High levels of complement activating enzyme MASP-2 are associated with risk of future incident venous thromboembolism

Short tittle: MASP-2 and venous thromboembolism

Christabel Esi Damoah,[1] Omri Snir,[1] Kristian Hindberg,[1] Peter Garred,[2] Judith K. Ludviksen,[3] Sigrid K. Brækkan,[1,4] Vânia M. Morelli,[1,4] Tom Eirik Mollnes,[1,3,5,6] John-Bjarne Hansen,[1,4] and the INVENT Consortium

[1]Thrombosis Research Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway; [2]Laboratory of Molecular Medicine, Department of Clinical Immunology, Section 7631, Rigshospitalet, Copenhagen, Denmark; [3]Reaserch Laboratory, Nordland Hospital, Bodø, Norway; [4]Division of Internal Medicine, University Hospital of North Norway, Tromsø, Norway; [5]Department of Immunology, Oslo University Hospital and University of Oslo, Norway; [6]Centre of Molecular Inflammation Research, Norwegian University of Science and Technology, Trondheim, Norway.

Corresponding author: Christabel Esi Damoah, Thrombosis Research Center (TREC), Department of Clinical Medicine, UiT- The Arctic University of Norway, N-9037 Tromsø, Norway. Tel: +47 77625105. Email: christabel.e.damoah@uit.no

Total word count: 8227 **(**Title Page, Abstract, Text, References, Tables and Figures Legends)

Abstract word count: 248

Number of tables and figures: 9 (3 tables and 6 figures)

Scientific Category: Thrombosis and Hemostasis

Authorship contributions

C.E.D. contributed to statistical analysis, interpreted data and drafted the manuscript; O.S.,P.G., and J.K.L. interpreted data and revised the manuscript; K.H. performed the Mendelian randomization analysis, interpreted data and revised the manuscript; S.K.B. organized data collection, contributed to statistical analyses, interpreted data, drafted and revised the manuscript; V.M.M. contributed to statistical analyses, interpreted data, drafted and revised the manuscript; T.E.M. performed the laboratory analysis, interpreted data and revised the manuscript; J-B.H. conceived and designed the study, organized data collection, interpreted data, drafted and revised the manuscript.

ABSTRACT

Background: Experimental studies have shown that the complement activating enzyme MBL-associated serine protease 2 (MASP-2) exhibits a thrombin-like activity, and that inhibition of MASP-2 protects against thrombosis. In this study, we investigated whether plasma MASP-2 levels were associated with risk of future venous thromboembolism (VTE), and whether genetic variants linked to MASP-2 levels were associated with VTE risk.

Methods: We conducted a population-based nested case-control study involving 410 VTE patients and 842 age- and sex-matched controls derived from the Norwegian Tromsø Study. Logistic regression was used to estimate odds ratios (ORs) of VTE across MASP-2 quartiles. Whole exome sequencing and protein quantitative trait loci (pQTL) analyses were performed to assess genetic variants associated with MASP-2 levels. A two-sample Mendelian randomization study, also including data from the INVENT consortium, was performed to assess causality.

Results: Subjects with plasma MASP-2 in the highest quartile had a 48% higher OR of VTE (OR:1.48; 95% CI:1.06-2.06) and 83% higher OR of deep vein thrombosis (OR:1.83, 95% CI:1.23-2.73) compared with those with MASP-2 levels in the lowest quartile. The pQTL analysis revealed that three previously described gene variants, rs12711521 (minor allele frequency (MAF)=0.153) and rs72550870 (MAF=0.045) (missense variants in MASP2 gene) and rs2275527 (MAF=0.220) (exon-variant in the adjacent MTOR gene) explained 39% of the variation of MASP-2 plasma concentration. The OR of VTE per 1 SD increase in genetically predicted MASP-2 was 1.03 (95% CI:1.01-1.05, p=0.0011).

Conclusions: Our findings suggest that high plasma MASP-2 levels are causally associated with risk of future VTE.

Graphic Abstract: A graphic abstract is available for this article.

68

69 **Non-standard Abbreviations and Acronyms**

BMI	**Body mass index**
CI	**Confidence interval**
CRP	**C-reactive protein**
CTPA	**CT Pulmonary angiogram**
CVD	**Cardiovascular disease**
DVT	**Deep vein thrombosis**
EIA	**Enzyme-immunoassay**
FVL	**Factor V Leiden**
GWAS	**Genome-wide association study**
INVENT	**International Network of Venous Thromboembolism Clinical Research Network**
LD	**Linkage disequilibrium**
MAF	**Minor allele frequency**
MBL	**Mannose-binding lectin**
MR	**Mendelian randomization**
MASP-1	**Mannose binding lectin associated serine protease 1**

MASP-2	**Mannose-binding lectin associated serine protease 2**
OR	**Odds ratio**
PE	**Pulmonary embolism**
pQTL	**Protein quantitative trait loci**
PRM	**Pattern recognition molecules**
SNP	**Single nucleotide polymorphism**
UNN	**University Hospital of North Norway**
VTE	**Venous thromboembolism**

70

INTRODUCTION

Venous thromboembolism (VTE), including deep vein thrombosis (DVT) and pulmonary embolism (PE), is a frequent disease affecting 1-2 per 1000 individuals annually.[1] VTE is associated with severe complications, including post-thrombotic syndrome, post-PE syndrome, recurrence and death.[1-3] Although medical thromboprophylaxis is provided to patients at high risk (e.g. after major surgery),[4,5] the incidence of VTE has slightly increased over the past decades.[6-8] As the prevalence of major VTE-risk factors, such as aging, cancer, and obesity is increasing,[9-11] the incidence of VTE is expected to continue to increase during the coming years. To lower the burden of VTE in the society, new insights into biomarkers and pathophysiological mechanisms are crucial to improve risk stratification and targeted VTE prevention.

The complement system is an important part of the innate immune system, and several points of intersection between the complement and coagulation systems may potentially contribute to a prothrombotic phenotype upon complement activation.[12,13] Growing evidence accumulated over the last years suggests that components of the complement system are associated with VTE. Plasma C3 levels are associated with the risk of future VTE in observational studies derived from the general population,[14] and C3 deficient mice displayed lower thrombus frequency and thrombus weight compared to wild-type mice in the inferior vena cava stenosis model.[15] Furthermore, we recently reported that complement activation *in vivo*, assessed by measurement of sC5b-9, the soluble form of the terminal complement complex (TCC) in plasma,[16,17] was associated with risk of future VTE and unprovoked VTE events in particular.[18] These findings suggest that components of the complement system are not only predictive biomarkers of VTE risk but have the potential to be involved in the pathogenesis of the disease.

Pattern recognition molecules (PRMs) of the lectin pathway of the complement system comprise two protein families, namely collectins and ficolins. The former includes mannose-binding lectin (MBL), collectin-10, collectin-11 and the latter ficolin-1, ficolin-2 and ficolin-3. These proteins circulate in the blood in complexes with three associated serine proteases named MASPs (1-3), and are activated when the PRMs bind to particular carbohydrate and/or acetylated moieties on pathogens or altered host cells.[19,20] We recently reported that subjects with low plasma MBL levels had lower VTE risk.[21] Apart from its canonical role in activating the complement system, both MASP-1 and MASP-2 have the ability to cleave prothrombin to thrombin with subsequent fibrin formation.[20,22-24] While MASP-1 has several substrates in the hemostatic system, including prothrombin, factor XIII, fibrinogen, and thrombin activatable fibrinolysis inhibitor, the activity of MASP-2 seems to be specific towards prothrombin.[25] Hence, the assessment of MASP-2 might provide novel insights into the pathogenesis of VTE that is particularly mediated by thrombin generation, which is probably a key mechanism of venous thrombus formation.[26] Moreover, elevated plasma MASP-2 levels have been reported in patients with acute ischemic stroke compared to healthy controls.[27] Additionally, inhibition of MASP-2 protects against stroke[28,29] and myocardial infarction[30,31] in animal models.

Based on these findings, we hypothesized that elevated plasma MASP-2 levels might be associated with an increased risk of future VTE. In the present nested case-control study derived from the general population comprising 410 VTE patients and 842 age- and sex-matched controls, we aimed to (i) investigate whether plasma MASP-2 levels were associated with risk of future VTE, (ii) identify genetic variants that regulated plasma MASP-2 levels, and (iii) explore whether these variants were associated with VTE risk in a Mendelian randomization framework.

METHODS

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Study Population

The Tromsø study is a population-based cohort with repeated health surveys of residents in the municipality Tromsø in the northern part of Norway[32]. To the fourth survey in 1994-1995, all inhabitants aged ≥25 years living in the municipality were invited to participate, and 27,158 subjects participated (77% response rate). These participants formed a cohort and were followed from their survey inclusion date (1994/95) until September 1, 2007. All first lifetime events of VTE occurring among the participants during follow-up were identified by searching the hospital discharge diagnosis registry, the autopsy registry, and the radiology procedure registry of the University Hospital of North Norway (UNN), the sole provider of hospital care in the Tromsø region. Trained personnel systematically reviewed the medical records and recorded each adjudicated VTE event, as previously described in detail.[33] In brief, an episode of VTE was adjudicated based on the presence of signs and symptoms of DVT or PE in combination with objective confirmation by radiological procedures (compression ultrasonography of the whole leg, venography, CTPA, perfusion-ventilation, pulmonary angiography or autopsy), that resulted in the initiation of treatment (unless contraindications were specified). A VTE event was further classified as unprovoked or provoked based on provoking factors closely preceding the VTE diagnosis. A VTE occurring in the presence of one or more of the following provoking factors was defined as provoked: recent hospitalization, surgery or trauma (within 8 weeks before the event), cancer, acute medical condition (acute myocardial infarction, acute ischemic stroke, acute infections), immobilization (bed rest >3 days, long distance travel of more than 4 hours duration during

the last 14 days, or confinement to a wheelchair within the last 8 weeks), or other factors described explicitly as provoking by a physician in the medical record (e.g. intravascular catheter).

