ancestries, and were known associations in a previous study from ARIC European and African ancestries[25-27].

The signal at chromosome position 20q11.22 may represent the effect from both *PROCR* and *EDEM2*. In ARIC, signals at *EDEM2* were found to be related with the level change of PC regardless of variants at *PROCR*. And evidence from conditional analysis along with some biological theory further indicated the possible role of *EDEM2* in changes of PC plasma level. However, more evidence is still required for this gene to be a confirmed candidate gene regulating PC plasma level.

Genome-wide association signal rs12740374 found at *CELSR2* was only significant in African ancestry and transethnic analysis. *CELSR2* is primarily involved in the coding of non-classic type cadherins.Chromosomal position 1p13.3 was related with LDL-C levels and coronary disease[65]. According to the meta-analysis of ARIC cohort conducted in 2017, rs12740374 which was defined to be located at *CELSR2-PSRC1-SORT1* region, showed significance in both ancestries[25]. This study concluded that the influence on PC levels from *CELSR2-PSRC1-SORT1* region could possibly derive from its regulation toward LDL-C. And evidence of LDL-C being the upstream determinant of PC had been raised from conditional analysis and mendelian randomization[25]. However, in our European cohorts, this association was not genome-wide significant, and the frequency of this SNP was still around 0.2 in European population. There is a possibility that some hospital-based cohorts (MARTHA, LURIC, RETROVE) had some selection criterion on coronary disease history or diseases associated with LDL-C levels. But the actual role of *CELSR2* to the PC plasma level remains unclear in European population.

Genome-wide associations at chromosomal position 2p23.3 were within *SNX17*, and the top SNP rs4665972 was in high LD with SNPs in *GCKR*. Similar to previous analyses on ARIC, this association was not observed in African ancestry in our analysis [25, 26]. It was still unclear of the actual role of these two genes. *GCKR* was assumed to have some pleiotropy effect on PC through its regulating effect of liver glucose in the GWAS on ARIC[25, 26], and other thrombosis related phenotypes[45, 76-80]. There is no direct evidence of *SNX17* being related with PC plasma level, but interesting findings about how its product protein snx17 could affect the generation of thrombosis were reported before. It was found to be related with level of LPR-1 in previous study [74]. Gertz and Bouchard[92] found that LPR-1 was involved in the binding process of Factor V and its receptor, which could further influence thrombin generation. In conclusion, more evidence was found in *GCKR* variants but still could not support further assumption around these two genes about their role in the association with PC plasma level.

We also found signals at chromosomal position 7q11.23 within *BAZ1B* associated with PC levels.

This association was also been investigated by previous analysis on ARIC cohorts[25, 26]. There has been some assumption of pleiotropic effects within the GWASs on ARIC regarding the finding of its association with triglyceride levels[26, 93]. Nevertheless, association of variants in *BAZ1B* and AT levels were observed in our analysis of European specified and transethnic analysis. This gives a rise of possible pleiotropic effect of variants in *BAZ1B.* A large proportion of our PC data was from ARIC, and some overlap of our genome-wide findings on PC is expected.

Genetic Findings of Protein S

The most significant association found at PS was located at *ORM1*, which showed genome-wide significant signals for both PS free and PS total. There were also 2 signals from *PROS1* and an intergenic SNP between *PROS1* and *EPHA3*. The *PROS1* is the structural gene coding for PS and has been previously studied in relation with the levels of PS. Our findings further support this discovery. Signal at rs566931451 was not in high LD with rs121918472 (D' =0.4985, $r^2$ =0.1655, Genome1000 European population), indicating that *EPHA3* may independently affect the level of PS. But confirmation of this findings requires further analysis with adjustment to genome-wide signals at *PROS1*.

Genome-wide association found at *ORM1* represent novel findings. *ORM1* is responsible for encoding acute phase plasma protein orosomucoid (ORM, also known as $\alpha_1$-acid-glycoprotein, AGP)[94], which is increased with acute inflammation. The research conducted by M.V Osikov et. al found that ORM could affect the hyper coagulation pathway and further affect the level of AT and platelet aggregation during the development of septic peritonitis[95]. In their research, ORM was examined to produce a positive effect on all pathogenetic components during the development of thromboembolic disorders. Also, previous genetic results suggested that ORM1 was associated with thrombin generation potential[69], and the discovery was further confirmed with in vitro experiments. Although we still lack direct evidence on whether ORM could affect the level of PS with in vivo or animal research, our results from PS free and total analyses strongly suggest an involvement of ORM1 in the regulating pathway of PS.

Strengths and Limitations

We hold our most important strength as relatively larger sample size, which is approximately 30,000 individuals with European or African ancestry. We are expected to be more powered compared with previous studies (e.g. GWASs on single cohort) and could be able to detect variants with lower MAF by this large sample size. There were several genome-wide association signals found at rare variants in our analyses. Take rare variant rs2227624 at *SERPINC1* found in AT analysis as an example. With 27,783 participants in the analysis, the statistical power to detect such variant with extremely low MAF (0.063) is estimated as 0.9372. While under the same setting of main effect (beta-estimate) and phenotype level (mean and standard deviation), the power for an analysis with 9000 (e.g. ARIC) sample would only be 0.0698.The novel imputation panel we used for all cohorts included in the analyses is also one of our strengths, which has better imputation quality at variants of lower frequencies compared with previous Genome1000 or Haplotype Reference Consortium panel. Also, non-European population could be represented better in TOPMed panel.

The major limitation of our study was introduced by the absence of replication analysis. Some of our findings could be false positive since we used all available data for the primary analysis and the power of detecting false positives was limited. We provided forest plots of AT and PS genome-wide findings trying to figure out the effect brough by single cohorts in our analysis. And there was a significant sign of CHS influencing findings in AT analysis. This should be further investigated by conditional analysis or other sensitive analysis. Also, to date, mostly rare familial variants have been found in *PROC, PROS1,* and *SERPINC1* genes and we will be underpowered to detected rare variants, which will lead to false negative findings. There was also some limitation about the absence of conditional analysis that we lack effective ways to investigate the existence of multiple independent signals within one locus.

In general, we used summary statistics from nine cohorts to conduct transethnic meta-analyses on AT, PC, and independent meta-analyses on PS free and total. We report potential novel associations found at *BAZ1B, GCKR, TXNL4B-HP* on AT, *ORM1* on PS. Conditional analyses are needed for

further confirmation of these findings. Our findings of known association as *SERPINC1* on AT, *CELSR2*, *PROC*, *BAZ1B*, *GCKR, PROCR* and *EDEM2* on PC, and *PROS1* on PS further support previous their role in the regulating pathways of these phenotypes. Moreover, findings on *HP* and *ORM1* suggest possible novel regulatory mechanisms for AT and PS.

