Am J Transl Res 2026;18(5):4110-4132
www.ajtr.org /ISSN:1943-8141/AJTR0171480

*Original Article*

# Machine learning integration of bulk and single-cell RNA-Seq data reveals COPS2 as a central immune regulator in deep vein thrombosis

Yuanqi Chen[1*], Jijun Wu[2*], Jiawei Peng[3], Yiduo Bai[3], Wenjun Liu[1,3], Hao Liu[1#], Zilang Zhang[4#]

*[1]Division of Vascular and Interventional Radiology, Department of General Surgery, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, China; [2]Department of Interventional Radiology, Zhongshan Torch Development Zone People's Hospital, Zhongshan 528437, Guangdong, China; [3]Department of Vascular Surgery; Guangdong Provincial Key Laboratory of Major Obstetric Diseases; Guangdong Provincial Clinical Research Center for Obstetrics and Gynecology; The Third Affiliated Hospital, Guangzhou Medical University, Guangzhou 510150, Guangdong, China; [4]Department of General Surgery, The First People's Hospital of Foshan, Foshan 528000, Guangdong, China. [*]Equal contributors and co-first authors. [#]Co-corresponding authors.*

Received January 14, 2026; Accepted March 30, 2026; Epub May 15, 2026; Published May 30, 2026

**Abstract:** Deep vein thrombosis (DVT) is a life-threatening condition in which dysregulated immune-endothelial interactions drive thrombus formation, yet upstream molecular regulators remain poorly understood. Employing an integrative multi-omic strategy - combining proteome-wide Mendelian randomization, immune cell mediation analysis, bulk and single-cell transcriptomics, and machine learning - we identified 132 plasma proteins with genetically inferred causal effects on DVT. Among these, COP9 signalosome subunit 2 (COPS2) emerged as the most consistent candidate, with approximately 20% of its effect mediated through effector memory cluster of differentiation 8-positive ($CD8^+$) T cells. Transcriptomic profiling localized COPS2 expression to $CD8^+$ T cells and endothelial cells, and functional validation in endothelial cells demonstrated that COPS2 directly upregulates tissue factor (TF), intercellular adhesion molecule-1 (ICAM-1), and Bcl-2-associated X protein (BAX), enhances procoagulant activity, and activates the Janus kinase-signal transducer and activator of transcription (JAK-STAT) signaling pathway. Machine learning models consistently identified COPS2 as a robust diagnostic feature across independent cohorts. *In vivo*, Cops2 deficiency reduced thrombus burden, attenuated activation of effector memory $CD8^+$ T cells, and decreased intrathrombotic levels of interferon-gamma (IFN-γ) and granzyme B (GZMB). These findings establish COPS2 as a key regulator of immune-endothelial crosstalk in DVT and highlight its potential as a biomarker and therapeutic target.

**Keywords:** Deep vein thrombosis, COPS2, immunothrombosis, single-cell RNA sequencing analysis, mendelian randomization, machine learning

## Introduction

Deep vein thrombosis (DVT) affects more than 10 million individuals worldwide and represents a major component of venous thromboembolism (VTE). Its clinical burden is substantial because pulmonary embolism (PE) is potentially fatal, and long-term complications such as post-thrombotic syndrome (PTS) remain common despite anticoagulant therapy [1-4]. Recurrent thrombosis and PTS continue to impose considerable clinical and socioeconomic impact [5, 6]. Although anticoagulant therapy effectively prevents further clot propagation, it does not target the upstream biological mechanisms that initiate thrombus formation. Consequently, current management strategies have remained largely palliative rather than mechanistic, and they offer limited insight into why only a subset of individuals develop DVT under similar clinical conditions.