During the cohort follow-up (1994-2007), 462 participants experienced a VTE event. We created a nested case-control study for the assessment of MASP-2 from stored blood samples from this cohort. In a nested case-control, the temporal sequence between exposure and outcome is preserved, and this design is therefore efficient to study biological precursors of disease. For each case, two age- and sex-matched controls (*n* =924), who were alive at the index date of the corresponding VTE case, were randomly sampled from the source cohort (Figure 1). A total of 52 cases and 82 controls were excluded because plasma samples were not available or of inadequate quality for the analyses. Thus, the final study population consisted of 410 cases and 842 controls. All participants provided written consent for participation in the study, and the regional committee for medical and health research ethics approved the study.

Baseline measurements

Baseline information at inclusion in the fourth survey in the Tromsø study (1994/95) was collected by physical examination, blood samples and a self-administered questionnaire. Height (to the nearest centimeter) and weight (to the nearest 0.5 kg) were measured in participants wearing light clothing and no shoes. Body mass index (BMI) was calculated as weight divided by the square of height in meters (kg/m^2). Information on smoking status, history of previous cardiovascular disease (CVD) events (stroke, angina pectoris, transient

ischemic attack, and myocardial infarction), diabetes mellitus and cancer were retrieved from the questionnaire.

Blood and DNA sample collection and storage

At inclusion in 1994/95, non-fasting blood was collected from an antecubital vein into 5-mL vacutainers (Becton Dickinson, Le Pont de Claix, France) containing ethylenediaminetetraacetic acid (K_3-EDTA 40 µL, 0.37 mol/L per tube) as an anticoagulant. Platelet-poor plasma was prepared by centrifugation at 3000 *x* g for 10 min at room temperature, after which the supernatant was transferred into cryovials (Greiner Labotechnik, Nürtingen, Germany) in 1-mL aliquots and stored at -80°C. DNA isolated from blood was stored at the National CONOR Biobank.[32]

Measurements of plasma levels of MASP-2 and C-reactive protein (CRP)

To measure biomarkers in plasma, samples were thawed in a water bath at 37°C for 5 min, followed by centrifugation for 2 min at 13,500 *x* g to obtain platelet-free plasma.

MASP-2 was measured using a sandwich MASP-2 ELISA (Hycult Biotech, Uden, The Netherlands). The assay was performed according to the instructions from the manufacturer. Optical density was measured using a microplate reader (Infinite M200 pro from Tecan Trading AG, Switzerland). The intra- and inter-assay coefficients of variation were < 5% and 12.5%, respectively. Each sample was normalized with respect to a commercially available control provided by the manufacturer, and the global median of MASP-2 levels derived from all control subjects of the nested case-control study. The value of the manufacturer's control obtained in each plate of the MASP-2 ELISA was subtracted from the raw value, and the global median was added ($MASP\text{-}2_{Normalized\ value} = MASP\text{-}2_{Raw\ value} - MASP\text{-}2_{Control\ value} + MASP\text{-}2_{global\ median\ value}$).

CRP was measured by the high sensitive technique ("hsCRP") in duplicates by enzyme-immunoassay (EIA) using commercially available reagents (R&D Systems, Minneapolis, MN) in a 384 format using the combination of a SELMA (Jena, Germany) pipetting robot and a BioTek (Winooski, VT) dispenser/washer (EL406). Absorption was read at 450 nm with a wavelength correction set to 540 nm using an EIA plate reader (Synergy H1 Hybrid, BioTek, Winooski, VT). The intra- and inter-assay coefficients of variation were 2.6% and 9.1%, respectively.

Exome sequencing

Whole exome sequencing at high-coverage (≈100×) was carried out in a random subset of the nested case-control study population (353 VTE patients and 354 control subjects) by the use of the Agilent SureSelect 50Mb capture kit. The subsequently retrieved genotypes were effectively filtered[34] and imputations performed as previously described in detail.[35] In brief, using the information from the exome sequencing data, genotypes were imputed to the whole genome using Beagle[36] and haplotypes from unrelated individuals from the European (EUR) and East Asian (EAS) superpopulations of the 1000 Genomes Project Phase 3[37] for sites with a minor allele frequency [MAF] >1%.

Statistical Analysis

Association between MASP-2 levels and risk of VTE in the nested case-control study

Statistical analyses were carried out using Stata version 16 (StataCorp LLC, College Station, TX, USA) and R version 4 (The R Foundation for Statistical Computing, Vienna, Austria. https://cran.r-project.org). Plasma MASP-2 was categorized according to quartile cutoffs in the control population (<302, 302-549, 550-823, ≥824 ng/mL). Means and proportions of baseline characteristics across quartiles of MASP-2 were calculated using descriptive statistics. Logistic regression models were used to estimate odds ratio (OR) of VTE with 95% confidence intervals (CIs) according to quartiles of MASP-2 adjusted for the matching factors,[38] with the addition of BMI and CRP as adjustment variables in a second model. The lowest quartile of MASP-2 was used as the reference group. P-values for linear trend across increasing quartiles of MASP-2 were estimated. Separate analyses were additionally conducted with unprovoked VTE, DVT, and PE as outcomes.

Due to the long follow-up time (≥ 12 years for many individuals) in the source cohort, the results based on baseline MASP-2 measurements could be influenced by regression dilution bias. To address this, we performed analyses that restricted the maximum follow-up time from blood sampling to the VTE events, while keeping all controls in the analyses. The logistic regression analyses on time restrictions were set to require at least 10 VTE events, and ORs were generated at every time point a new VTE event occurred and plotted as a function of this maximum time.

To assess potential non-linearity between plasma MASP-2 levels and risk of VTE, a generalized additive regression plot was generated to visualize the association by modelling MASP-2 with a smoothing spline fit in a logistic model adjusted for age, sex, BMI and CRP. We created one plot for the full follow-up and one plot restricted to the first five years of follow-up. The MASP-2 levels were transformed to follow a perfect standard normal distribution with a mean value of zero and a standard deviation (SD) of one before entering the analyses.

Identification of single nucleotide polymorphisms (SNPs) associated with MASP-2 plasma levels in the pQTL analysis

After filtering and imputation, the whole-exome dataset contained 1,033,970 variants. A protein quantitative trait loci (pQTL) analysis was applied to identify genetic variants associated with regulation of MASP-2 plasma levels using samples collected at cohort baseline, when all participants were VTE-free individuals. This pQTL analysis was performed both in a genome-wide setting and restricted to the loci within ± 500 kb of the different genes involved in the complement system. The commonly used significance threshold of 5 ×

10^{-8} was used to adjust for multiple testing in the genome-wide setting. As the cis analysis in total contained 11,829 variants, a Bonferroni-based adjustment for multiple testing corresponded to a significance threshold of $-\log_{10}(0.05/11829) = 5.37$. The plasma MASP-2 values transformed to follow a perfect standard normal distribution were used in the pQTL analysis. The pQTL analysis was performed with the EPACTS (Efficient and Parallelizable Association Container Toolbox) software.[39] The EMMAX[40] (Efficient Mixed Model Association eXpedited) linear mixed model approach implemented within EPACTS was used to test for associations between MASP-2 and genetic variants while adjusting for covariates (age, sex, BMI, CRP and VTE status) and genetic relatedness between individuals in the cohort. Because the Tromsø study, which is the source of our nested case-control study, is a population based-cohort, it may naturally include some proportion of related individuals[35]. Of the 707 exome sequenced individuals, 6% were related to another individual in the study at an identity-by-descent value of 0.1. To search for independent genetic variants, we applied linkage disequilibrium (LD) pruning.