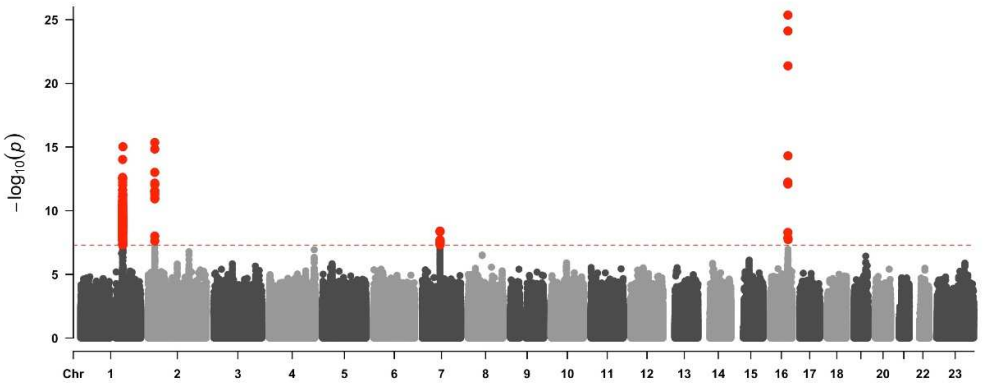



**Figure 1. Manhattan plot of transethnic meta-analysis of Antithrombin**

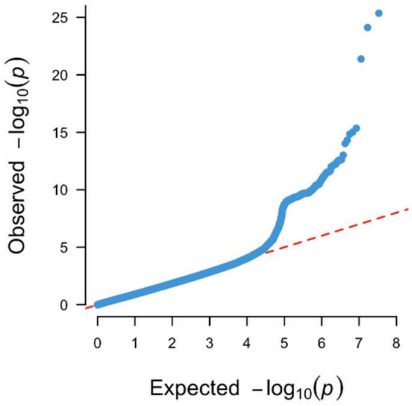



**Figure 2. Q-Q plot of transethnic meta-analysis of antithrombin**

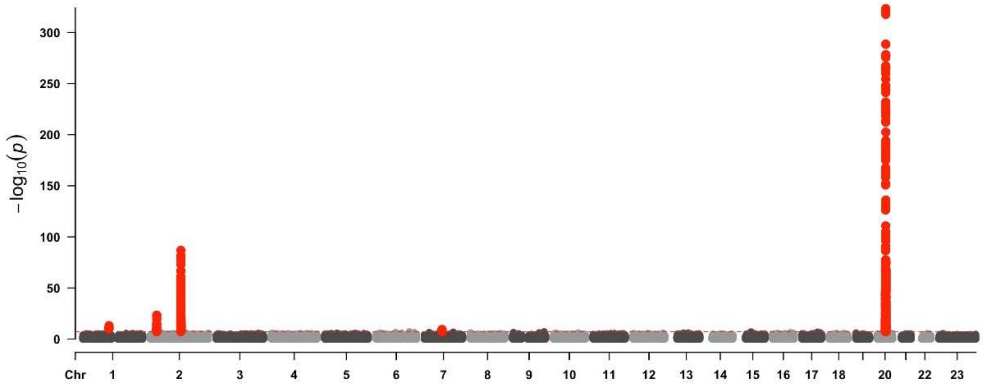



**Figure 3. Manhattan plot of transethnic meta-analysis of Protein C**

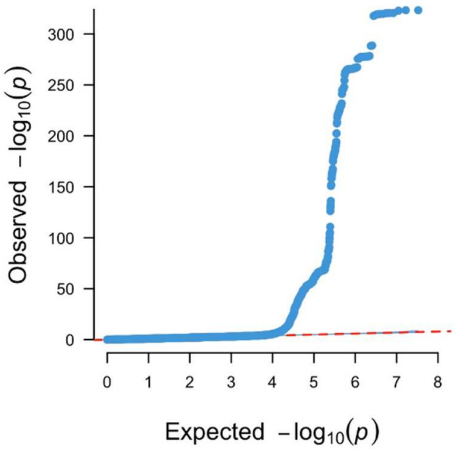



**Figure 4. Q-Q plot of transethnic meta-analysis of Protein C**

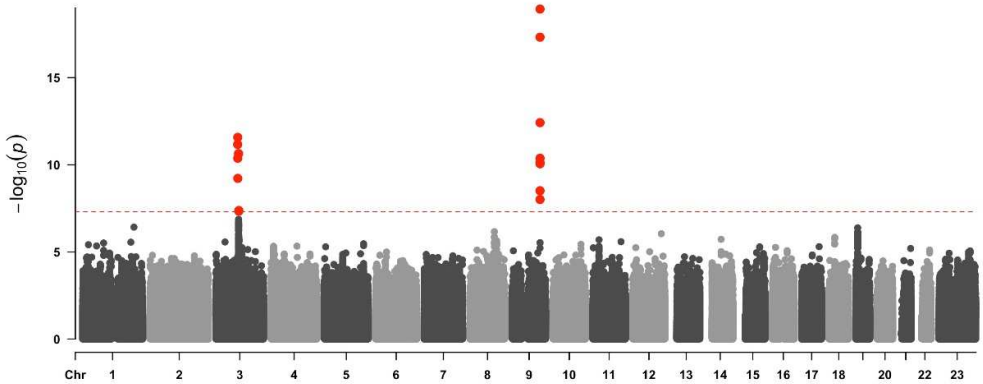



**Figure 5. Manhattan plot of meta-analysis of Protein S Free**

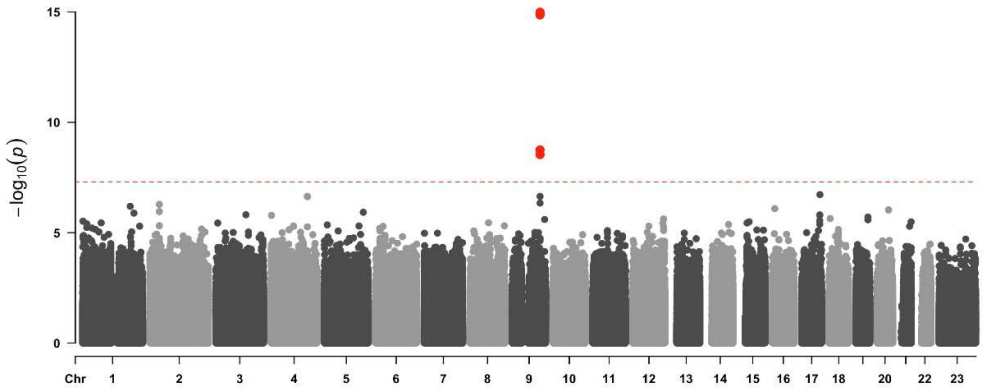



**Figure 6. Manhattan plot of meta-analysis of Protein S Total**

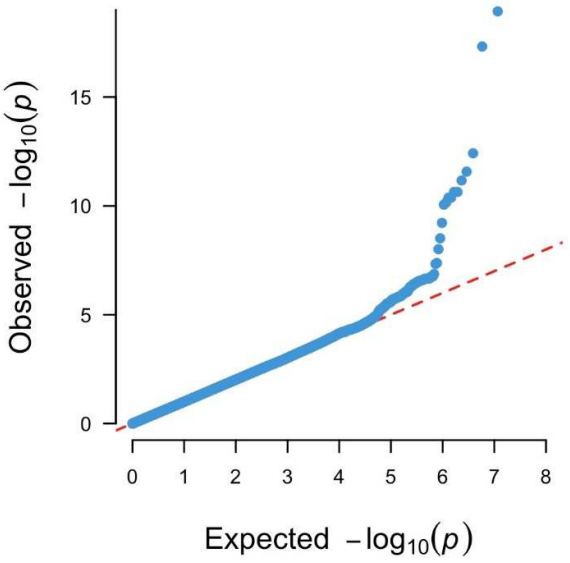



**Figure 6. Manhattan plot of meta-analysis of Protein S Free**

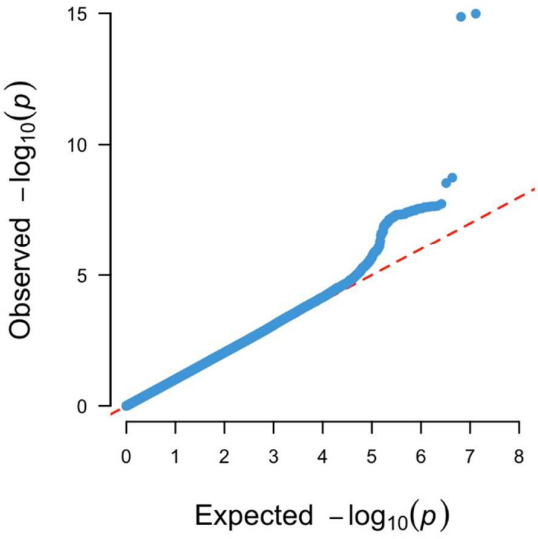



**Figure 6. Q-Q plot of meta-analysis of Protein S Total**

## Table 1. Phenotype information of selected cohorts

| Measure | Ancestry | Antithrombin | | | Protein C | | | Protein S Free | | | Protein |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | Mean (SD) | Unit | N | Mean (SD) | Unit | N | Mean (SD) | Unit | N | Mean (SI |
| Activity | EA | 9179 | 110.2(20.9) | % | / | / | / | / | / | / | / | / |
| Antigen | EA | / | / | / | 9180 | 3.2(0.6) | Ug/mL | / | / | / | / | / |
| Activity | AA | 2688 | 115.3(23.1) | % | / | / | / | / | / | / | / | / |
| Antigen | AA | / | / | / | 2688 | 3.1(0.6) | Ug/mL | / | / | / | / | / |
| Activity | EA | 9012 | 102.5 (11.5) | % | / | / | / | / | / | / | / | / |
| Activity | EA | 681 | 109.4 (14.1) | % | / | / | / | 288 | 101.3 (18.9) | % | / | / |
| Antigen | EA | / | / | / | 683 | 4.2 (1.6) | ug/mL | / | / | / | 290 | 99.7 (13. |
| Activity | EA | 932 | 36.32(6.5) | IU/mL | 919 | 118.7(20.21) | IU/dL | / | / | / | 932 | 83.49(26. |
| Activity | EA | 890 | 112.1(13.5) | IU/mL*100 | 890 | 115.8(26.6) | IU/mL*100 | / | / | / | 922 | 102.7(24 |
| Antigen | EA | / | / | / | / | / | / | 922 | 94.7 (23.0) | IU/mL*100 | 911 | 96.5(20 |
| Activity | EA | 2705 | 97.1 (13.4) | % | 1041 | 110.5 (22.6) | % | 1049 | 101 (27.4) | % | 1040 | 118.9 (31 |
| Activity | EA | 897 | 102.9 (11.6) | % | 949 | 111.6 (24.7) | % | / | / | / | / | / |
| Activity | EA | / | / | / | 2136 | 128.33 (19.55) | IU/dL | / | / | / | 2272 | 84.44 (27 |
| Activity | EA | 399 | 106.9(13.4) | IU/mL*100 | 399 | 127.498 (24.67) | IU/mL*100 | / | / | / | 398 | 109.944 (2 |
| Antigen | EA | / | / | / | / | / | / | 398 | 100.6 (21.3) | IU/mL*100 | / | / |
| Activity | EA | 400 | 107.81(10.8) | IU/mL*100 | 400 | 120.49 (23.855) | IU/mL*100 | / | / | / | 400 | 108.3(21. |
| Antigen | EA | / | / | / | / | / | / | 400 | 99.6(21.3) | IU/mL*100 | / | / |

**Table 2. Main association found in antithrombin meta-analyses**

| Region | Rs ID | Variant | MAF EA | MAF AA | Trans-ethnic beta (SE) | Trans-ethnic P-value | EA p-value | AA p-value | Closest Gene(s) |
|---|---|---|---|---|---|---|---|---|---|
| 1q25.1 | rs61828917 | C/T | 0.1262 | 0.0222 | 1.14(0.18) | 6.33e-10 | 7.86e-10 | 0.5151 | *TEX50* |
| 1q25.1 | rs28825469 | T/G | 0.1695 | 0.2758 | -1.00(0.16) | 3.91e-10 | 1.27e-09 | 0.1262 | *ANKRD45* |
| 1q25.1 | rs61826815 | T/A | 0.1273 | 0.0214 | 1.15(0.18) | 4.26e-10 | 5.87e-10 | 0.381 | *KLHL20* |
| 1q25.1 | rs2227624 | A/T | 0.0063 | N/A | 7.52(0.97) | 9.49e-15 | 9.49e-15 | N/A | *SERPINC1* |
| 1q25.1 | rs190938056 | C/T | 0.0113 | 0.0018 | -4.28(0.65) | 5.25e-11 | 7.39e-11 | 0.374 | *RABGAP1L* |
| 1q25.1 | rs557150901 | A/G | 0.0073 | 0.0014 | -6.89(0.85) | 9.50e-16 | 1.41e-15 | 0.37 | *SDCCAG3P2 - ENST0000* |
| 2p23.3 | rs4665972 | T/C | 0.4471 | 0.0858 | 1.00(0.12) | 4.41e-16 | 5.01e-16 | 0.5463 | *SNX-17\|GCKR* |
| 2p23.3 | rs11127048 | G/A | 0.4465 | 0.2902 | -0.97(0.12) | 1.45e-15 | 1.26e-16 | 0.35 | *GCKR* |
| 7q11.23 | rs13244268 | T/C | 0.1400 | 0.0432 | 1.14(0.19) | 4.01e-9 | 6.00e-09 | 0.36 | *BAZ1B* |
| 16q22.2 | rs5471 | A/C | 0.0084 | 0.1341 | 9.75(0.92) | 4.37e-26 | 0.02 | 7.76e-25 | *TXNL4B\|HP* |

**Table 3. Main association found in Protein C meta-analyses**

| n | Region | Rs ID | Variant | MAF EA | MAF AA | Trans-ethnic P-value | EA p-value | AA p-value | Closest Gene(s) |
|---|---|---|---|---|---|---|---|---|---|
| 968 | 1p13.3 | rs12740374 | T/G | 0.2585 | 0.2209 | 7.15e-14 | 2.1e-04 | 4.67e-11 | *CELSR2* |
| 299 | 2q14.3 | rs1799809 | A/G | 0.4306 | 0.3198 | 1.40e-87 | 1.07e-80 | 1.14e-08 | *PROC* |
| 30 | 2p23.3 | rs4665972 | T/C | 0.4165 | 0.0858 | 3.73e-24 | 1.12e-23 | 0.14 | *SNX17\|GCKR* |
| 19 | 7q11.23 | rs34594435 | C/T | 0.1892 | 0.0765 | 4.72e-10 | 2.05e-10 | 0.8945 | *BAZ1B* |
| 67 | 20q11.22 | rs11907011 | T/C | 0.0938 | 0.0635 | 2.57e-390 | 1.40e-33 | 1.51e-358 | *PROCR\|MMP24-AS21-EDEM2* |

**Table 4. Main association found in Protein S meta-analyses**

| Region | Rs ID | Variant | Related Phenotype | MAF EA | Beta (SE) | EA p-value | Closest Gene(s) | Status |
|---|---|---|---|---|---|---|---|---|
| 3q11.1 | rs566931451 | C/A | Protein S free | 0.0102 | 29.08(4.16) | 2.67e-12 | *EPHA3-PROS1* | Known |
| 3q11.1 | rs121918472 | A/C | Protein S free | 0.0159 | 31.17(4.66) | 2.30e-11 | *PROS1* | Known |
| 9q32 | rs150611042 | C/A | Protein S free | 0.0883 | -8.12(0.89) | 1.16e-19 | *ORM1* | Novel |
| | | | Protein S total | 0.0160 | -6.14(0.77) | 1.03e-15 | | |

## Supplemental Table S1

| Ancestry | Phenotype | Study design | Population | District/Area | Time | Genotype measurement |
|---|---|---|---|---|---|---|
| European/African | AT, , PC ant | Population-based cohort study | Men and Women aged 45-65 years in the selected communities | North Carolina, Mississippi, Minnesota, Maryland | 1987-2017 | Affymetrix SNP array 6.0 |
| European | AT | Population-based cohort study | Adults aged in middle and Upper Vinschgau/Val Venosta area of South Tyrol, Italy | Middle and upper Vinschgau/Val Venosta area of South Tyrol, Italy | 2011- | Illumina Human Omni Express Exome Bead Chip. |
| European/African | AT, PC ant, PST act, PSF act | Population-based cohort study | Adults aged over 65 years in the selected communities | North Carolina, California, Maryland, Pennsylvania | 1989-2010 | Illumina 370CNV BeadChip/HumanOmnil-Quad_v1 Bead Chip |
| European | AT, PC act, PST act | Population-based cohort study | Healthy sibilings aged between 14 and 35 from University of Michigan | Michigan | 2006-2009 | Illumina HumanOmni1-Quad_v1-0_B array. |
| European | AT, PC act, PSF ant, PST act. PST ant | Family-based study | 397 individuals from 21 families selected from a proband with VTE and pedigree size | Barcelona, Spain | 1999-2000 | HumanOmniExpressExome-8v1.2 (324 individuals and coverage 964,193 variants) and HumanCoreExome-12v1.1 (610 individuals and coverage 542,585 variants) |
| European | AT, PC act, PSF act, PST act | Hospital-based cohort study | Patients diagnosed with coronary angiography in a care center | Southwestern Germany | 1997-2000 | Affymetrix SNP array 6.0 |
| European | AT, PC act | Hospital-based cohort study | People with documented history VT and free of AT, PS, PC deficiency or other coagulation related status | La Timone hospital (Marseille, France) | 1994-2012 | Illumina Human610-Quad Beadchip/Illumina Human660W-Quad Beadchip |
| European | PC act, PST act | Population-based cohort study | Irish individuals aged 18 to 28 years attending University of Dulbin | Trinity, Ireland | 2003-2004 | he Illumina HumanOmni1-Quad_v1-0_B array |
| European | AT, PC act, PC ant, PSF ant, PST act | Hospital-based case-control study | VTE patients and controls from the hospital | Barcelona**,** Spain | | Illumina Infinium Global Screening Array-24 v2.0 |