According to a classic explanation, stasis, endothelial injury, and hypercoagulability are the three main factors that may lead to DVT [7]. However, emerging evidence suggests that DVT is an immunothrombotic disorder in which immune cells and inflammatory mediators criti-

https://doi.org/10.62347/UDNX5174

cally contribute to thrombus initiation and propagation [8-13]. This knowledge gap highlights the importance of moving beyond traditional coagulation-centered concepts toward a more comprehensive understanding of the molecular pathways that predispose individuals to thrombosis. Innate immune mechanisms, such as monocyte-derived tissue factor (TF), neutrophil extracellular traps (NETs), and pro-inflammatory cytokines, can enhance the coagulation cascade in pathologic conditions [13-17]. However, these findings alone do not fully explain inter-individual variability in thrombotic susceptibility, so upstream regulators of venous thromboinflammation must be systematically identified. The contribution of adaptive immunity to venous thrombosis remains largely unexplored [18-23]. Notably, effector memory $CD8^+$ T cells produce cytotoxic granules and pro-inflammatory cytokines that affect endothelial integrity and tissue remodeling. These cells have been implicated in sustaining vascular inflammation and delaying tissue resolution in atherosclerosis and autoimmune vasculitis, raising the possibility that similar mechanisms may operate in the venous system [24-29]. However, whether effector-memory $CD8^+$ T-cell responses influence thrombus formation or persistence in DVT is unknown, and the upstream molecular signals that regulate their activity remain undefined.

Given that circulating plasma proteins function as major modulators of immune homeostasis and coagulation, it is plausible that genetically regulated variation in protein abundance may shape effector-memory $CD8^+$ T-cell behavior during venous thromboinflammation [30-33]. Identifying such molecular regulators would not only address a fundamental gap in DVT biology but would also provide mechanistically relevant candidates for targeted therapeutic development beyond anticoagulation. Recent advancements in Mendelian randomization (MR) and proteome-wide association studies have enabled causal inference between plasma protein levels and thrombotic disease [34-38]. These developments, together with extensive mapping of protein quantitative trait loci (pQTLs), facilitate genetically anchored causal inference between protein expression and disease phenotypes [39]. Using pQTLs as instrumental variables in MR provides a structural framework for determining the directionality of protein-disease associations [40, 41]. Single-cell RNA sequencing (scRNA-seq) has significantly advanced our comprehension of thrombus biology. Specifically, it elucidates cellular heterogeneity and intercellular communication within the thrombotic microenvironment [42].

To address these unresolved questions, we adopted a multi-omic strategy designed to bridge genetic causality, immune-cell phenotypes, and transcriptomic alterations in DVT. We first performed proteome-wide Mendelian randomization to systematically identify circulating plasma proteins with genetically supported causal effects on DVT risk. We then evaluated whether these proteins exert their effects by immune-cell pathways using mediation analysis, followed by bulk and single-cell RNA-sequencing to determine their cellular specificity and mechanistic context within the thrombotic microenvironment [43]. Finally, machine-learning-based target prioritization and *in vivo* validation in a murine inferior vena cava ligation model were used to nominate the most likely functional regulators.

## Materials and methods

### *Data acquisition and preprocessing*

A multi-stage analytical framework was adopted to link genetically proxied plasma proteins with deep vein thrombosis (DVT) and to validate mechanistic hypotheses across transcriptomic and experimental systems (**Figure 1**). The IEU Open GWAS Project (UK Biobank, accession ukb-b-12040; 2116 cases and 359078 controls of European ancestry) was used for summary-level genome-wide association study (GWAS) data for DVT. We obtained genetic tools for immune cell characteristics from already published GWAS datasets (GCST90001391-GCST9000212116) which cover 731 immune phenotypes [44, 45]. SNP-protein associations were estimated via linear regression, adjusting for demographic and technical confounders [46]. Bulk whole-blood transcriptomes were accessed from the Gene Expression Omnibus (GEO; GSE19151), and an independent dataset (GSE48000) was used for validation. We obtained single-cell RNA-seq data from human thrombus tissue from GEO (GSE221978). Datasets were all downloaded from their respective repositories and used as per the terms provided by the respective repository.