SNPs regulating MASP-2 levels and risk of VTE by Mendelian Randomization

A two-sample Mendelian Randomization (MR) study was performed to investigate the association between MASP-2 levels and risk of VTE from a causal perspective. The effect size of each single nucleotide polymorphism (SNP) on MASP-2 plasma levels was obtained from the pQTL analysis. We used genome-wide association study (GWAS) summary data from the International Network on Venous Thrombosis (INVENT) consortium meta-analysis, including 30,234 VTE cases and 172,122 controls from 18 studies,[41] to obtain the effect size estimates of the association between the individual SNPs and VTE. For each SNP, the two effect sizes

(i.e., SNP on MASP-2 and SNP on VTE) with corresponding standard errors were calculated, and based on these effect sizes, the estimated increase in OR of VTE per SD increase in genetically predicted MASP-2 levels was estimated using the inverse-variance weighted method of MRBase.[42] The obtained estimates of the causal inference based on MR were interpreted with the assumptions that the identified SNPs (i) were truly predictive of MASP-2 in study participants, (ii) were not associated with confounders that influenced both MASP-2 and VTE risk, and (iii) affected VTE risk only through their effects on MASP-2, as described elsewhere.[43]

RESULTS

The distribution of baseline characteristics of the study participants across quartiles of plasma MASP-2 is shown in Table 1. The mean age slightly decreased, while the mean BMI and the proportion of smokers slightly increased, with increasing quartiles of plasma MASP-2. The proportion of men was highest in the two upper quartiles. Predictably, the plasma levels of high-sensitivity CRP slightly increased across quartiles of MASP-2 from 1.43 ± 1.2 mg/L in the lowest quartile to 1.87 ± 1.6 mg/L in the highest quartile. The baseline characteristics of VTE cases and controls are shown in Supplementary Table 1. VTE patients had higher BMI and higher proportion with history of cancer than controls, whereas the proportion of smokers was somewhat lower in cases versus controls. The distribution of raw and normalized values of MASP-2 in cases and controls is shown in Supplementary Figure 1.

The characteristics of the VTE patients, measured at the time of the VTE event, are shown in Table 2. The mean age at the time of VTE was 67 years, and 49% were men. Out of the total VTE events, 62% were DVTs and 38% were PEs, and 42% of the cases were classified as unprovoked.

The ORs of VTE, DVT and PE across quartiles of plasma MASP-2 levels are shown in Table 3. For overall VTE, the OR increased across quartiles of plasma MASP-2 (*P* for trend = 0.01), with the exception of the second lowest quartile. Subjects with plasma MASP-2 ≥824 ng/mL had a 48% higher OR of VTE compared with those with MASP-2 <302 ng/mL (OR: 1.48; 95% CI: 1.06-2.06) in the model adjusted for age and sex. Plasma levels of MASP-2 were more strongly associated with risk of DVT (OR for upper vs. lower quartile: 1.83, 95% CI: 1.23-2.73) than with risk of PE (OR for upper vs. lower quartile: 1.04, 95% CI: 0.64-1.69). Further adjustment for BMI and CRP did not considerably influence the risk estimates (Table 3). The risk estimates for unprovoked events were essentially similar to those observed for overall VTE, DVT and PE (Supplementary Table 2). The addition of smoking as a covariate to the regression models did not virtually change the risk estimates for overall VTE and subgroups (data not shown).

The association between MASP-2 levels, entered as a continuous variable, and risk of VTE is depicted in Figure 2. In the analysis which included the full follow-up time (Figure 2A), the OR of VTE started to increase for MASP-2 levels above the 50th percentile, indicating that the 50th percentile cut-off could be appropriate for assessing VTE risk. However, when the follow-up time was restricted to <5 years from blood sampling to VTE diagnosis, a linear association throughout the continuum of MASP-2 levels was more prominent (Figure 2B). As depicted in Figure 2A and 2B, estimates of VTE risk were imprecise with wide 95% CIs at more extreme levels of MASP-2 due to the limited number of individuals in the analysis.

To consider the possibility of underestimating the true association due to regression dilution bias, we estimated ORs (highest vs. lowest quartile of MASP-2) of VTE and subgroups (DVT and PE) as a function of time between blood sampling and the events (Figure 3). The

313 ORs of overall VTE and DVT by high plasma MASP-2 were considerably higher with shortened
314 time between blood sampling and VTE. In contrast, no association was observed between
315 MASP-2 and PE over time (Figure 3).

316 The results of the pQTL analysis are described in Figure 4 and Supplementary Figure
317 2. The pQTL analysis revealed three SNPs that were significantly associated with MASP-2
318 plasma levels at the fixed genome-wide threshold of $p < 5x10^{-8}$. The identified SNPs
319 rs12711521 (minor allele frequency [MAF] = 0.153) and rs72550870 (MAF = 0.045) are
320 missense variants in exons of the MASP2 gene on chromosome 1, while rs2275527 (MAF =
321 0.220) is an exon-variant in the MTOR gene, which is a few genes away from *MASP2* (Figure
322 4B). The SNPs individually accounted for 25%, 17%, and 16% of the variance of MASP-2
323 levels, respectively. Together, these SNPs explained 39% of the variance of MASP-2 in the
324 model adjusted for age, sex, BMI, CRP and VTE status (adjusted r^2: 0.392, p-value < $2.2x10^{-16}$). The rs2275527 was linked with the rs12711521 with an r^2 of 0.44, and the MR analysis
326 was therefore performed with and without inclusion of rs2275527 for sensitivity. The
327 detailed information on the 3 SNPs used in the MR analysis is described in Supplementary
328 Table 3, along with the effect size estimates and standard errors for the SNP-exposure (i.e.
329 plasma MASP-2) association and the SNP-outcome (i.e. VTE) association obtained from the
330 pQTL and the INVENT consortium[41], respectively.

331 The inverse-variance weighted MR analysis revealed a statistically significant
332 association between genetically predicted MASP-2 and VTE. The forest plot of the MR
333 analysis with point estimates (log (odds ratio) per SD of MASP-2) and 95% CIs of causal effect
334 of MASP-2 levels on VTE for each of the three identified SNPs is shown in Figure 5. The OR of
335 VTE per 1 SD increase in genetically predicted MASP-2 was 1.03 (95% CI 1.01-1.05, p=

0.0011) (Figure 5A). Exclusion of rs2275527 showed essentially similar results (Figure 5B).

DISCUSSION

Plasma MASP-2 levels were associated with risk of future VTE in our nested case-control study derived from a population-based cohort. The risk of VTE increased across quartiles of plasma MASP-2 levels, and subjects with plasma MASP-2 levels in the highest quartile had a 48% higher risk of overall VTE and 83% higher risk of DVT compared to those with MASP-2 levels in the lowest quartile. The ORs for VTE and DVT by elevated plasma MASP-2 were substantially higher when the time between blood sampling and the VTE events was shorter. The risk estimates were modestly attenuated by further adjustments for BMI and CRP. Moreover, when MASP-2 was analyzed as a continuous variable, a linear association with VTE risk was noted, mainly in analyses restricted to the first 5 years of follow-up. In the pQTL analysis, we confirmed three previously identified genetic variants associated with plasma MASP-2 levels,[44,45] and estimated that these variants explained 39% of the variance of plasma MASP-2 levels. By applying an MR approach, genetically predicted MASP-2 levels were weakly, but significantly, associated with VTE risk. Thus, the present results suggest that plasma MASP-2 levels are genetically regulated and causally associated with the risk of VTE.

Although our observations are unchallenged, circumstantial evidence supports a role of the complement lectin pathway, and in particular MASP-2, in the pathogenesis of VTE. First, as thrombus formation originates in the valvular sinuses of the deep veins in a milieu characterized by severe hypoxia,[46,47] endothelial cells are exposed to oxidative stress which facilitates binding of MBL and/or other PRMs from the lectin pathway to the endothelial cell

surface with subsequent activation of MASP-2.[48-50] Accordingly, we recently reported that the risk of VTE increased with higher plasma MBL levels.[21] Second, *in vitro* studies have shown that activated MASP-2 can cleave prothrombin to thrombin with subsequent fibrin formation,[20,22,23] and several observational studies have shown that a high degree of coagulation activation is associated with the risk of future VTE.[51-54] Third, experimental studies in mouse models of arterial thrombosis have shown that inhibition of MASP-2, either by genetic deficiency or antibody neutralization, caused smaller myocardial infarct sizes[30,31] and less cerebral infarct volumes and neurological deficits.[28] Supporting these observations, the multiple interactions between the lectin complement pathway and the coagulation system, as well as key complement factors (e.g. C3, C5 and the terminal complement complex) that have been shown to associate with the risk of future VTE, are summarized in Figure 6 and Supplementary Table 4. Furthermore, elevated levels of C3, a central component of the complement system, are associated with cardiovascular risk factors, such as obesity, hypertension, and insulin resistance,[55,56] findings that strengthen the notion of a relationship between the complement system and CVD.

In our study, the association between plasma MASP-2 levels and VTE was entirely driven by the relationship between MASP-2 levels and risk of DVT. The explanation(s) for this observation is uncertain but could potentially be related to site of action and involved mechanisms. As the PRMs-MASP-2 complexes may bind to the endothelial surface and activate MASP-2 under hypoxic conditions,[47-50] valvular sinuses in the deep veins could be predilection sites for coagulation activation by MASP-2. Furthermore, hypercoagulable states associated with higher risk of DVT than PE, often referred to as the "Factor V Leiden (FVL) paradox",[57] has been explained by the formation of stable clots less susceptible for embolization. Indeed, an experimental study in mice reported that thrombi in FVL carriers

were larger and embolized less than in wild-type mice.[58] Accordingly, our observation of a preponderance of DVT over PE in those with high MASP-2 levels may suggest that MASP-2 activation promotes formation of thrombi less fragile to embolization.