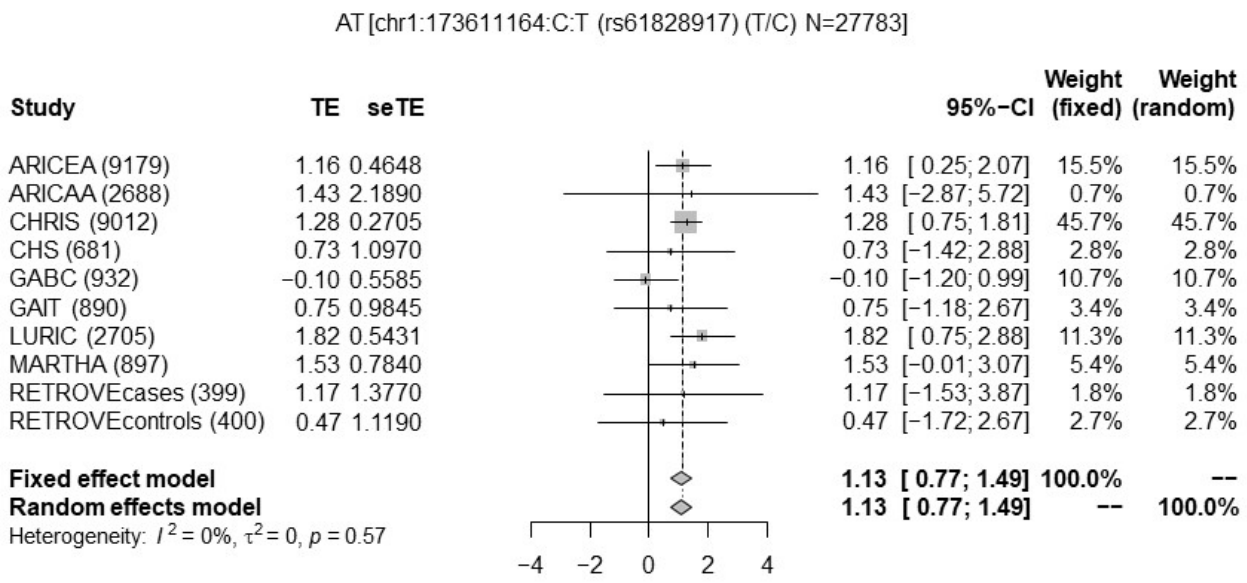
Supplemental Materials:

Meta-level quality control and meta-analysis:

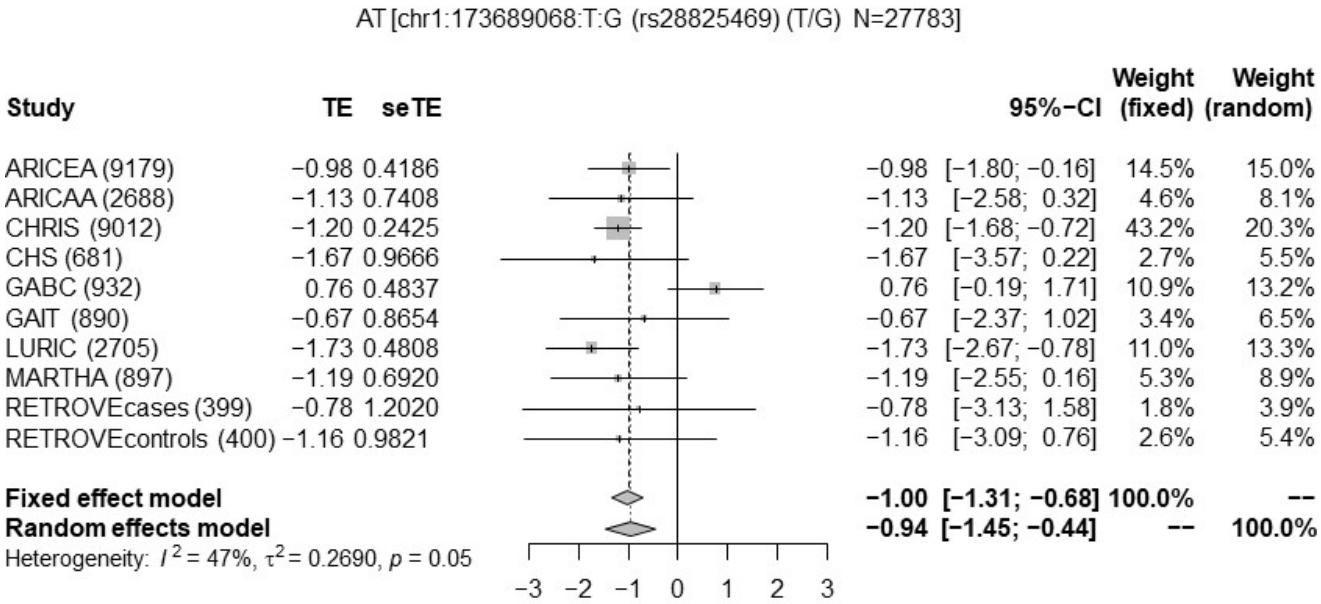
For meta-level statistics, we proposed to provide SE-N (standard error against sample size) plots for different meta-analyses, and deviation from general trend (usually a straight line) could reveal inconsistence of phenotype transformations between studies as well as problems with imputation, effective sample size and issues could affect MAFs across studies.. The P-Z (P-value against Z-value) plot would be utilized to investigate study specified computation problems such as beta, standard errors, and P-value. We also generated effective allele frequency (EAF) plots uncovering problematic issues of strand and allele coding, which is practical and informative for most SNPs. Subsequently, the Lambda-N plot reveals problems with inflation of P-values triggered by problems of stratification. Also, problematic SNPs would be filtered out during the process of quality control and output of EasyQC is cleaned dataset which can be meta-analyzed directly.

Meta-analysis:

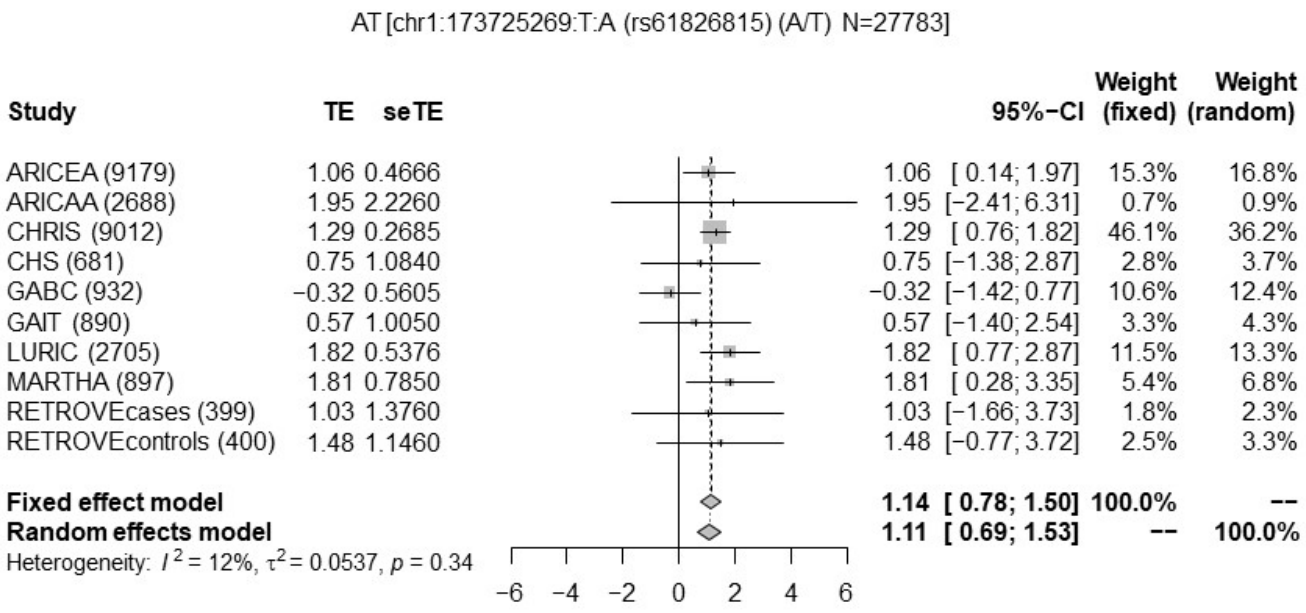
Inverse variance weighted fixed-effect meta-analysis model was utilized to meta-analyze summary data from selected studies for each phenotype. Summarized coefficients and P-values for each phenotype was presented as the result of analyses on AT an PS. Genome-wide association significant level was set as $5\times10^{-8}$ initially. Primary focus in our plan was transethnic analyses across African and European ancestral population. Inclusion-exclusion criteria of the variants were determined before the meta-analysis was completed. Variants with imputation quality smaller than 0.3 were removed before the association analysis and subsequently we also excluded variants that had missing value of coefficients, frequency, and P-value. Also, values would be checked and removed from subsequent analysis. For analyses combining PC activity and PC antigen level, P-values would be utilized to indicate the association instead of coefficients.