A prerequisite for causal inference of the apparent association between plasma MASP-2 levels and VTE risk is a clear temporal sequence where the presence of the exposure, i.e. elevated MASP-2 level, occurs prior to the outcome, i.e. the VTE. In our study, the association between MASP-2 and VTE risk was demonstrated in a nested case-control study, a study design that would not be susceptible to reverse causation.[59] Although multivariable adjustments for potential confounders only modestly attenuated the association between MASP-2 and VTE risk, residual confounding cannot be ruled out due to the observational nature of the study.[60] MR analysis is a method designed to uncover causal relationships between exposure and outcome in observational studies.[61,62] MR exploits the fact that gene variants robustly associated with modifiable exposures are fixed at conception and follow Mendel's Laws for inheritance.[61,62] We identified two missense mutations in the MASP2 gene (rs12711521 and rs72550870) and one variant in an adjacent gene named *MTOR* (rs2275527). All three variants have previously been described,[44,45] and these gene variants explained 39% of the variation of plasma MASP-2 levels in our study. Even after excluding the variant in the adjacent gene (rs2275527), which is partly in LD with one of the missense mutations in *MASP2* (rs12711521), the remaining variants in the model explained an equal variation of plasma MASP-2 levels. Using summary data obtained from the INVENT consortium,[41] we found that one SD increase in genetically predicted MASP-2 showed a weak but significantly increased VTE risk, suggesting a causal relationship between MASP-2

and VTE risk. The MR estimates remained essentially the same upon exclusion of the variant in *MTOR*. Importantly, because rs12711521 and rs72550870 are missense variants in *MASP2*, the risk of biased MR estimates due to horizontal pleiotropy is low, as it is unlikely that these variants would influence VTE through a pathway other than plasma MASP-2. Of note, in previous genome-wide association analyses of VTE involving the GWAS summary data from the INVENT consortium[41] and the Million Veteran Program and UK Biobank[63] no signal at the *MASP2* locus was detected. In the present study, because the effect size of the association between the SNPs in *MASP-2* and VTE was modest at most (see Supplementary Table 3 for details), it might be speculated that such association did not reach a genome-wide significant level in the previous GWAS.[41,63] In light of the currently available data, future studies are warranted to confirm our findings from the MR analysis on MASP-2 and VTE.

Although our pQTL analysis confirmed that plasma MASP-2 levels are under a strong genetic regulation,[44,45,64] the OR for VTE according to high versus low MASP-2 levels increased with shorter time between blood sampling and VTE. This implies that biological fluctuations of plasma MASP-2 during the long follow-up resulted in underestimation of the true association, a phenomenon called regression dilution bias.[65,66] In analysis restricted to <5 years from blood sampling to VTE diagnosis, a linear association between MASP-2 levels and VTE risk was displayed throughout the continuum of MASP-2 levels, which reinforces the notion of a biological gradient between plasma MASP-2 levels and risk of VTE. Additionally, we observed a significant, albeit weak association between genetically predicted MASP-2 and VTE in our MR analysis, further strengthening the hypothesis of a causal relationship.

Strengths of this study include the temporal sequence of exposure and outcome in a sample recruited from the general adult population with validated VTE events and access to

exome sequencing data and measured plasma MASP-2 levels in the same population. The study also has limitations. Changes in MASP-2 levels during follow-up could result in underestimation of the OR, as indicated by the regression dilution plot showing higher ORs when analyses were restricted to the first years after follow-up. Blood samples were drawn in 1994-95 and stored at -80 °C for up to 22 years. The long storage time could potentially affect the plasma MASP-2 levels. However, plasma MASP-2 levels in our study population were similar to those in previous reports among healthy individuals and blood donors.[67,68] Additionally, as all samples were stored under the same conditions and for the same amount of time for cases and controls, the storage effect is assumed to be similar in the two groups, and any misclassification would be non-differential with regards to VTE status. Even though imputation expanded the investigation beyond the exome and allowed for the identification of variants in intergenic or intronic regions, the power to detect trans-acting pQTLs was limited, as previously described[35]. Finally, the limitations of an MR approach[61,62] should be considered when interpreting the results.

In conclusion, the current results indicate that high plasma MASP-2 levels are causally associated with risk of future VTE. Further studies are warranted to confirm our findings, and to unravel molecular mechanisms and explore potential targets for intervention.

Acknowledgements

Thrombosis Research Center (TREC) was supported by an independent grant from Stiftelsen Kristian Gerhard Jebsen (2014-2020).

Source of Funding

This work was supported by a former grant from Stiftelsen Kristian Gerhard Jebsen.

Conflict of Interest Disclosures

The authors declare no competing financial interests.

The INVENT Consortium provided the data for the Mendelian randomization analysis. A complete list of the members of the INVENT Consortium appears in the supplemental appendix.

Supplemental Material

Tables S1–S4

Figure S1-S2

References 69-75

Supplemental Appendix

REFERENCES

1. Naess IA, Christiansen SC, Romundstad P, Cannegieter SC, Rosendaal FR, Hammerstrøm J. Incidence and mortality of venous thrombosis: a population-based study. *J Thromb Haemost*. 2007;5(4):692-699.
2. Schulman S, Lindmarker P, Holmstrom M, et al. Post-thrombotic syndrome, recurrence, and death 10 years after the first episode of venous thromboembolism treated with warfarin for 6 weeks or 6 months. *J Thromb Haemost*. 2006;4(4):734-742.
3. Klok FA, van der Hulle T, den Exter PL, Lankeit M, Huisman MV, Konstantinides S. The post-PE syndrome: a new concept for chronic complications of pulmonary embolism. *Blood Rev*. 2014;28(6):221-226.
4. Gould MK, Garcia DA, Wren SM, et al. Prevention of VTE in nonorthopedic surgical patients: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012;141(2 Suppl):e227S-e277S.

5. Falck-Ytter Y, Francis CW, Johanson NA, et al. Prevention of VTE in orthopedic surgery patients: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012;141(2 Suppl):e278S-e325S.
6. Heit JA, Crusan DJ, Ashrani AA, Petterson TM, Bailey KR. Effect of a near-universal hospitalization-based prophylaxis regimen on annual number of venous thromboembolism events in the US. *Blood*. 2017;130(2):109-114.
7. Arshad N, Isaksen T, Hansen JB, Braekkan SK. Time trends in incidence rates of venous thromboembolism in a large cohort recruited from the general population. *Eur J Epidemiol*. 2017;32(4):299-305.
8. Munster AM, Rasmussen TB, Falstie-Jensen AM, et al. A changing landscape: Temporal trends in incidence and characteristics of patients hospitalized with venous thromboembolism 2006-2015. *Thromb Res*. 2019;176:46-53.
9. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.
10. Ng M, Fleming T, Robinson M, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2014;384(9945):766-781.
11. Stein PD, Hull RD, Kayali F, Ghali WA, Alshab AK, Olson RE. Venous thromboembolism according to age: the impact of an aging population. *Arch Intern Med*. 2004;164(20):2260-2265.
12. Foley JH, Conway EM. Cross Talk Pathways Between Coagulation and Inflammation. *Circ Res*. 2016;118(9):1392-1408.
13. Markiewski MM, Nilsson B, Ekdahl KN, Mollnes TE, Lambris JD. Complement and coagulation: strangers or partners in crime? *Trends Immunol*. 2007;28(4):184-192.
14. Norgaard I, Nielsen SF, Nordestgaard BG. Complement C3 and High Risk of Venous Thromboembolism: 80517 Individuals from the Copenhagen General Population Study. *Clin Chem*. 2016;62(3):525-534.
15. Subramaniam S, Jurk K, Hobohm L, et al. Distinct contributions of complement factors to platelet activation and fibrin formation in venous thrombus development. *Blood*. 2017;129(16):2291-2302.
16. Mollnes TE, Lea T, Froland SS, Harboe M. Quantification of the terminal complement complex in human plasma by an enzyme-linked immunosorbent assay based on monoclonal antibodies against a neoantigen of the complex. *Scand J Immunol*. 1985;22(2):197-202.
17. Harboe M, Thorgersen EB, Mollnes TE. Advances in assay of complement function and activation. *Adv Drug Deliv Rev*. 2011;63(12):976-987.
18. Hoiland, II, Liang RA, Braekkan SK, et al. Complement activation assessed by the plasma terminal complement complex and future risk of venous thromboembolism. *J Thromb Haemost*. 2019;17(6):934-943.
19. Garred P, Genster N, Pilely K, et al. A journey through the lectin pathway of complement-MBL and beyond. *Immunol Rev*. 2016;274(1):74-97.
20. Kozarcanin H, Lood C, Munthe-Fog L, et al. The lectin complement pathway serine proteases (MASPs) represent a possible crossroad between the coagulation and complement systems in thromboinflammation. *J Thromb Haemost*. 2016;14(3):531-545.
21. Liang RA, Hoiland, II, Ueland T, et al. Plasma levels of mannose-binding lectin and future risk of venous thromboembolism. *J Thromb Haemost*. 2019;17(10):1661-1669.