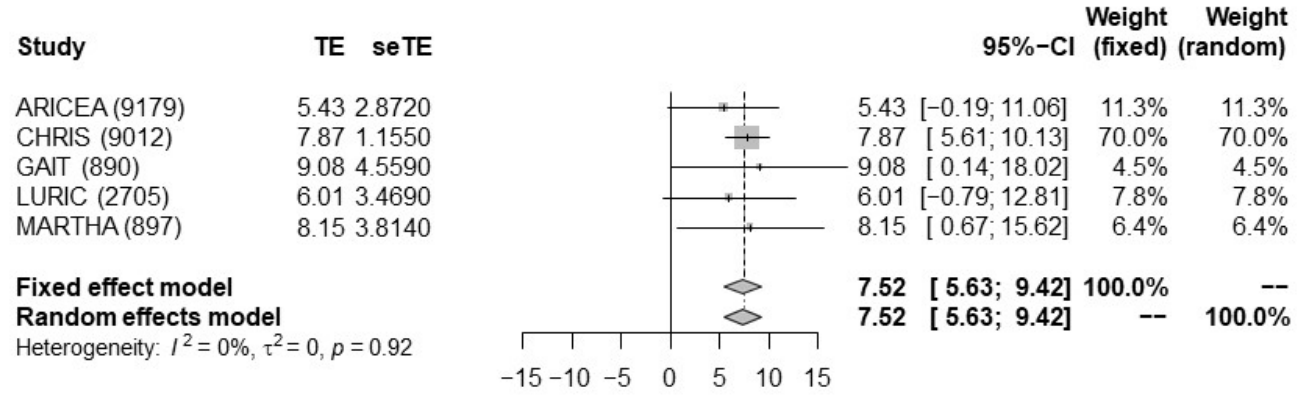

**Supplementary Figure 1. Forest plot for genome-wide signal at rs61828917 in AT analysis**



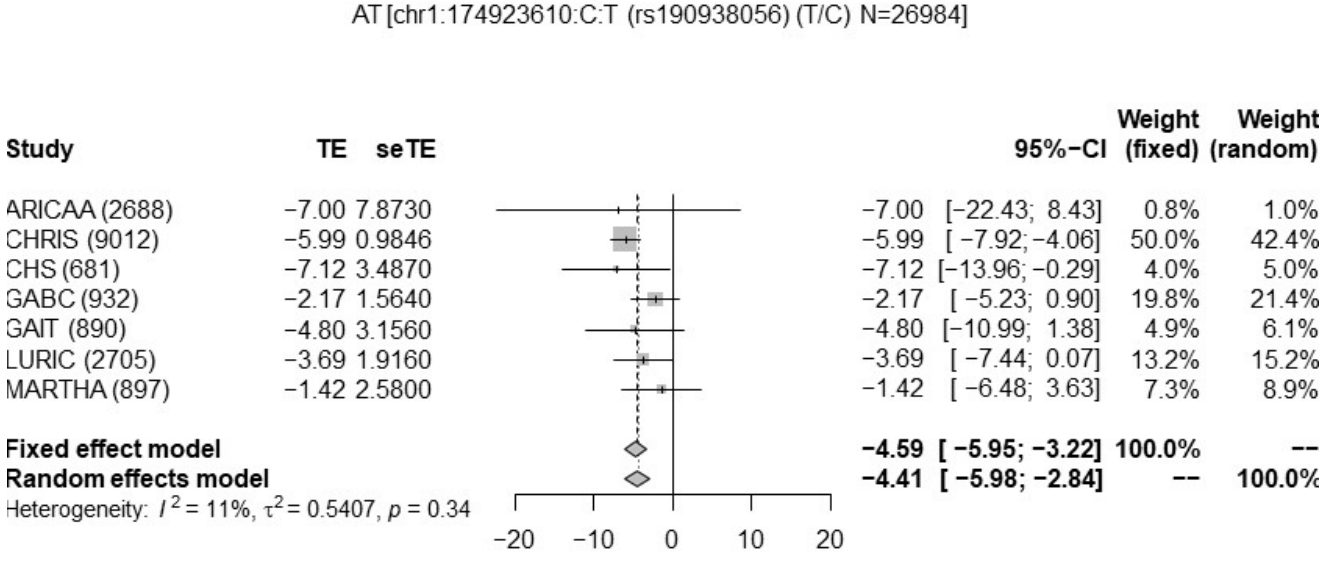

**Supplementary Figure 2. Forest plot for genome-wide signal at rs28825469 in AT analysis**



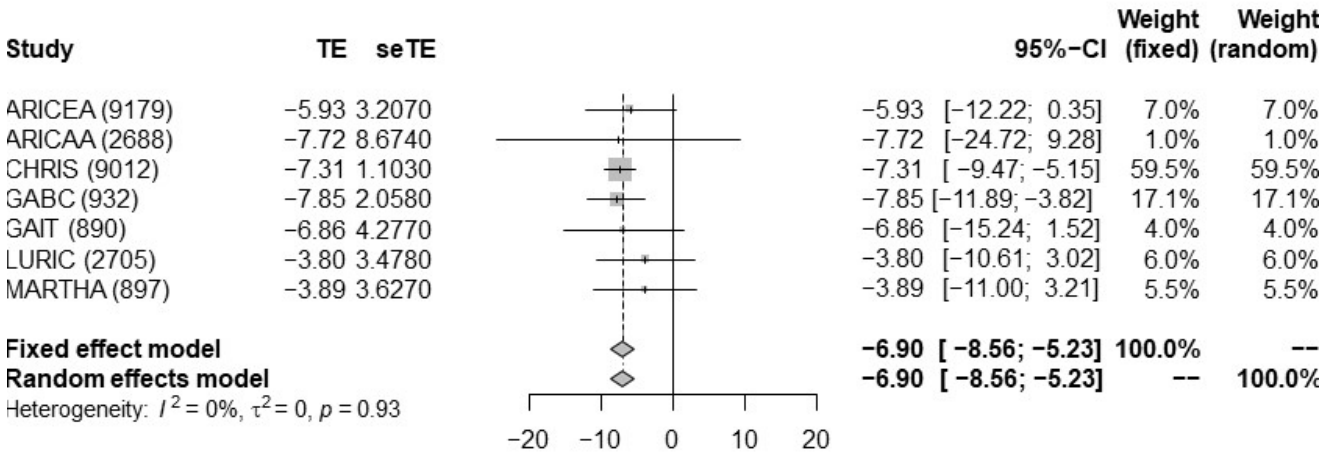

**Supplementary Figure 3. Forest plot for genome-wide signal at rs61826815 in AT analysis**



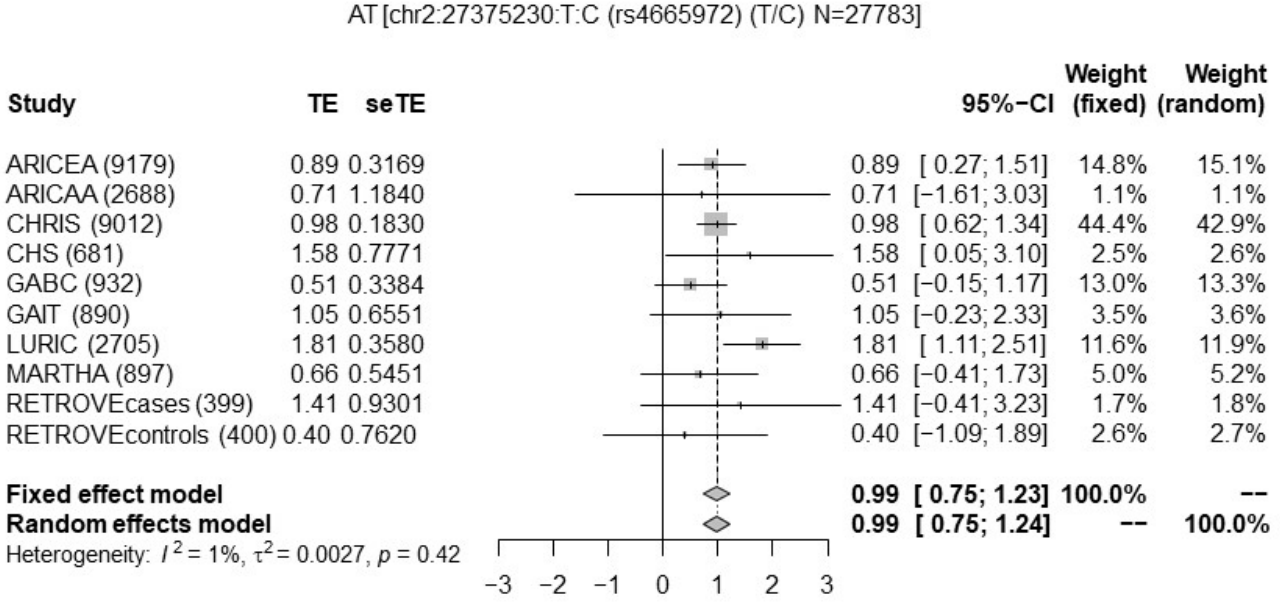

**Supplementary Figure 4. Forest plot for genome-wide signal at rs2227624 in AT analysis**



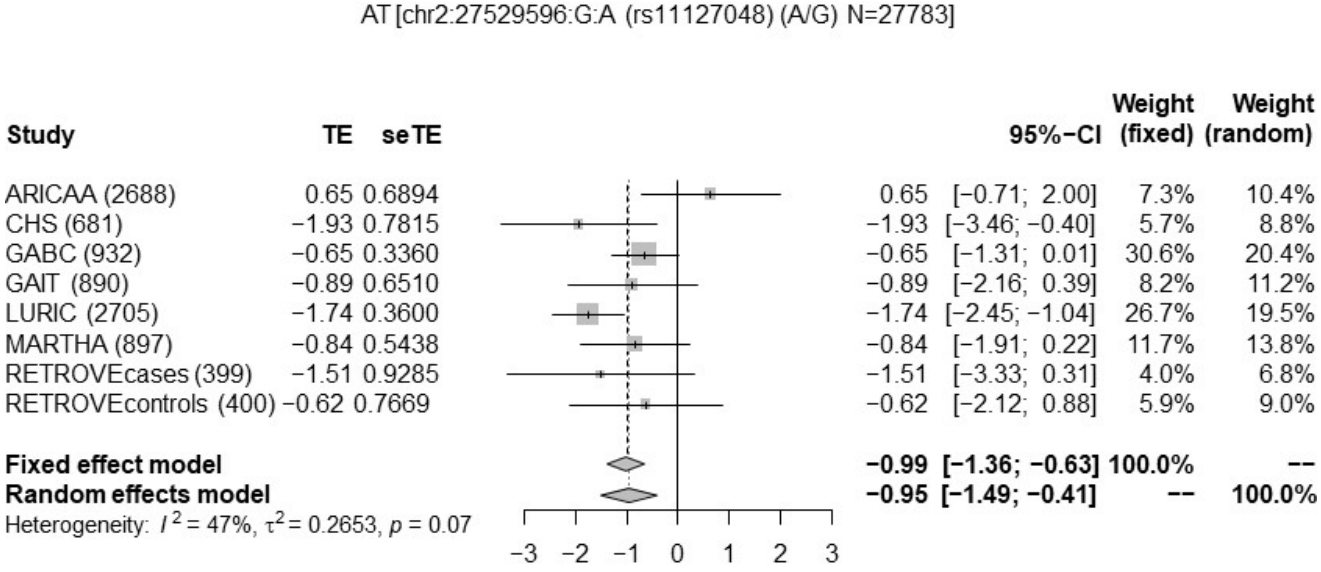

**Supplementary Figure 5. Forest plot for genome-wide signal at rs190938056 in AT analysis**



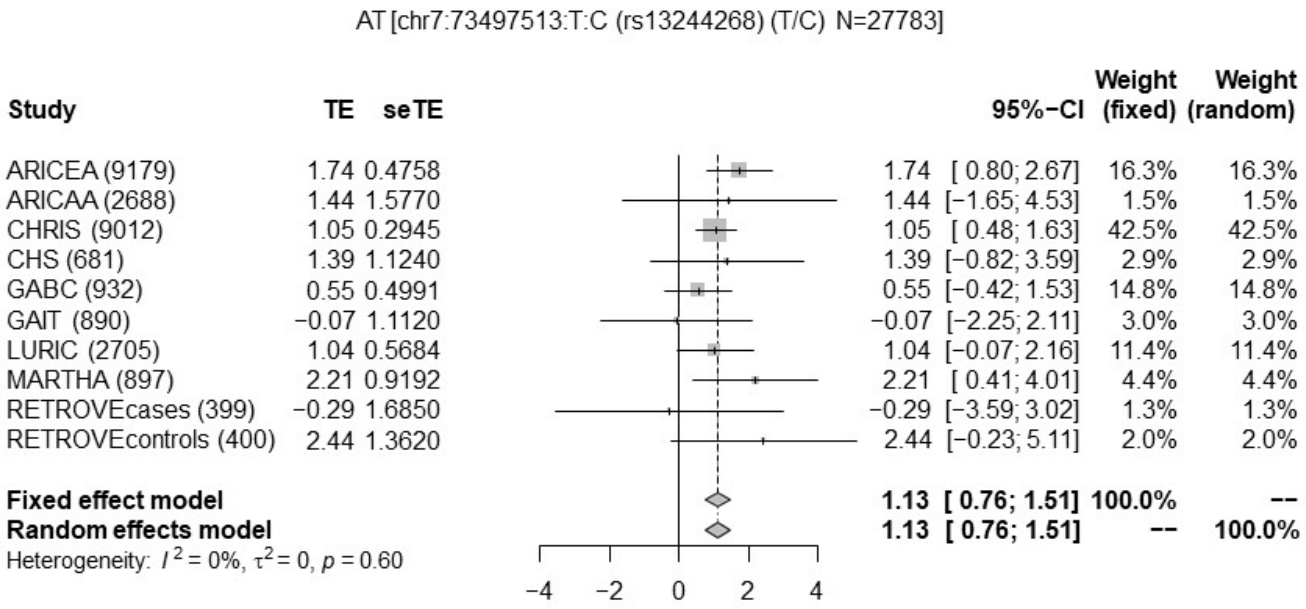

**Supplementary Figure 6. Forest plot for genome-wide signal at rs557150901 in AT analysis**



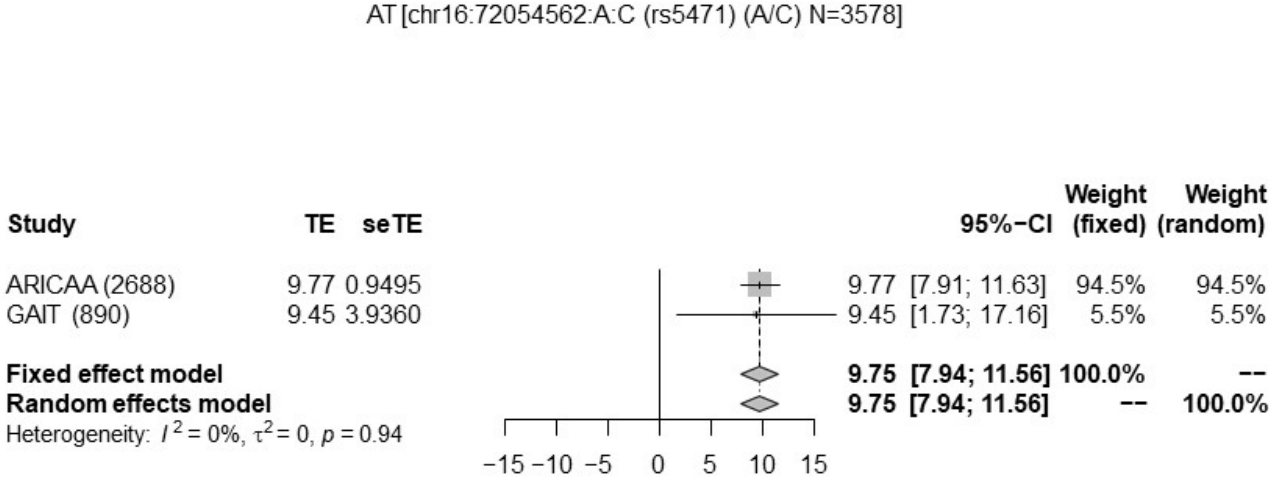

**Supplementary Figure 7. Forest plot for genome-wide signal at rs4665972 in AT analysis**



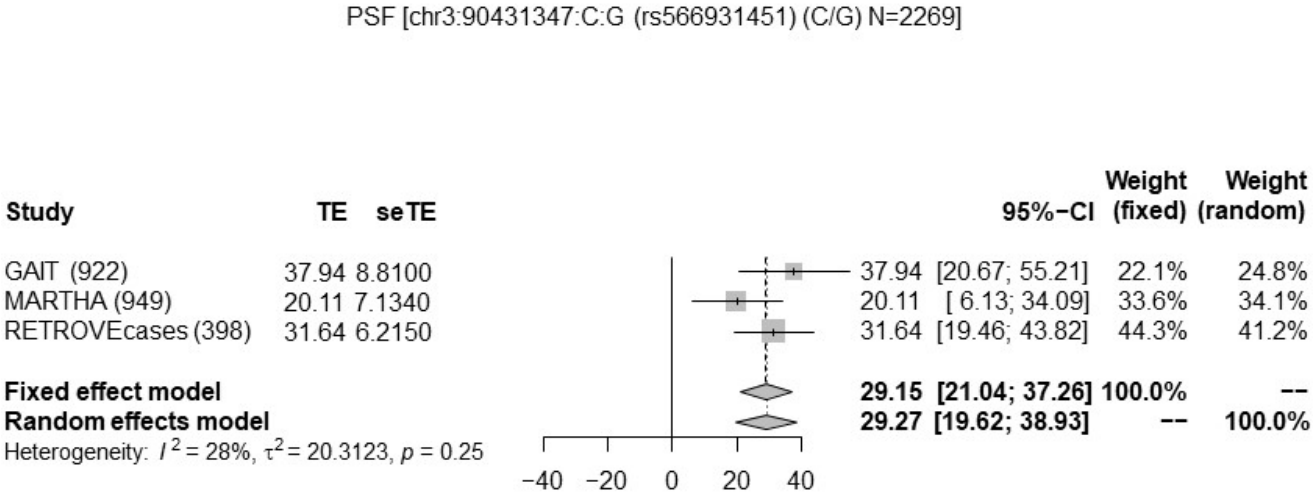

**Supplementary Figure 8. Forest plot for genome-wide signal at rs27529596 in AT analysis**



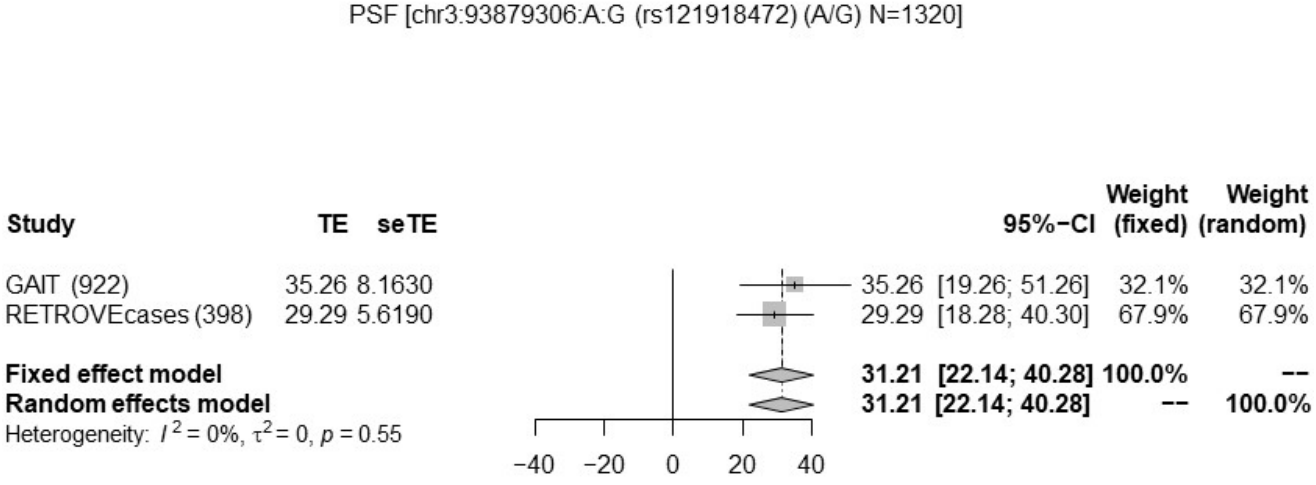

**Supplementary Figure 9. Forest plot for genome-wide signal at rs13244268 in AT analysis**



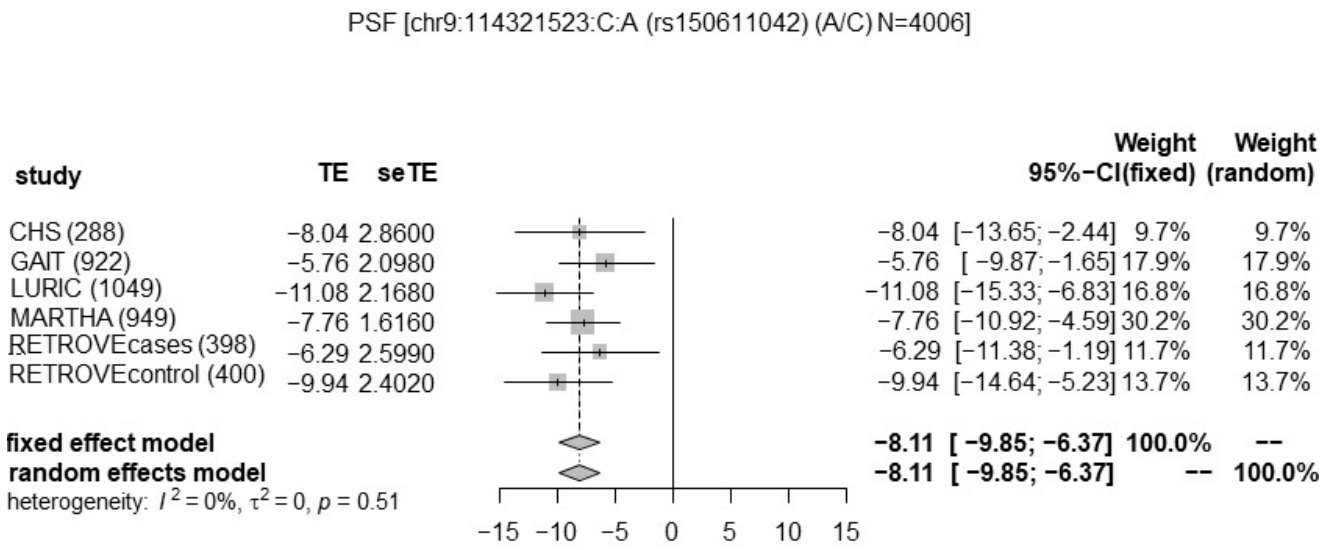

**Supplementary Figure 10. Forest plot for genome-wide signal at rs72054562 in AT analysis**



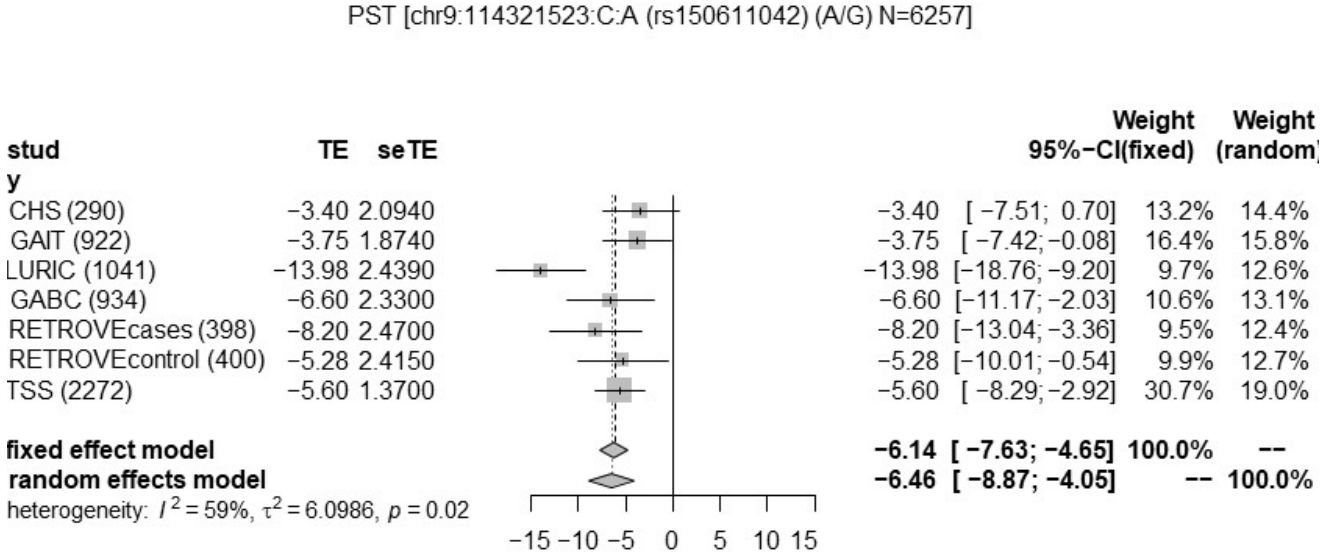

**Supplementary Figure 11. Forest plot for genome-wide signal at rs566931451 in PSF analysis**




**Supplementary Figure 12. Forest plot for genome-wide signal at rs121918472 in PSF analysis**




**Supplementary Figure 13. Forest plot for genome-wide signal at rs150611042 in PSF analysis**




**Supplementary Figure 14. Forest plot for genome-wide signal at rs150611042 in PST analysis**

REFERENCE


1. André W. Broekmans, Jan J. Veltkamp, and Rogier M. Bertina, *Congenital Protein C Deficiency and Venous Thromboembolism.* New England Journal of Medicine, 1983. **309**(6): p. 340-344.
2. John H Griffin, Bruce Evatt, Theodore S Zimmerman, Alice J Kleiss, and C Wideman, *Deficiency of protein C in congenital thrombotic disease.* The Journal of clinical investigation, 1981. **68**(5): p. 1370-1373.
3. Hans Peter Schwarz, Michael Fischer, Pierre Hopmeier, Mary Ann Batard, and John H Griffin, *Plasma protein S deficiency in familial thrombotic disease.* 1984.
4. O Egeberg, *Inherited antithrombin deficiency causing thrombophilia.* Thrombosis and Haemostasis, 1965. **13**(01): p. 516-530.
5. A. R. Folsom, N. Aleksic, L. Wang, M. Cushman, K. K. Wu, and R. H. White, *Protein C, antithrombin, and venous thromboembolism incidence: a prospective population-based study.* Arterioscler Thromb Vasc Biol, 2002. **22**(6): p. 1018-22.
6. Charles T Esmon, *The roles of protein C and thrombomodulin in the regulation of blood coagulation.* J Biol Chem, 1989. **264**: p. 4743-4746.
7. Phd Charles T. Esmon, *Protein C pathway.* CHEST, 2013. **124**(3): p. 26-32.
8. Lisbeth Fa Maurissen, M Christella Lgd Thomassen, Gerry Af Nicolaes, Björn Dahlbäck, Guido Tans, Jan Rosing, and Tilman M Hackeng, *Re-evaluation of the role of the protein S-C4b binding protein complex in activated protein C-catalyzed factor Va-inactivation.* Blood, The Journal of the American Society of Hematology, 2008. **111**(6): p. 3034-3041.
9. S. M. Rezende, R. E. Simmonds, and D. A. Lane, *Coagulation, inflammation, and apoptosis: different roles for protein S and the protein S-C4b binding protein complex.* Blood, 2004. **103**(4): p. 1192-201.
10. Joost Meijers, Laurent Mosnier, and Bonno Bouma, *The Role of Protein S in the Activation of Thrombin Activatable Fibrinolysis Inhibitor (TAFI) and Regulation of Fibrinolysis.* Thrombosis and Haemostasis, 2017. **86**(10): p. 1040-1046.
11. Elena M Faioni, Carla Valsecchi, Alessandra Palla, Emanuela Taioli, Cristina Razzari, and Pier Mannuccio Mannucci, *Free protein S deficiency is a risk factor for venous thrombosis.* Thrombosis and haemostasis, 1997. **78**(01): p. 1343-1346.
12. Noelene S. Quinsey, Ainslie L. Greedy, Stephen P. Bottomley, James C. Whisstock, and Robert N. Pike, *Antithrombin: in control of coagulation.* The International Journal of Biochemistry & Cell Biology, 2004. **36**(3): p. 386-389.