22. Gulla KC, Gupta K, Krarup A, et al. Activation of mannan-binding lectin-associated serine proteases leads to generation of a fibrin clot. *Immunology*. 2010;129(4):482-495.
23. Krarup A, Wallis R, Presanis JS, Gal P, Sim RB. Simultaneous activation of complement and coagulation by MBL-associated serine protease 2. *PLoS One*. 2007;2(7):e623.
24. Jenny L, Dobo J, Gal P, Schroeder V. MASP-1 Induced Clotting--The First Model of Prothrombin Activation by MASP-1. *PLoS One*. 2015;10(12):e0144633.
25. Conway EM. Reincarnation of ancient links between coagulation and complement. *Journal of Thrombosis and Haemostasis*. 2015;13:S121-S132.
26. Mackman N. New insights into the mechanisms of venous thrombosis. *Journal of Clinical Investigation*. 2012;122(7):2331-2336.
27. Tsakanova G, Stepanyan A, Nahapetyan K, Sim RB, Arakelyan A, Boyajyan A. Serine proteases of the complement lectin pathway and their genetic variations in ischaemic stroke. *J Clin Pathol*. 2018;71(2):141-147.
28. Orsini F, Chrysanthou E, Dudler T, et al. Mannan binding lectin-associated serine protease-2 (MASP-2) critically contributes to post-ischemic brain injury independent of MASP-1. *J Neuroinflammation*. 2016;13(1):213.
29. Orsini F, Villa P, Parrella S, et al. Targeting mannose-binding lectin confers long-lasting protection with a surprisingly wide therapeutic window in cerebral ischemia. *Circulation*. 2012;126(12):1484-1494.
30. Schwaeble WJ, Lynch NJ, Clark JE, et al. Targeting of mannan-binding lectin-associated serine protease-2 confers protection from myocardial and gastrointestinal ischemia/reperfusion injury. *Proc Natl Acad Sci U S A*. 2011;108(18):7523-7528.
31. Clark JE, Dudler T, Marber MS, Schwaeble W. Cardioprotection by an anti-MASP-2 antibody in a murine model of myocardial infarction. *Open Heart*. 2018;5(1):e000652.
32. Jacobsen BK, Eggen AE, Mathiesen EB, Wilsgaard T, Njolstad I. Cohort profile: the Tromso Study. *Int J Epidemiol*. 2012;41(4):961-967.
33. Braekkan SK, Borch KH, Mathiesen EB, Njolstad I, Wilsgaard T, Hansen JB. Body height and risk of venous thromboembolism: The Tromso Study. *Am J Epidemiol*. 2010;171(10):1109-1115.
34. Carson AR, Smith EN, Matsui H, et al. Effective filtering strategies to improve data quality from population-based whole exome sequencing studies. *Bmc Bioinformatics*. 2014;15.
35. Solomon T, Smith EN, Matsui H, et al. Associations Between Common and Rare Exonic Genetic Variants and Serum Levels of 20 Cardiovascular-Related Proteins The Tromso Study. *Circulation-Cardiovascular Genetics*. 2016;9(4):375-383.
36. Browning BL, Browning SR. Genotype Imputation with Millions of Reference Samples. *Am J Hum Genet*. 2016;98(1):116-126.
37. Genomes Project C, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56-65.
38. Pearce N. Analysis of matched case-control studies. *BMJ*. 2016;352:i969.
39. EPACTS 3.3.0 edn Arbor A. *MI: University of Michigan, Center for Statistical Genetics* 2016.
40. Kang HM, Sul JH, Service SK, et al. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet*. 2010;42(4):348-354.
41. Lindstrom S, Wang L, Smith EN, et al. Genomic and transcriptomic association studies identify 16 novel susceptibility loci for venous thromboembolism. *Blood*. 2019;134(19):1645-1657.

42. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet*. 2018;27(R2):R195-R208.
43. Didelez VaS, N. . Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res*. 2007;16:309-330.
44. Silva AA, Catarino SJ, Boldt ABW, Pedroso MLA, Beltrame MH, Messias-Reason IJ. Effects of MASP2 haplotypes and MASP-2 levels in hepatitis C-infected patients. *Int J Immunogenet*. 2018;45(3):118-127.
45. Boldt AB, Goeldner I, Stahlke ER, Thiel S, Jensenius JC, de Messias-Reason IJ. Leprosy association with low MASP-2 levels generated by MASP2 haplotypes and polymorphisms flanking MAp19 exon 5. *PLoS One*. 2013;8(7):e69054.
46. Reitsma PH, Versteeg HH, Middeldorp S. Mechanistic view of risk factors for venous thromboembolism. *Arterioscler Thromb Vasc Biol*. 2012;32(3):563-568.
47. Bovill EG, van der Vliet A. Venous valvular stasis-associated hypoxia and thrombosis: what is the link? *Annu Rev Physiol*. 2011;73:527-545.
48. Collard CD, Montalto MC, Reenstra WR, Buras JA, Stahl GL. Endothelial oxidative stress activates the lectin complement pathway: role of cytokeratin 1. *Am J Pathol*. 2001;159(3):1045-1054.
49. Kataoka H, Hamilton JR, McKemy DD, et al. Protease-activated receptors 1 and 4 mediate thrombin signaling in endothelial cells. *Blood*. 2003;102(9):3224-3231.
50. Durigutto P, Macor P, Pozzi N, et al. Complement Activation and Thrombin Generation by MBL Bound to beta2-Glycoprotein I. *J Immunol*. 2020;205(5):1385-1392.
51. Cushman M, Folsom AR, Wang L, et al. Fibrin fragment D-dimer and the risk of future venous thrombosis. *Blood*. 2003;101(4):1243-1248.
52. Folsom AR, Alonso A, George KM, Roetker NS, Tang W, Cushman M. Prospective study of plasma D-dimer and incident venous thromboembolism: The Atherosclerosis Risk in Communities (ARIC) Study. *Thromb Res*. 2015;136(4):781-785.
53. Puurunen MK, Enserro D, Xanthakis V, et al. Biomarkers for the prediction of venous thromboembolism in the community. *Thromb Res*. 2016;145:34-39.
54. Evensen LH, Folsom AR, Pankow JS, et al. Hemostatic factors, inflammatory markers, and risk of incident venous thromboembolism: The Multi-Ethnic Study of Atherosclerosis. *J Thromb Haemost*. 2021.
55. Hernandez-Mijares A, Jarabo-Bueno MM, Lopez-Ruiz A, Sola-Izquierdo E, Morillas-Arino C, Martinez-Triguero ML. Levels of C3 in patients with severe, morbid and extreme obesity: its relationship to insulin resistance and different cardiovascular risk factors. *Int J Obes (Lond)*. 2007;31(6):927-932.
56. Torres T, Bettencourt N, Mendonca D, Vasconcelos C, Silva BM, Selores M. Complement C3 as a marker of cardiometabolic risk in psoriasis. *Archives of Dermatological Research*. 2014;306(7):653-660.
57. van Langevelde K, Flinterman LE, van Hylckama Vlieg A, Rosendaal FR, Cannegieter SC. Broadening the factor V Leiden paradox: pulmonary embolism and deep-vein thrombosis as 2 sides of the spectrum. *Blood*. 2012;120(5):933-946.
58. Shaya SA, Westrick RJ, Gross PL. Thrombus stability explains the factor V Leiden paradox: a mouse model. *Blood Adv*. 2019;3(21):3375-3378.
59. Smith GD, Ebrahim S. Data dredging, bias, or confounding. *BMJ*. 2002;325(7378):1437-1438.

60. Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol*. 2007;166(6):646-655.
61. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89-98.
62. Zheng J, Baird D, Borges MC, et al. Recent Developments in Mendelian Randomization Studies. *Curr Epidemiol Rep*. 2017;4(4):330-345.
63. Klarin D, Busenkell E, Judy R, et al. Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. *Nat Genet*. 2019;51(11):1574-1579.
64. Boldt AB, Grisbach C, Steffensen R, et al. Multiplex sequence-specific polymerase chain reaction reveals new MASP2 haplotypes associated with MASP-2 and MAp19 serum levels. *Hum Immunol*. 2011;72(9):753-760.
65. Clarke R, Shipley M, Lewington S, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol*. 1999;150(4):341-353.
66. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. *BMJ*. 2010;340:c2289.
67. Moller-Kristensen M, Jensenius JC, Jensen L, et al. Levels of mannan-binding lectin-associated serine protease-2 in healthy individuals. *J Immunol Methods*. 2003;282(1-2):159-167.
68. Ytting H, Christensen IJ, Thiel S, Jensenius JC, Nielsen HJ. Serum mannan-binding lectin-associated serine protease 2 levels in colorectal cancer: relation to recurrence and mortality. *Clin Cancer Res*. 2005;11(4):1441-1446.

Supplemental Reference

69. Norgaard I, Nielsen SF, Nordestgaard BG. Complement C3 and High Risk of Venous Thromboembolism: 80517 Individuals from the Copenhagen General Population Study. *Clinical Chemistry*. 2016;62(3):525-534.
70. Hoiland, II, Liang RA, Hindberg K, et al. Associations between complement pathways activity, mannose-binding lectin, and odds of unprovoked venous thromboembolism. *Thromb Res*. 2018;169:50-56.
71. Hoiland II, Liang RA, Braekkan SK, et al. Complement activation assessed by the plasma terminal complement complex and future risk of venous thromboembolism. *Journal of Thrombosis and Haemostasis*. 2019;17(6):934-943.
72. Liang RA, Hoiland II, Ueland T, et al. Plasma levels of mannose-binding lectin and future risk of venous thromboembolism. *Journal of Thrombosis and Haemostasis*. 2019;17(10):1661-1669.
73. Skjeflo EW, Braekkan SK, Ludviksen JK, et al. Elevated plasma concentration of complement factor C5 is associated with risk of future venous thromboembolism. *Blood*. 2021;138(21):2129-2137.
74. Foley JH, Walton BL, Aleman MM, et al. Complement Activation in Arterial and Venous Thrombosis is Mediated by Plasmin. *EBioMedicine*. 2016;5:175-182.