13. Paolo Bucciarelli, Frits R Rosendaal, Armando Tripodi, Pier Mannuccio Mannucci, Valerio De Stefano, Gualtiero Palareti, Guido Finazzi, Francesco Baudo, and Roberto Quintavalla, *Risk of venous thromboembolism and clinical manifestations in carriers of antithrombin, protein C, protein S deficiency, or activated protein C resistance: a multicenter collaborative family study.* Arteriosclerosis, thrombosis, and vascular biology, 1999. **19**(4): p. 1026-1033.
14. C Leroy-Matheron, J Duchemin, M Levent, and M Gouault-Heilmann, *Genetic modulation of plasma protein S levels by two frequent dimorphisms in the PROS1 gene.* Thrombosis and haemostasis, 1999. **82**(09): p. 1088-1092.
15. Mrinal M Patnaik and Stephan Moll, *Inherited antithrombin deficiency: a review.* Haemophilia, 2008. **14**(6): p. 1229-1239.
16. M. K. Shamsher, N. A. Chuzhanova, B. Friedman, D. A. Scopes, A. Alhaq, D. S. Millar, D. N. Cooper, and L. P. Berg, *Identification of an intronic regulatory element in the human protein C (PROC) gene.* Hum Genet, 2000. **107**(5): p. 458-65.
17. C. B. Grundy, M. Chisholm, V. V. Kakkar, and D. N. Cooper, *A novel homozygous missense mutation in the protein C (PROC) gene causing recurrent venous thrombosis.* Hum Genet, 1992. **89**(6): p. 683-4.
18. Ds Millar, Cb Grundy, P Bignell, Dc Mitchell, D Corden, P Woods, Vv Kakkar, and Dn Cooper, *A novel nonsense mutation in the protein C (PROC) gene (Trp-29→ Term) causing recurrent venous thrombosis.* Human genetics, 1993. **91**(2): p. 196-196.
19. Martine Aiach, Viviane Nicaud, Martine Alhenc-Gelas, Sophie Gandrille, Emmanuel Arnaud, Jean Amiral, Louis Guize, Jean-NoëL Fiessinger, and Joseph Emmerich, *Complex association of protein C gene promoter polymorphism with circulating protein C levels and thrombotic risk.* Arteriosclerosis, thrombosis, and vascular biology, 1999. **19**(6): p. 1573-1576.
20. D. Wu, Z. Zhong, Y. Chen, H. Ding, M. Yang, N. Lian, Z. Huang, Q. Zhang, J. Zhao, and C. Deng, *Analysis of PROC and PROS1 single nucleotide polymorphisms in a thrombophilia family.* Clin Respir J, 2019. **13**(8): p. 530-537.
21. T. C. El-Galaly, M. T. Severinsen, K. Overvad, R. Steffensen, A. K. Vistisen, A. Tjonneland, and S. R. Kristensen, *Single nucleotide polymorphisms and the risk of venous thrombosis: results from a Danish case-cohort study.* Br J Haematol, 2013. **160**(6): p. 838-41.
22. C. Boinot, D. Borgel, A. Kitzis, M. Guicheteau, M. Aiach, and M. Alhenc-Gelas, *Familial thrombophilia is an oligogenetic disease: involvement of the prothrombin G20210A, PROC and PROS gene mutations.* Blood Coagul Fibrinolysis, 2003. **14**(2): p. 191-6.
23. Robert J Klein, Caroline Zeiss, Emily Y Chew, Jen-Yue Tsai, Richard S Sackler, Chad Haynes, Alice K Henning, John Paul Sangiovanni, Shrikant M Mane, and Susan T Mayne, *Complement factor H polymorphism in age-related macular degeneration.* Science, 2005. **308**(5720): p. 385-389.
24. Vivian Tam, Nikunj Patel, Michelle Turcotte, Yohan Bossé, Guillaume Paré, and David Meyre, *Benefits and limitations of genome-wide association studies.* Nature Reviews Genetics, 2019. **20**(8): p. 467-484.
25. J. S. Pankow, W. Tang, N. Pankratz, W. Guan, L. C. Weng, M. Cushman, E. Boerwinkle, and A. R. Folsom, *Identification of Genetic Variants Linking Protein C and Lipoprotein Metabolism: The ARIC Study (Atherosclerosis Risk in Communities).* Arterioscler Thromb Vasc Biol, 2017. **37**(3): p. 589-597.
26. W. Tang, S. Basu, X. Kong, J. S. Pankow, N. Aleksic, A. Tan, M. Cushman, E. Boerwinkle, and A. R. Folsom, *Genome-wide association study identifies novel loci for plasma levels of protein C: the ARIC study.* Blood, 2010. **116**(23): p. 5032-6.
27. M. S. Munir, L. C. Weng, W. Tang, S. Basu, J. S. Pankow, N. Matijevic, M. Cushman, E. Boerwinkle, and A. R. Folsom, *Genetic markers associated with plasma protein C level in African Americans: the atherosclerosis risk in communities (ARIC) study.* Genet Epidemiol, 2014. **38**(8): p. 709-13.
28. T. Oudot-Mellakh, W. Cohen, M. Germain, N. Saut, C. Kallel, D. Zelenika, M. Lathrop, D. A. Tregouet, and P. E. Morange, *Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project.* Br J Haematol, 2012. **157**(2): p. 230-9.
29. G. Athanasiadis, A. Buil, J. C. Souto, M. Borrell, S. López, A. Martinez-Perez, M. Lathrop, J. Fontcuberta, L. Almasy, and J. M. Soria, *A genome-wide association study of the Protein C anticoagulant pathway.* PLoS One, 2011. **6**(12): p. e29168.
30. María Eugenia De La Morena-Barrio, Alfonso Buil, Ana Isabel Antón, Irene Martínez-Martínez, Antonia Miñano, Ricardo Gutiérrez-Gallego, José Navarro-Fernández, Sonia Aguila, Juan Carlos Souto, and Vicente Vicente, *Identification of*