75. Subramaniam S, Jurk K, Hobohm L, et al. Distinct contributions of complement factors to platelet activation and fibrin formation in venous thrombus development. *Blood*. 2017;129(16):2291-2302.

Highlights

- In a population-based nested case-control study derived from the Tromsø cohort, high plasma MASP-2 levels are associated with increased risk of future incident venous thromboembolism (VTE)
- According to protein quantitative trait loci analysis**,** plasma levels of MASP-2 are genetically regulated
- Mendelian randomization suggests that the association between MASP-2 and VTE is causal

Tables

Table 1 Distribution of baseline characteristics across quartiles of plasma levels of mannose-binding lectin-associated serine protein 2 (MASP-2).

	Q1 (<302 ng/mL)	**Q2 (302-549 ng/mL)**	**Q3 (550-823 ng/mL)**	**Q4 (≥824ng/mL)**
n	297	296	315	344
Age, years (±SD)	62.4 ± 14.3	61.4 ± 13.8	59.1 ± 13.7	58.4 ± 13.0
Sex, % men (n)	41.4 (123)	46.0 (136)	51.1 (161)	49.4 (170)
BMI, Kg/m^2	25.9 ± 4.0	26.3 ± 4.2	26.5 ± 4.2	26.9 ± 4.5
Smoking, % (n)	28.3 (84)	30.1 (89)	31.8 (100)	34.3 (118)
*hs*CRP, mg/L (±SD)	1.43 ± 1.2	1.47 ± 1.3	1.66 ± 1.3	1.87 ± 1.6
WBC, 10^9/L (±SD)	6.89 ± 1.9	6.95 ± 3.1	7.09 ± 1.8	7.09 ± 2.0
CVD, % (n)*	16.5 (49)	18.9 (56)	13.7 (43)	13.7 (47)
Cancer, % (n)*	5.4 (16)	5.4 (16)	4.4 (14)	3.2 (11)
Diabetes, % (n)*	4.7 (14)	3.0 (9)	3.5 (11)	4.1 (14)

BMI: Body mass index; CVD: Cardiovascular disease (history of myocardial infarction, stroke, angina pectoris), hsCRP: C-reactive protein measured by a high sensitive technique; WBC: White blood cell count; SD: Standard deviation.

*Self-reported history of CVD, cancer or diabetes at baseline.

Table 2 Characteristics of patients at VTE diagnosis (n=410). Values are % (n) or means ± 1 SD.

	% (*n*)
Age at VTE (years)	67.4 ± 13.6
Sex (males)	48.5 (199)
Deep vein thrombosis	61.7 (253)
Pulmonary Embolism	38.3 (157)
Unprovoked VTE	42.0 (172)
Provoked VTE	58.0 (238)
Surgery/trauma	22.4 (92)
Cancer	21.7 (89)
Immobilization	17.8 (73)
Acute medical condition	15.6 (64)
Other factors	3.9 (16)

Table 3 Odds ratios (OR) with 95% confidence intervals (CI) for venous thromboembolism (VTE), deep vein thrombosis (DVT) and pulmonary embolism (PE) according to quartiles of plasma levels of mannose-binding lectin-associated serine protein 2 (MASP-2).

Quartiles of MASP-2 (ng/mL)	Controls	Cases	Model 1 OR (95% CI)	Model 2 OR (95% CI)
Overall VTE				
<302	209	88	Ref.	Ref.
302-549	210	86	0.97 (0.68-1.38)	0.94 (0.66-1.35)
550-823	211	104	1.17 (0.83-1.65)	1.10 (0.78-1.56)
≥824	212	132	1.48 (1.06-2.06)	1.36 (0.97-1.91)
P for trend			*0.01*	*0.04*
DVT				
<302	209	49	Ref.	Ref.
302-549	210	51	1.04 (0.67-1.60)	1.01 (0.65-1.57)
550-823	211	62	1.25 (0.82-1.91)	1.20 (0.78-1.84)
≥824	212	91	1.83 (1.23-2.73)	1.72 (1.14-2.58)
P for trend			*0.001*	*0.004*
PE				
<302	209	39	Ref.	Ref.
302-549	210	35	0.89 (0.54-1.46)	0.85 (0.52-1.41)
550-823	211	42	1.07 (0.66-1.73)	0.99 (0.61-1.61)
≥824	212	41	1.04 (0.64-1.69)	0.94 (0.57-1.53)
P for trend			*0.7*	*1.0*

Model 1: adjusted for age and sex.

Model 2: adjusted for age, sex, body mass index and C-reactive protein.

Figure legends

Figure 1. Flowchart of the study population. The flowchart illustrates the nested case-control study derived from the fourth survey of the Tromsø Study (1994-1995). VTE, venous thromboembolism.

Figure 2. Odds ratios (OR) of venous thromboembolism (VTE) as a function of MASP-2 plasma levels adjusted for age, sex, body mass index and C-reactive protein in a generalized additive regression model. Panel **A** shows the results for the full follow-up, while panel **B** shows the results of analyses restricted to the first 5 years of follow-up. The solid lines show ORs surrounded by shaded areas showing 95% confidence intervals. The distributions of MASP-2 plasma levels are shown as density plots (light grey) at the bottom and white vertical lines indicate quartile cutoff.

Figure 3. Plots of estimated odds ratios (OR) of venous thromboembolism (VTE, panel A), deep vein thrombosis (DVT, panel B) and pulmonary embolism (PE, panel C) as a function of maximum time from blood sampling in Tromsø 4 (1994-1995) to events in analyses adjusted for age, sex, body mass index and C-reactive protein. Subjects with plasma mannose binding lectin associated serine protease-2 (MASP-2) in the highest quartile (Q4) were compared to those with MASP-2 levels in the lowest quartile (Q1, reference category). The number of VTE, DVT and PE events are depicted above the plot. Large, solid circles indicate ORs with *P* values <0.05.

Figure 4. Protein quantitative trait loci (pQTL) analysis results. Panel 4A shows the Manhattan plot of pQTL analysis (GRCh37/hg19 was used as reference human genome). The upper, dashed line indicates the $5x10^{-8}$ *P* value significance threshold. The purple triangles indicate complement-related genes. In the genome-wide plot, the blue dots are the cis-region around *MASP2*. In the inserted cis-region plot, those above the genome-wide threshold are marked in blue. Panel 4B shows the regional plot for the associated region near *MASP2* on Chr1, with r^2 value for the linkage disequilibrium of variants. rs12711521 and rs72550870 were independent missense single nucleotide polymorphisms in the *MASP2* exons.

Figure 5. Forest plot of the Mendelian randomization (MR) analysis. Forest plot of the MR analysis with point estimates (log(odds ratio) per SD of MASP-2) and 95% confidence intervals of causal effect of plasma MASP-2 levels on VTE for each single nucleotide polymorphism (SNP) and collectively (i.e. inverse variance-weighted analysis) in regression analyses. Panel 5A shows the forest plot of MR analysis with the inclusion of all 3 SNPs (rs12711521 and rs72550870 in *MASP2*, and rs2275527 in *MTOR)*. Panel 5B shows the forest plot of MR analysis with the exclusion of the SNP in *MTOR* (rs2275527).

Figure 6. A simplified overview of multiple interactions between factors of the lectin pathway as well as central factors of the complement system with the coagulation system. The complement factors that have been shown to associate with risk of venous thromboembolism are colored with gray shades (deeper shade for the lectin pathway factors and lighter shade for the factors of the common pathway) and are also summarized in Supplementary Table 4. Gray arrows indicate activation of coagulation factors by lectin pathway factors. MBL, mannose-binding lectin; MASP, mannose binding lectin associated serine protease.

SUPPLEMENTAL MATERIALS

High levels of complement activating enzyme MASP-2 are associated with risk of future incident venous thromboembolism

Short tittle: MASP-2 and venous thromboembolism

Christabel Esi Damoah,[1] Omri Snir,[1] Kristian Hindberg,[1] Peter Garred,[2] Judith K. Ludviksen,[3] Sigrid K. Brækkan,[1,4] Vânia M. Morelli,[1,4] Tom Eirik Mollnes,[1,3,5,6] John-Bjarne Hansen,[1,4] and the INVENT Consortium

[1]Thrombosis Research Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway; [2]Laboratory of Molecular Medicine, Department of Clinical Immunology, Section 7631, Rigshospitalet, Copenhagen, Denmark; [3]Reaserch Laboratory, Nordland Hospital, Bodø, Norway; [4]Division of Internal Medicine, University Hospital of North Norway, Tromsø, Norway; [5]Department of Immunology, Oslo University Hospital and University of Oslo, Norway; [6]Centre of Molecular Inflammation Research, Norwegian University of Science and Technology, Trondheim, Norway.