*antithrombin-modulating genes. Role of LARGE, a gene encoding a bifunctional glycosyltransferase, in the secretion of proteins?* PloS one, 2013. **8**(5): p. e64998.

31. Me De La Morena-Barrio and A Buil, *Antó n AI, Martınez-Martınez I, Min ano A, et al.(2013) Identification of Antithrombin-Modulating Genes.* Role of LARGE, a.
32. M. E. Kleber, I. Seppala, S. Pilz, M. M. Hoffmann, A. Tomaschitz, N. Oksala, E. Raitoharju, L. P. Lyytikainen, K. M. Makela, R. Laaksonen, M. Kahonen, O. T. Raitakari, J. Huang, K. Kienreich, A. Fahrleitner-Pammer, C. Drechsler, V. Krane, B. O. Boehm, W. Koenig, C. Wanner, T. Lehtimaki, W. Marz, and A. Meinitzer, *Genome-wide association study identifies 3 genomic loci significantly associated with serum levels of homoarginine: the AtheroRemo Consortium.* Circ Cardiovasc Genet, 2013. **6**(5): p. 505-13.
33. Nicholas L Smith, Lucia A Hindorff, Susan R Heckbert, Rozenn N Lemaitre, Kristin D Marciante, Kenneth Rice, Thomas Lumley, Joshua C Bis, Kerri L Wiggins, and Frits R Rosendaal, *Association of genetic variations with nonfatal venous thrombosis in postmenopausal women.* Jama, 2007. **297**(5): p. 489-498.
34. J. C. Souto, L. Almasy, M. Borrell, F. Blanco-Vaca, J. Mateo, J. M. Soria, I. Coll, R. Felices, W. Stone, J. Fontcuberta, and J. Blangero, *Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. Genetic Analysis of Idiopathic Thrombophilia.* Am J Hum Genet, 2000. **67**(6): p. 1452-9.
35. Bruce M Psaty, Christopher J O'donnell, Vilmundur Gudnason, Kathryn L Lunetta, Aaron R Folsom, Jerome I Rotter, André G Uitterlinden, Tamara B Harris, Jacqueline Cm Witteman, and Eric Boerwinkle, *Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts.* Circulation: Cardiovascular Genetics, 2009. **2**(1): p. 73-80.
36. Jens Baumert, Jie Huang, Barbara Mcknight, Maria Sabater-Lleal, Maristella Steri, Audrey Y Chu, Stella Trompet, Lorna M Lopez, Myriam Fornage, and Alexander Teumer, *No evidence for genome-wide interactions on plasma fibrinogen by smoking, alcohol consumption and body mass index: results from meta-analyses of 80,607 subjects.* PloS one, 2014. **9**(12).
37. Paul S De Vries, Daniel I Chasman, Maria Sabater-Lleal, Ming-Huei Chen, Jennifer E Huffman, Maristella Steri, Weihong Tang, Alexander Teumer, Riccardo E Marioni, and Vera Grossmann, *A meta-analysis of 120 246 individuals identifies 18 new loci for fibrinogen concentration.* Human molecular genetics, 2016. **25**(2): p. 358-370.
38. Paul S De Vries, Maria Sabater-Lleal, Daniel I Chasman, Stella Trompet, Tarunveer S Ahluwalia, Alexander Teumer, Marcus E Kleber, Ming-Huei Chen, Jie Jin Wang, and John R Attia, *Comparison of HapMap and 1000 genomes reference panels in a large-scale genome-wide association study.* PloS one, 2017. **12**(1).
39. Paul S De Vries, Maria Sabater-Lleal, Jennifer E Huffman, Jonathan Marten, Ci Song, Nathan Pankratz, Traci M Bartz, Hugoline G De Haan, Graciela E Delgado, and John D Eicher, *A genome-wide association study identifies new loci for factor VII and implicates factor VII in ischemic stroke etiology.* Blood, The Journal of the American Society of Hematology, 2019. **133**(9): p. 967-977.
40. Abbas Dehghan, Qiong Yang, Annette Peters, Saonli Basu, Joshua C Bis, Alicja R Rudnicka, Maryam Kavousi, Ming-Huei Chen, Jens Baumert, and Gordon Do Lowe, *Association of novel genetic Loci with circulating fibrinogen levels: a genome-wide association study in 6 population-based cohorts.* Circulation: Cardiovascular Genetics, 2009. **2**(2): p. 125-133.
41. Jie Huang, Jennifer E Huffman, Munekazu Yamakuchi, Stella Trompet, Folkert W Asselbergs, Maria Sabater-Lleal, David-Alexandre Trégouët, Wei-Min Chen, Nicholas L Smith, and Marcus E Kleber, *Genome-wide association study for circulating tissue plasminogen activator levels and functional follow-up implicates endothelial STXBP5 and STX2.* Arteriosclerosis, thrombosis, and vascular biology, 2014. **34**(5): p. 1093-1101.
42. Jie Huang, Maria Sabater-Lleal, Folkert W Asselbergs, David Tregouet, So-Youn Shin, Jingzhong Ding, Jens Baumert, Tiphaine Oudot-Mellakh, Lasse Folkersen, and Andrew D Johnson, *Genome-wide association study for circulating levels of PAI-1 provides novel insights into its regulation.* Blood, The Journal of the American Society of Hematology, 2012. **120**(24): p. 4873-4881.
43. Jennifer E Huffman, Paul S De Vries, Alanna C Morrison, Maria Sabater-Lleal, Tim Kacprowski, Paul L Auer, Jennifer A Brody, Daniel I Chasman, Ming-Huei Chen, and Xiuqing Guo, *Rare and low-frequency variants and their association with plasma levels of fibrinogen, FVII, FVIII, and vWF.* Blood, 2015. **126**(11): p. e19-e29.
44. Maria Sabater-Lleal, Jie Huang, Daniel Chasman, Silvia Naitza, Abbas Dehghan, Andrew D Johnson, Alexander Teumer, Alex P Reiner, Lasse Folkersen, and Saonli Basu, *Multiethnic meta-analysis of genome-wide association studies in> 100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease.* Circulation, 2013. **128**(12): p. 1310-1324.
45. Bengt Sennblad, Saonli Basu, Johanna Mazur, Pierre Suchon, Angel Martinez-Perez, Astrid Van Hylckama Vlieg, Vinh Truong, Yuhuang Li, Jesper R Gådin, and Weihong Tang, *Genome-wide association study with additional genetic and post-transcriptional analyses reveals novel regulators of plasma factor XI levels.* Human molecular genetics, 2017. **26**(3): p. 637-649.
46. Nicholas L Smith, Ming-Huei Chen, Abbas Dehghan, David P Strachan, Saonli Basu, Nicole Soranzo, Caroline Hayward, Igor Rudan, Maria Sabater-Lleal, and Joshua C Bis, *Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE Consortium.* Circulation, 2010. **121**(12): p. 1382.
47. Nicholas L Smith, Jennifer E Huffman, David P Strachan, Jie Huang, Abbas Dehghan, Stella Trompet, Lorna M Lopez, So-Youn Shin, Jens Baumert, and Veronique Vitart, *Genetic predictors of fibrin D-dimer levels in healthy adults.* Circulation, 2011. **123**(17): p. 1864-1872.
48. Weihong Tang, Christine Schwienbacher, Lorna M Lopez, Yoav Ben-Shlomo, Tiphaine Oudot-Mellakh, Andrew D Johnson, Nilesh J Samani, Saonli Basu, Martin Gögele, and Gail Davies, *Genetic associations for activated partial thromboplastin time and prothrombin time, their gene expression profiles, and risk of coronary artery disease.* The American Journal of Human Genetics, 2012. **91**(1): p. 152-162.
49. Janine Van Loon, Abbas Dehghan, Tang Weihong, Stella Trompet, Wendy L Mcardle, Folkert Fw Asselbergs, Ming-Huei Chen, Lorna M Lopez, Jennifer E Huffman, and Frank Wg Leebeek, *Genome-wide association studies identify genetic loci for low von Willebrand factor levels.* European Journal of Human Genetics, 2016. **24**(7): p. 1035-1040.
50. Janine E Van Loon, Frank Wg Leebeek, Jaap W Deckers, Diederik Wj Dippel, Don Poldermans, David P Strachan, Weihong

Tang, Christopher J O'donnell, Nicholas L Smith, and Moniek Pm De Maat, *Effect of genetic variations in syntaxin-binding protein-5 and syntaxin-2 on von Willebrand factor concentration and cardiovascular risk.* Circulation: Cardiovascular Genetics, 2010. **3**(6): p. 507-512.

51. Cavin K Ward-Caviness, Paul S De Vries, Kerri L Wiggins, Jennifer E Huffman, Lisa R Yanek, Lawrence F Bielak, Franco Giulianini, Xiuqing Guo, Marcus E Kleber, and Tim Kacprowski, *Mendelian randomization evaluation of causal effects of fibrinogen on incident coronary heart disease.* PloS one, 2019. **14**(5).
52. Cavin K Ward-Caviness, Jennifer E Huffman, Karl Everett, Marine Germain, Jenny Van Dongen, W David Hill, Min A Jhun, Jennifer A Brody, Mohsen Ghanbari, and Lei Du, *DNA methylation age is associated with an altered hemostatic profile in a multiethnic meta-analysis.* blood, 2018. **132**(17): p. 1842-1850.
53. Juan Carlos Souto, Laura Almasy, Montserrat Borrell, Francisco Blanco-Vaca, José Mateo, José Manuel Soria, Inma Coll, Rosa Felices, William Stone, and Jordi Fontcuberta, *Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study.* The American Journal of Human Genetics, 2000. **67**(6): p. 1452-1459.
54. Karl C Desch, Ayse B Ozel, David Siemieniak, Yossi Kalish, Jordan A Shavit, Courtney D Thornburg, Anjali A Sharathkumar, Caitlin P Mchugh, Cathy C Laurie, and Andrew Crenshaw, *Linkage analysis identifies a locus for plasma von Willebrand factor undetected by genome-wide association.* Proceedings of the National Academy of Sciences, 2013. **110**(2): p. 588-593.
55. Linda P Fried, Nemat O Borhani, Paul Enright, Curt D Furberg, Julius M Gardin, Richard A Kronmal, Lewis H Kuller, Teri A Manolio, Maurice B Mittelmark, and Anne Newman, *The cardiovascular health study: design and rationale.* Annals of epidemiology, 1991. **1**(3): p. 263-276.
56. Gordon Do Lowe, Ann Rumley, Peter M Sweetnam, John Wg Yarnell, and Joseph Rumley, *Fibrin D-dimer, markers of coagulation activation and the risk of major ischaemic heart disease in the caerphilly study.* Thrombosis and haemostasis, 2001. **86**(09): p. 822-827.
57. Cristian Pattaro, Martin Gögele, Deborah Mascalzoni, Roberto Melotti, Christine Schwienbacher, Alessandro De Grandi, Luisa Foco, Yuri D'elia, Barbara Linder, and Christian Fuchsberger, *The Cooperative Health Research in South Tyrol (CHRIS) study: rationale, objectives, and preliminary results.* Journal of translational medicine, 2015. **13**(1): p. 348.
58. Madeline H Kowalski, Huijun Qian, Ziyi Hou, Jonathan D Rosen, Amanda L Tapia, Yue Shan, Deepti Jain, Maria Argos, Donna K Arnett, and Christy Avery, *Use of> 100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium whole genome sequences improves imputation quality and detection of rare variant associations in admixed African and Hispanic/Latino populations.* bioRxiv, 2019: p. 683201.
59. T. W. Winkler, F. R. Day, D. C. Croteau-Chonka, A. R. Wood, A. E. Locke, R. Magi, T. Ferreira, T. Fall, M. Graff, A. E. Justice, J. Luan, S. Gustafsson, J. C. Randall, S. Vedantam, T. Workalemahu, T. O. Kilpelainen, A. Scherag, T. Esko, Z. Kutalik, I. M. Heid, R. J. Loos, and Consortium Genetic Investigation of Anthropometric Traits, *Quality control and conduct of genome-wide association meta-analyses.* Nat Protoc, 2014. **9**(5): p. 1192-212.
60. Cristen J Willer, Yun Li, and Gonçalo R Abecasis, *METAL: fast and efficient meta-analysis of genomewide association scans.* Bioinformatics, 2010. **26**(17): p. 2190-2191.
61. Panagis Filippakopoulos and Stefan Knapp, *The bromodomain interaction module.* FEBS letters, 2012. **586**(17): p. 2692-2704.
62. M. Soejima, K. Teye, and Y. Koda, *The haptoglobin promoter polymorphism rs5471 is the most definitive genetic determinant of serum haptoglobin level in a Ghanaian population.* Clin Chim Acta, 2018. **483**: p. 303-307.
63. T. M. Teslovich, K. Musunuru, A. V. Smith, A. C. Edmondson, I. M. Stylianou, M. Koseki, J. P. Pirruccello, S. Ripatti, D. I. Chasman, C. J. Willer, C. T. Johansen, S. W. Fouchier, A. Isaacs, G. M. Peloso, M. Barbalic, S. L. Ricketts, J. C. Bis, Y. S. Aulchenko, G. Thorleifsson, M. F. Feitosa, J. Chambers, M. Orho-Melander, O. Melander, T. Johnson, X. Li, X. Guo, M. Li, Y. Shin Cho, M. Jin Go, Y. Jin Kim, J. Y. Lee, T. Park, K. Kim, X. Sim, R. Twee-Hee Ong, D. C. Croteau-Chonka, L. A. Lange, J. D. Smith, K. Song, J. Hua Zhao, X. Yuan, J. Luan, C. Lamina, A. Ziegler, W. Zhang, R. Y. Zee, A. F. Wright, J. C. Witteman, J. F. Wilson, G. Willemsen, H. E. Wichmann, J. B. Whitfield, D. M. Waterworth, N. J. Wareham, G. Waeber, P. Vollenweider, B. F. Voight, V. Vitart, A. G. Uitterlinden, M. Uda, J. Tuomilehto, J. R. Thompson, T. Tanaka, I. Surakka, H. M. Stringham, T. D. Spector, N. Soranzo, J. H. Smit, J. Sinisalo, K. Silander, E. J. Sijbrands, A. Scuteri, J. Scott, D. Schlessinger, S. Sanna, V. Salomaa, J. Saharinen, C. Sabatti, A. Ruokonen, I. Rudan, L. M. Rose, R. Roberts, M. Rieder, B. M. Psaty, P. P. Pramstaller, I. Pichler, M. Perola, B. W. Penninx, N. L. Pedersen, C. Pattaro, A. N. Parker, G. Pare, B. A. Oostra, C. J. O'donnell, M. S. Nieminen, D. A. Nickerson, G. W. Montgomery, T. Meitinger, R. Mcpherson, M. I. Mccarthy, W. Mcardle, D. Masson, N. G. Martin, F. Marroni, M. Mangino, P. K. Magnusson, G. Lucas, R. Luben, R. J. Loos, M. L. Lokki, G. Lettre, C. Langenberg, L. J. Launer, E. G. Lakatta, R. Laaksonen, K. O. Kyvik, F. Kronenberg, I. R. König, K. T. Khaw, J. Kaprio, L. M. Kaplan, A. Johansson, M. R. Jarvelin, A. C. Janssens, E. Ingelsson, W. Igl, G. Kees Hovingh, J. J. Hottenga, A. Hofman, A. A. Hicks, C. Hengstenberg, I. M. Heid, C. Hayward, A. S. Havulinna, N. D. Hastie, T. B. Harris, T. Haritunians, A. S. Hall, U. Gyllensten, C. Guiducci, L. C. Groop, E. Gonzalez, C. Gieger, N. B. Freimer, L. Ferrucci, J. Erdmann, P. Elliott, K. G. Ejebe, A. Döring, A. F. Dominiczak, S. Demissie, P. Deloukas, E. J. De Geus, U. De Faire, G. Crawford, F. S. Collins, Y. D. Chen, M. J. Caulfield, H. Campbell, N. P. Burtt, L. L. Bonnycastle, D. I. Boomsma, S. M. Boekholdt, R. N. Bergman, I. Barroso, S. Bandinelli, C. M. Ballantyne, T. L. Assimes, T. Quertermous, D. Altshuler, M. Seielstad, T. Y. Wong, E. S. Tai, A. B. Feranil, C. W. Kuzawa, L. S. Adair, H. A. Taylor, Jr., I. B. Borecki, S. B. Gabriel, J. G. Wilson, H. Holm, U. Thorsteinsdottir, V. Gudnason, R. M. Krauss, K. L. Mohlke, J. M. Ordovas, P. B. Munroe, J. S. Kooner, A. R. Tall, R. A. Hegele, J. J. Kastelein, E. E. Schadt, J. I. Rotter, E. Boerwinkle, D. P. Strachan, V. Mooser, K. Stefansson, M. P. Reilly, N. J. Samani, H. Schunkert, L. A. Cupples, M. S. Sandhu, P. M. Ridker, D. J. Rader, C. M. Van Duijn, L. Peltonen, G. R. Abecasis, M. Boehnke and S. Kathiresan, *Biological, clinical and population relevance of 95 loci for blood lipids.* Nature, 2010. **466**(7307): p. 707-13.
64. F. Suarez-Sanchez, M. Vazquez-Moreno, E. Herrera-Lopez, J. H. Gomez-Zamudio, J. J. Peralta-Romero, O. D. Castelan-Martinez, M. Cruz, E. J. Parra, and A. Valladares-Salgado, *Association of rs2000999 in the haptoglobin gene with total cholesterol, HDL-C, and LDL-C levels in Mexican type 2 diabetes patients.* Medicine (Baltimore), 2019. **98**(39): p. e17298.
65. N. J. Samani, J. Erdmann, A. S. Hall, C. Hengstenberg, M. Mangino, B. Mayer, R. J. Dixon, T. Meitinger, P. Braund, H. E. Wichmann, J. H. Barrett, I. R. König, S. E. Stevens, S. Szymczak, D. A. Tregouet, M. M. Iles, F. Pahlke, H. Pollard, W. Lieb, F. Cambien, M. Fischer, W. Ouwehand, S. Blankenberg, A. J. Balmforth, A. Baessler, S. G. Ball, T. M. Strom, I. Braenne, C. Gieger, P. Deloukas, M. D. Tobin, A. Ziegler, J. R. Thompson, and H. Schunkert, *Genomewide association analysis of coronary*