Corresponding author: Christabel Esi Damoah, Thrombosis Research Center (TREC), Department of Clinical Medicine, UiT- The Arctic University of Norway, N-9037 Tromsø, Norway. Tel: +47 77625105. Email: christabel.e.damoah@uit.no

Supplemental tables

Supplementary Table 1. Distribution of baseline characteristics in venous thromboembolism (VTE) cases and controls

Supplementary Table 2. Odds ratios (OR) with 95% confidence intervals (CI) for unprovoked venous thromboembolic events and unprovoked events in VTE subgroups (DVT and PE) according to quartiles of plasma levels of mannose-binding lectin associated serine protein 2 (MASP-2)

Supplemental excel file

Supplementary Table 3. The association of single nucleotide polymorphisms (SNPs) with plasma MASP-2 and venous thromboembolism

Supplementary Table 4. Epidemiological and experimental studies showing key lectin pathway /complement factors and their association with venous thromboembolism (VTE)

Table S1. Distribution of baseline characteristics in venous thromboembolism (VTE) cases and controls

Variables	**VTE Cases (n= 410)**	**Controls (n= 842)**
MASP-2 (ng/mL)	616.7 ± 403.6	568.8 ± 395.7
Age, years (±SD)	60.2 ± 13.8	60.3 ± 13.8
Sex, % men (n)	48.5 (199)	46.4 (391)
BMI, Kg/m^2	27.2 ± 4.5	26.1 ± 4.1
Smoking, % (n)	29.5 (121)	32.1 (270)
hsCRP, mg/L (±SD)	1.72 ± 1.37	1.58 ± 1.36
WBC, 10^9/L (±SD)	7.19 ± 2.92	6.95 ± 1.79
CVD, % (n)*	15.9 (65)	15.4 (130)
Cancer, % (n)*	6.3 (26)	3.7 (31)
Diabetes, % (n)*	3.9 (16)	3.8 (32)

BMI: Body mass index; CVD: Cardiovascular disease (history of myocardial infarction, stroke, angina pectoris); hsCRP: C-reactive protein measured by a high sensitive technique; MASP-2: mannose-binding lectin-associated serine protein 2; WBC: White blood cell count; SD: Standard deviation.
*Self-reported history of diabetes, cancer or CVD at baseline.

Table S2. Odds ratios (OR) with 95% confidence intervals (CI) for unprovoked venous thromboembolic events and unprovoked events in VTE subgroups (DVT and PE) according to quartiles of plasma levels of mannose-binding lectin associated serine protein 2 (MASP-2)

Quartiles of MASP2 (ng/mL)	Controls	Cases	Model 1 OR (95% CI)	Model 2 OR (95% CI)
Unprovoked VTE				
<302	209	38	Ref.	Ref.
302-549	210	36	0.93 (0.57-1.52)	0.90 (0.54-1.48)
550-823	211	41	1.04 (0.64-1.68)	0.97 (0.59-1.58)
≥824	212	57	1.43 (0.91-2.26)	1.29 (0.81-2.05)
P for trend			0.09	0.2
Unprovoked DVT				
<302	207	19	Ref.	Ref.
302-549	211	17	0.87 (0.44-1.73)	0.85 (0.43-1.70)
550-823	210	24	1.21 (0.64-2.29)	1.15 (0.60-2.17)
≥824	214	40	2.00 (1.11-3.59)	1.84 (1.02-3.32)
P for trend			0.007	0.02
Unprovoked PE				
<302	207	19	Ref.	Ref.
302-549	211	19	0.98 (0.50-1.91)	0.92 (0.47-1.81)
550-823	210	17	0.86 (0.43-1.71)	0.79 (0.39-1.58)
≥824	214	17	0.86 (0.43-1.70)	0.75 (0.37-1.51)
P for trend			0.6	0.4

DVT, deep vein thrombosis; PE, pulmonary embolism; VTE, venous thromboembolism.

Model 1: adjusted for age and sex.

Model 2: adjusted for age, sex, body mass index and C-reactive protein.

Table S4. Epidemiological and experimental studies showing key lectin pathway /complement factors and their association with venous thromboembolism (VTE)

First author, Year of publication	Study Design and Population	Complement pathways/factors	Main findings
Nørgaard *et al*., (2016) [69]	Cohort study in Danish population, with 80,517 participants, of whom 1176 developed VTE during follow-up	C3	High levels of complement C3 were associated with risk of future development of VTE
Høiland *et al.,* (2018) [70]	Case-control study (Tromsø study), with 24 unprovoked VTE cases and 24 age- and sex-matched controls	Classical, lectin and alternative pathways	High activity in the classical pathway and MBL-deficiency were associated with increased odds of unprovoked VTE
Høiland *et al.,* (2019) [71]	Nested case-control study (Tromsø study), with 415 VTE cases and 848 age- and sex-matched controls	Terminal complement complex (TCC)	High levels of plasma TCC were associated with risk of VTE, particularly unprovoked events
Liang *et al.,* (2019) [72]	Nested case-control study (Tromsø study), with 417 VTE patients and 849 age-matched and sex-matched controls	Mannose binding lectin (MBL)	Low plasma MBL levels were associated with reduced risk of VTE and deep vein thrombosis
Skjeflo *et al.,* (2021) [73]	Nested case-control study (Tromsø study), with 415 VTE patients and 848 age- and sex-matched controls	C5	High levels of C5 were associated with increased risk of VTE, particularly unprovoked events
First authors (Year)	**Murine model of VTE**	**Complement factors studied**	**Main findings**
Foley *et al.,* (2016) [74]	Inferior vena cava stasis model	C3a and C5a	Clot weight strongly correlate with C5a
Subramaniam. *et al.,* (2017) [75]	Inferior vena cava stenosis model	C3 and C5	C3 deficiency resulted in reduced thrombus incidence and size; C5 deficiency resulted in reduced thrombus stability

Supplemental figures

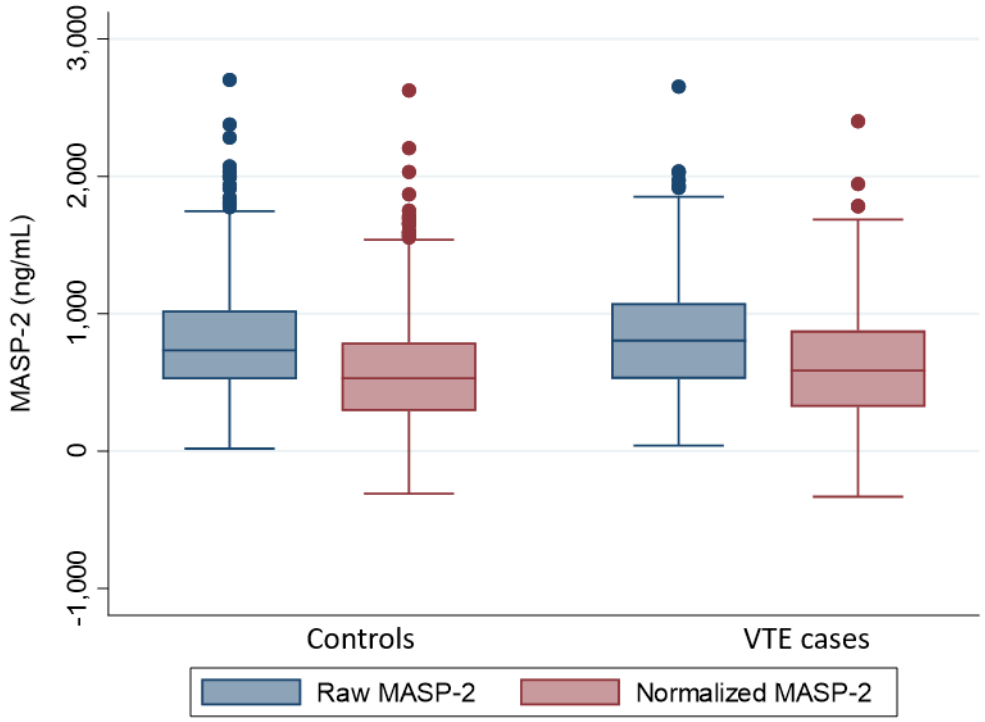

Figure S1. Boxplots of plasma MASP-2 levels in controls and venous thromboembolism (VTE) cases. The boxplots represent the raw (blue) and normalized (red) distribution of MASP-2 levels in controls (n =842) and VTE cases (n = 410). Boxplots display the minimum, the maximum, the median and the 25th and 75th percentiles of MASP-2 levels, with outliers plotted as individual points. Each sample was normalized with respect to a commercially available control provided by the manufacturer, and the global median of MASP-2 levels derived from all control subjects of the nested case-control study.