*artery disease.* N Engl J Med, 2007. **357**(5): p. 443-53.
66. M. A. Coram, Q. Duan, T. J. Hoffmann, T. Thornton, J. W. Knowles, N. A. Johnson, H. M. Ochs-Balcom, T. A. Donlon, L. W. Martin, C. B. Eaton, J. G. Robinson, N. J. Risch, X. Zhu, C. Kooperberg, Y. Li, A. P. Reiner, and H. Tang, *Genome-wide characterization of shared and distinct genetic components that influence blood lipid levels in ethnically diverse human populations.* Am J Hum Genet, 2013. **92**(6): p. 904-16.
67. M. Alhenc-Gelas, M. Canonico, P. E. Morange, and J. Emmerich, *Protein S inherited qualitative deficiency: novel mutations and phenotypic influence.* J Thromb Haemost, 2010. **8**(12): p. 2718-26.
68. Y. Espinosa-Parrilla, G. Navarro, M. Morell, E. Abella, X. Estivill, and N. Sala, *Homozygosity for the protein S Heerlen allele is associated with type I PS deficiency in a thrombophilic pedigree with multiple risk factors.* Thromb Haemost, 2000. **83**(1): p. 102-6.
69. A. Rocanin-Arjo, W. Cohen, L. Carcaillon, C. Frère, N. Saut, L. Letenneur, M. Alhenc-Gelas, A. M. Dupuy, M. Bertrand, M. C. Alessi, M. Germain, P. S. Wild, T. Zeller, F. Cambien, A. H. Goodall, P. Amouyel, P. Y. Scarabin, D. A. Trégouët, and P. E. Morange, *A meta-analysis of genome-wide association studies identifies ORM1 as a novel gene controlling thrombin generation potential.* Blood, 2014. **123**(5): p. 777-85.
70. Rene Mulder, F Nanne Croles, Andre B Mulder, James A Huntington, Karina Meijer, and Michael V Lukens, *SERPINC 1 gene mutations in antithrombin deficiency.* British Journal of Haematology, 2017. **178**(2): p. 279-285.
71. S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W. W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, and H. L. Rehm, *Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.* Genet Med, 2015. **17**(5): p. 405-24.
72. K. Downes, K. Megy, D. Duarte, M. Vries, J. Gebhart, S. Hofer, O. Shamardina, S. V. V. Deevi, J. Stephens, R. Mapeta, S. Tuna, N. Al Hasso, M. W. Besser, N. Cooper, L. Daugherty, N. Gleadall, D. Greene, M. Haimel, H. Martin, S. Papadia, S. Revel-Vilk, S. Sivapalaratnam, E. Symington, W. Thomas, C. Thys, A. Tolios, C. J. Penkett, W. H. Ouwehand, S. Abbs, M. A. Laffan, E. Turro, I. Simeoni, A. D. Mumford, Y. M. C. Henskens, I. Pabinger, K. Gomez, and K. Freson, *Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders.* Blood, 2019. **134**(23): p. 2082-2091.
73. Dandan Zhao, Xuelian Li, Haihai Liang, Nan Zheng, Zhenwei Pan, Yuhong Zhou, Xiao Liu, Ming Qian, Bozhi Xu, and Ying Zhang, *SNX17 produces anti-arrhythmic effects by preserving functional SERCA2a protein in myocardial infarction.* International journal of cardiology, 2018. **272**: p. 298-305.
74. Walter Stockinger, Beate Sailler, Vera Strasser, Burgi Recheis, Daniela Fasching, Larissa Kahr, Wolfgang J Schneider, and Johannes Nimpf, *The PX□domain protein SNX17 interacts with members of the LDL receptor family and modulates endocytosis of the LDL receptor.* The EMBO Journal, 2002. **21**(16): p. 4259-4267.
75. Volker Florian, Thomas Schlüter, and Ralf Bohnensack, *A new member of the sorting nexin family interacts with the C-terminus of P-selectin.* Biochemical and biophysical research communications, 2001. **281**(4): p. 1045-1050.
76. Salam Salloum-Asfar, Raúl Teruel-Montoya, Ana B Arroyo, Nuria García-Barberá, Amarjit Chaudhry, Erin Schuetz, Ginés Luengo-Gil, Vicente Vicente, Rocío González-Conejero, and Constantino Martínez, *Regulation of coagulation factor XI expression by microRNAs in the human liver.* PLoS One, 2014. **9**(11): p. e111713.
77. Weihong Tang, Mary Cushman, David Green, Stephen S Rich, Leslie A Lange, Qiong Yang, Russell P Tracy, Geoffrey H Tofler, Saonli Basu, and James G Wilson, *Gene□centric approach identifies new and known loci for F VIII activity and VWF antigen levels in E uropean A mericans and A frican A mericans.* American journal of hematology, 2015. **90**(6): p. 534-540.
78. Jaclyn Ellis, Ethan M Lange, Jin Li, Josee Dupuis, Jens Baumert, Jeremy D Walston, Brendan J Keating, Peter Durda, Ervin R Fox, and Cameron D Palmer, *Large multiethnic Candidate Gene Study for C-reactive protein levels: identification of a novel association at CD36 in African Americans.* Human genetics, 2014. **133**(8): p. 985-995.
79. Aaron R Folsom, Pamela L Lutsey, Brad C Astor, and Mary Cushman, *C-reactive protein and venous thromboembolism: a prospective investigation in the ARIC cohort.* Thrombosis and haemostasis, 2009. **102**(4): p. 615.
80. Valegh Faid, Nouha Denguir, Violaine Chapuis, Nicolas Bihoreau, and Guillaume Chevreux, *Site□specific N□glycosylation analysis of human factor XI: Identification of a noncanonical NXC glycosite.* Proteomics, 2014. **14**(21-22): p. 2460-2470.
81. Matthew A Lalli, Jiwon Jang, Joo-Hye C Park, Yidi Wang, Elmer Guzman, Hongjun Zhou, Morgane Audouard, Daniel Bridges, Kenneth R Tovar, and Sorina M Papuc, *Haploinsufficiency of BAZ1B contributes to Williams syndrome through transcriptional dysregulation of neurodevelopmental pathways.* Human molecular genetics, 2016. **25**(7): p. 1294-1306.
82. S. Ligthart, A. Vaez, Y. H. Hsu, R. Stolk, A. G. Uitterlinden, A. Hofman, B. Z. Alizadeh, O. H. Franco, and A. Dehghan, *Bivariate genome-wide association study identifies novel pleiotropic loci for lipids and inflammation.* BMC Genomics, 2016. **17**: p. 443.
83. John C Wejman, David Hovsepian, Joseph S Wall, James F Hainfeld, and Jonathan Greer, *Structure of haptoglobin and the haptoglobin-hemoglobin complex by electron microscopy.* Journal of molecular biology, 1984. **174**(2): p. 319-341.
84. R Clive Landis, Pandelis Philippidis, Jan Domin, Joseph J Boyle, and Dorian O Haskard, *Haptoglobin genotype-dependent anti-inflammatory signaling in CD163+ Macrophages.* International journal of inflammation, 2013. **2013**.
85. Rabea Asleh, Rachael Miller-Lotan, Michael Aviram, Tony Hayek, Michael Yulish, Joanne E Levy, Benjamin Miller, Shany Blum, Uzi Milman, and Chen Shapira, *Haptoglobin genotype is a regulator of reverse cholesterol transport in diabetes in vitro and in vivo.* Circulation research, 2006. **99**(12): p. 1419-1425.
86. Rainer Vormittag, Thomas Vukovich, Christine Mannhalter, Erich Minar, Verena Schonauer, Christine Bialonczyk, Mirko Hirschl, and Ingrid Pabinger, *Haptoglobin phenotype 2-2 as a potentially new risk factor for spontaneous venous thromboembolism.* haematologica, 2005. **90**(11): p. 1557-1561.
87. Yury I Miller, Svetlana M Altamentova, and Nurith Shaklai, *Oxidation of low-density lipoprotein by hemoglobin stems from a heme-initiated globin radical: antioxidant role of haptoglobin.* Biochemistry, 1997. **36**(40): p. 12189-12198.
88. Simin Rota, Nicola A Mcwilliam, Trevor P Baglin, and Christopher D Byrne, *Atherogenic lipoproteins support assembly of the prothrombinase complex and thrombin generation: modulation by oxidation and vitamin E.* Blood, The Journal of the American Society of Hematology, 1998. **91**(2): p. 508-515.
89. Takeshi Kumagai, Yukari Hoshi, Hiromu Tsutsumi, Keiichi Ebina, and Katsushi Yokota, *Inhibition of plasma coagulation*

*through interaction between oxidized low-density lipoprotein and blood coagulation factor VIII.* Biological and Pharmaceutical Bulletin, 2005. **28**(6): p. 952-956.

90. P. Froguel, N. C. Ndiaye, A. Bonnefond, N. Bouatia-Naji, A. Dechaume, G. Siest, B. Herbeth, M. Falchi, L. Bottolo, R. M. Guéant-Rodriguez, C. Lecoeur, M. R. Langlois, Y. Labrune, A. Ruokonen, S. El Shamieh, M. G. Stathopoulou, A. Morandi, C. Maffeis, D. Meyre, J. R. Delanghe, P. Jacobson, L. Sjöström, L. M. Carlsson, A. Walley, P. Elliott, M. R. Jarvelin, G. V. Dedoussis, and S. Visvikis-Siest, *A genome-wide association study identifies rs2000999 as a strong genetic determinant of circulating haptoglobin levels.* PLoS One, 2012. **7**(3): p. e32327.
91. C. J. Willer, E. M. Schmidt, S. Sengupta, G. M. Peloso, S. Gustafsson, S. Kanoni, A. Ganna, J. Chen, M. L. Buchkovich, S. Mora, J. S. Beckmann, J. L. Bragg-Gresham, H. Y. Chang, A. Demirkan, H. M. Den Hertog, R. Do, L. A. Donnelly, G. B. Ehret, T. Esko, M. F. Feitosa, T. Ferreira, K. Fischer, P. Fontanillas, R. M. Fraser, D. F. Freitag, D. Gurdasani, K. Heikkilä, E. Hyppönen, A. Isaacs, A. U. Jackson, Å Johansson, T. Johnson, M. Kaakinen, J. Kettunen, M. E. Kleber, X. Li, J. Luan, L. P. Lyytikäinen, P. K. E. Magnusson, M. Mangino, E. Mihailov, M. E. Montasser, M. Müller-Nurasyid, I. M. Nolte, J. R. O'connell, C. D. Palmer, M. Perola, A. K. Petersen, S. Sanna, R. Saxena, S. K. Service, S. Shah, D. Shungin, C. Sidore, C. Song, R. J. Strawbridge, I. Surakka, T. Tanaka, T. M. Teslovich, G. Thorleifsson, E. G. Van Den Herik, B. F. Voight, K. A. Volcik, L. L. Waite, A. Wong, Y. Wu, W. Zhang, D. Absher, G. Asiki, I. Barroso, L. F. Been, J. L. Bolton, L. L. Bonnycastle, P. Brambilla, M. S. Burnett, G. Cesana, M. Dimitriou, A. S. F. Doney, A. Döring, P. Elliott, S. E. Epstein, G. Ingi Eyjolfsson, B. Gigante, M. O. Goodarzi, H. Grallert, M. L. Gravito, C. J. Groves, G. Hallmans, A. L. Hartikainen, C. Hayward, D. Hernandez, A. A. Hicks, H. Holm, Y. J. Hung, T. Illig, M. R. Jones, P. Kaleebu, J. J. P. Kastelein, K. T. Khaw, E. Kim, N. Klopp, P. Komulainen, M. Kumari, C. Langenberg, T. Lehtimäki, S. Y. Lin, J. Lindström, R. J. F. Loos, F. Mach, W. L. Mcardle, C. Meisinger, B. D. Mitchell, G. Müller, R. Nagaraja, N. Narisu, T. V. M. Nieminen, R. N. Nsubuga, I. Olafsson, K. K. Ong, A. Palotie, T. Papamarkou, C. Pomilla, A. Pouta, D. J. Rader, M. P. Reilly, P. M. Ridker, F. Rivadeneira, I. Rudan, A. Ruokonen, N. Samani, H. Scharnagl, J. Seeley, K. Silander, A. Stančáková, K. Stirrups, A. J. Swift, L. Tiret, A. G. Uitterlinden, L. J. Van Pelt, S. Vedantam, N. Wainwright, C. Wijmenga, S. H. Wild, G. Willemsen, T. Wilsgaard, J. F. Wilson, E. H. Young, J. H. Zhao, L. S. Adair, D. Arveiler, T. L. Assimes, S. Bandinelli, F. Bennett, M. Bochud, B. O. Boehm, D. I. Boomsma, I. B. Borecki, S. R. Bornstein, P. Bovet, M. Burnier, H. Campbell, A. Chakravarti, J. C. Chambers, Y. I. Chen, F. S. Collins, R. S. Cooper, J. Danesh, G. Dedoussis, U. De Faire, A. B. Feranil, J. Ferrières, L. Ferrucci, N. B. Freimer, C. Gieger, L. C. Groop, V. Gudnason, U. Gyllensten, A. Hamsten, T. B. Harris, A. Hingorani, J. N. Hirschhorn, A. Hofman, G. K. Hovingh, C. A. Hsiung, S. E. Humphries, S. C. Hunt, K. Hveem, C. Iribarren, M. R. Järvelin, A. Jula, M. Kähönen, J. Kaprio, A. Kesäniemi, M. Kivimaki, J. S. Kooner, P. J. Koudstaal, R. M. Krauss, D. Kuh, J. Kuusisto, K. O. Kyvik, M. Laakso, T. A. Lakka, L. Lind, C. M. Lindgren, N. G. Martin, W. März, M. I. Mccarthy, C. A. Mckenzie, P. Meneton, A. Metspalu, L. Moilanen, A. D. Morris, P. B. Munroe, I. Njølstad, N. L. Pedersen, C. Power, P. P. Pramstaller, J. F. Price, B. M. Psaty, T. Quertermous, R. Rauramaa, D. Saleheen, V. Salomaa, D. K. Sanghera, J. Saramies, P. E. H. Schwarz, W. H. Sheu, A. R. Shuldiner, A. Siegbahn, T. D. Spector, K. Stefansson, D. P. Strachan, B. O. Tayo, E. Tremoli, J. Tuomilehto, M. Uusitupa, C. M. Van Duijn, P. Vollenweider, L. Wallentin, N. J. Wareham, J. B. Whitfield, B. H. R. Wolffenbuttel, J. M. Ordovas, E. Boerwinkle, C. N. A. Palmer, U. Thorsteinsdottir, D. I. Chasman, J. I. Rotter, P. W. Franks, S. Ripatti, L. A. Cupples, M. S. Sandhu, S. S. Rich, M. Boehnke, P. Deloukas, S. Kathiresan, K. L. Mohlke, E. Ingelsson and G. R. Abecasis, *Discovery and refinement of loci associated with lipid levels.* Nat Genet, 2013. **45**(11): p. 1274-1283.
92. Jacqueline M Gertz and Beth A Bouchard, *Mechanisms Regulating Acquisition of Platelet□Derived Factor V/Va by Megakaryocytes.* Journal of Cellular Biochemistry, 2015. **116**(10): p. 2121-2126.
93. P. J. Talmud, F. Drenos, S. Shah, T. Shah, J. Palmen, C. Verzilli, T. R. Gaunt, J. Pallas, R. Lovering, K. Li, J. P. Casas, R. Sofat, M. Kumari, S. Rodriguez, T. Johnson, S. J. Newhouse, A. Dominiczak, N. J. Samani, M. Caulfield, P. Sever, A. Stanton, D. C. Shields, S. Padmanabhan, O. Melander, C. Hastie, C. Delles, S. Ebrahim, M. G. Marmot, G. D. Smith, D. A. Lawlor, P. B. Munroe, I. N. Day, M. Kivimaki, J. Whittaker, S. E. Humphries, and A. D. Hingorani, *Gene-centric association signals for lipids and apolipoproteins identified via the HumanCVD BeadChip.* Am J Hum Genet, 2009. **85**(5): p. 628-42.
94. Luciana Dente, Mg Pizza, A Metspalu, and R Cortese, *Structure and expression of the genes coding for human alpha 1□acid glycoprotein.* The EMBO journal, 1987. **6**(8): p. 2289-2296.
95. Mv Osikov, Ev Makarov, and Lv Krivokhizhina, *Effects of α 1-acid glycoprotein on hemostasis in experimental septic peritonitis.* Bulletin of experimental biology and medicine, 2007. **144**(2): p. 178-180.