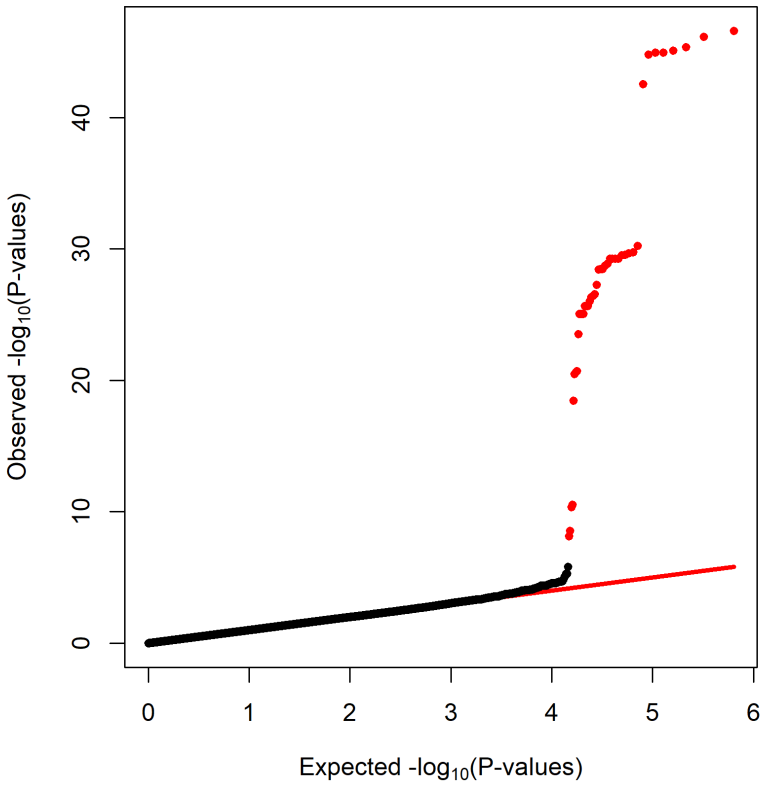

Figure S2. Quantile-quantile (Q-Q) plot of observed vs expected $-\log_{10}$(*P*-value) for MASP-2 protein quantitative trait loci (pQTL) analysis. The observed *P*-values < 5×10^{-8} are shown in red.

Major Resources Table

Animals (in vivo studies)

Species	Vendor or Source	Background Strain	Sex	Persistent ID / URL
NA	NA	NA	NA	NA

Genetically Modified Animals

	Species	Vendor or Source	Background Strain	Other Information	Persistent ID / URL
Parent - Male	NA	NA	NA	NA	NA
Parent - Female	NA	NA	NA	NA	NA

Antibodies

Target antigen	Vendor or Source	Catalog #	Working concentration	Lot # (preferred but not required)	Persistent ID / URL
NA	NA	NA	NA	NA	NA

DNA/cDNA Clones

Clone Name	Sequence	Source / Repository	Persistent ID / URL
NA	NA	NA	NA

Cultured Cells

Name	Vendor or Source	Sex (F, M, or unknown)	Persistent ID / URL
NA	NA	NA	NA

Data & Code Availability

Description	Source / Repository	Persistent ID / URL
The Data that support the findings of this study are available from the corresponding author upon reasonable request		

Other

Description	Source / Repository	Persistent ID / URL
NA	NA	NA

DOI [to be added]

NA, not applicable

Supplemental appendix

Members of the INVENT consortium

Sara Lindstrom, PhD,[1] Lu Wang, PhD,[2] Erin N. Smith, PhD,[3] William Gordon, MS,[4] Astrid van Hylckama Vlieg, PhD,[5] Mariza de Andrade, PhD,[6] Jennifer A. Brody, BA,[7] Jack W. Pattee, BA,[8] Jeffrey Haessler, MS,[9] Ben M. Brumpton, PhD, MPH,[10] Daniel I. Chasman, PhD,[11] Pierre Suchon, MD-PhD,[12] Ming-Huei Chen, PhD,[13] Constance Turman, MS,[14] Marine Germain,[15] Kerri L. Wiggins, MS, RD,[16] James MacDonald, MS,[17] Sigrid K. Braekkan, PhD,[18] Sebastian M. Armasu, MS,[19] Nathan Pankratz, PhD,[20] Rabecca D. Jackson, MD,[21] Jonas B. Nielsen, MD, PhD,[22] Franco Giulianini, PhD,[23] Marja K. Puurunen, MD, PhD,[24] Manal Ibrahim, MD,[25] Susan R. Heckbert, MD, PhD,[26] Theo K. Bammler, PhD,[27] Kelly A. Frazer, PhD,[28] Bryan M. McCauley, MS,[29] Kent Taylor, PhD,[30] James S. Pankow, PhD, MPH,[31] Alexander P. Reiner, MD, MPH,[32] Maiken E. Gabrielsen, PhD,[33] Jean-François Deleuze, PhD,[34] Chris J. O'Donnell, MD,[35] Jihye Kim, PhD, MPH,[36] Barbara McKnight, PhD,[37] Peter Kraft, PhD,[38] John-Bjarne Hansen, MD, PhD,[39] Frits R. Rosendaal, MD, PhD,[40] John A. Heit, MD,[41] Bruce M. Psaty, MD, PhD,[42] Weihong Tang, MD, PhD,[43] Charles Kooperberg, PhD,[44] Kristian Hveem, MD, PhD,[45] Paul M. Ridker, MD, MPH,[46] Pierre-Emmanuel Morange, MD-PhD,[47] Andrew D. Johnson, PhD,[48] Christopher Kabrhel, MD MPH,[49] David-AlexandreTrégouët, PhD,[50] Nicholas L. Smith, PhD[51]

Affiliations

[1]Department of Epidemiology, University of Washington, Seattle, Washington, USA; Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

[2]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA.

[3]Department of Pediatrics and Rady Children's Hospital, University of California San Diego, La Jolla, USA; K.G. Jebsen Thrombosis Research and Expertise Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway.

[4]Department of Epidemiology, University of Washington, Seattle, WA, USA.

[5]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. [6]Department of Health Sciences Research, Mayo Clinic, Rochester, MN USA.

[7]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle WA USA.

[8]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN USA. [9]Division of Public Health, Fred Hutchinson Cancer Research Center, Seattle WA, USA.

[10]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway; MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK; Clinic of Thoracic and Occupational Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway.

[11]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, USA; Harvard Medical School, Boston, USA.

[12]Laboratory of Haematology, La Timone Hospital, Marseille, France; C2VN, Aix Marseille University, INSERM, INRA, C2VN, Marseille, France.

[13]Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD, USA; NHLBI and Boston University's The Framingham Heart Study, Framingham, MA, USA.

[14]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

[15]INSERM UMR_S 1219, Bordeaux Population Health Research Center, University of Bordeaux, France.

[16]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle WA USA.

[17]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA.

[18]K.G. Jebsen Thrombosis Research and Expertise Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway; Division of Internal Medicine, University Hospital of North Norway, Tromsø, Norway.

[19]Health Sciences Research, Mayo Clinic, Rochester, MN USA.

[20]Department of Laboratory Medicine and Pathology, School of Medicine, University of Minnesota, Minneapolis, MN, USA.

[21]Division of Endocrinology, Diabetes and Metabolism, The Ohio State University, Columbus OH, USA.

[22]Department of Internal Medicine, Division of Cardiology, University of Michigan Medical School, Ann Arbor, Michigan, USA.

[23]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, USA.

[24]NHLBI and Boston University's The Framingham Heart Study, Framingham, MA, USA.

[25]Laboratory of Haematology, La Timone Hospital, Marseille, France.; C2VN, Aix Marseille University, INSERM, INRA, C2VN, Marseille, France.

[26]Department of Epidemiology, University of Washington, Seattle, Washington, USA; Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle WA USA. [27]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA.

[28]Department of Pediatrics and Rady Children's Hospital, University of California San Diego, La Jolla, USA; K.G. Jebsen Thrombosis Research and Expertise Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway; Institute of Genomic Medicine, University of California San Diego, La Jolla, California, USA.

[29]Health Sciences Research, Mayo Clinic, Rochester, MN USA.

[30]Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor-UCLA Medical Center, Torrence CA 90502, USA.

[31]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA.

[32]Department of Epidemiology, University of Washington, Seattle WA, United States; Division of Public Health, Fred Hutchinson Cancer Research Center, Seattle WA, United States.

[33]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway.

[34]Centre National de Recherche en Génomique Humaine, Direction de la Recherche Fondamentale, CEA, 91057 Evry, France; CEPH, Fondation Jean Dausset, Paris, France.

[35]Million Veteran's Program, Veteran's Administration, Boston, MA; Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD, USA; NHLBI and Boston University's The Framingham Heart Study, Framingham, MA, USA.

[36]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

[37]Department of Biostatistics, University of Washington, Seattle WA USA; Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

[38]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

[39]K.G. Jebsen Thrombosis Research and Expertise Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway; Division of Internal Medicine, University Hospital of North Norway, Tromsø, Norway.

[40]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. [41]Health Sciences Research, Mayo Clinic, Rochester, MN USA.

[42]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle WA USA; Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle WA USA.

[43]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA.

[44]Division of Public Health, Fred Hutchinson Cancer Research Center, Seattle WA, United States. [45]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway.

[46]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, USA; Harvard Medical School, Boston, USA.

[47]C2VN, Aix Marseille Univ, INSERM, INRA, C2VN, Marseille, France; Laboratory of Haematology, La Timone Hospital, Marseille, France; CRB Assistance Publique - Hôpitaux de Marseille, HemoVasc (CRB AP-HM HemoVasc), Marseille, France.

[48]Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD, USA; NHLBI and Boston University's The Framingham Heart Study, Framingham, MA, USA.

[49]Center for Vascular Emergencies, Department of Emergency Medicine, Massachusetts General Hospital; Channing Division of Network Medicine, Brigham and Women's Hospital; Harvard Medical School.

[50]INSERM UMR_S 1219, Bordeaux Population Health Research Center, University of Bordeaux, France.

[51]Department of Epidemiology, University of Washington, Seattle, Washington, USA; Kaiser Permanente Washington Health Research Institute, Kaiser