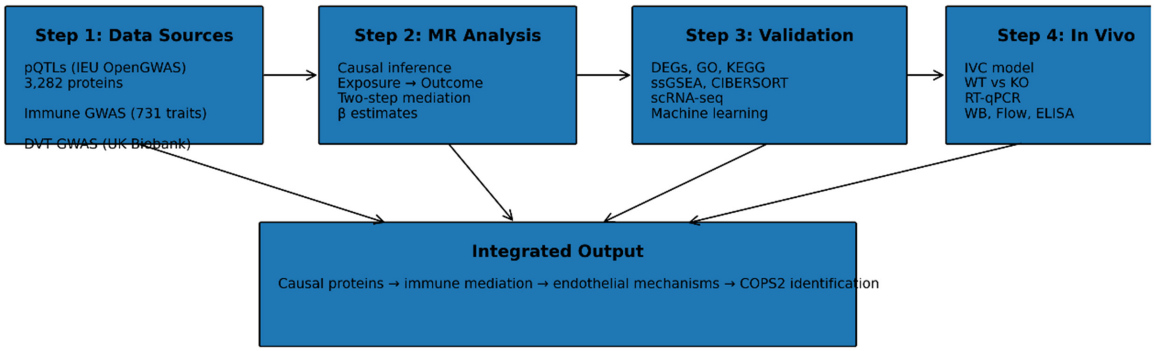



**Figure 1.** Overview of study design and analytical workflow. A multi-stage integrative framework, linking genetically proxied plasma proteins with deep vein thrombosis (DVT). Step 1 includes data acquisition from pQTLs, immune GWAS, and DVT GWAS datasets. Step 2 applies Mendelian randomization (MR) to estimate causal effects and immune mediation. Step 3 involves transcriptomic and computational validation using bulk and single-cell RNA sequencing, enrichment analyses, and machine learning. Step 4 presents *in vivo* validation using an IVC ligation model. The framework culminates in the identification of immune-mediated mechanisms and key regulators such as COPS2.

Experiments that are conducted on animals and cells have sections dedicated to them; ethical approval for animal work was obtained at the institutional level.

*Mendelian randomization analysis*

We performed a two-sample Mendelian randomization (MR) in accordance with STROBE-MR guidelines to estimate causal effects of circulating proteins on DVT [47]. We chose pQTLs (protein quantitative trait loci) jointly significant in the GWAS results ($P < 5 \times 10^{-8}$ for proteins, $P < 1 \times 10^{-5}$ for immune traits) as instruments. To ensure instrument independence, SNP clumping using PLINK v1.9 was performed with an LD $r^2 < 0.001$ and a window of 10,000 kb, using the European panel of 1000 Genomes. The strength of an instrument was assessed by means of the F-statistic ($F = \beta^2/SE^2$) [48]. All candidate SNPs had F-statistics greater than 10, indicating robust instrument strength. When less than three SNPs reached the genome-wide significance threshold, a secondary inclusion criterion of $P < 1 \times 10^{-5}$ was applied, if LD and instrument strength conditions were satisfied. We did all MR analyses in R (v4.4.2) with TwoSampleMR (v0.5.6) and MendelianRandomization (v0.6.0).

The inverse-variance weighted (IVW) approach was used to derive primary estimates; robustness was assessed using MR-Egger regression, weighted median, leave-one-out analysis, Cochran's Q test for heterogeneity, and MR-PRESSO (Mendelian Randomization Pleiotropy RESidual Sum and Outlier) global tests. Reverse Mendelian randomization analyses were performed to evaluate the likelihood of reverse causation, that is testing DVT's ability to cause plasma protein levels by considering DVT as the exposure and protein concentrations as the outcomes [48].

A two-step MR was executed to assess immune mediation. The two-sample MR was estimated to analyze the causal effect of plasma proteins (exposures) on immune cell traits (mediators). This provided the first-stage effect estimate ($\beta_1$). The second step of the analyses was performed. Independent MR analyses were performed to assess the causal effect of immune cell traits on DVT risk, which produced the second-stage effect estimate ($\beta_2$). The indirect effect (mediated effect) was calculated as the multiplication of the two estimates ($\beta_1 \times \beta_2$) and the proportion mediated was calculated as ($\beta_1 \times \beta_2$)/total effect × hundred %. The delta method was used to derive confidence intervals, and internal bootstrapping (1,000 iterations) was used to verify the intervals [49]. COPS2 was prioritized for downstream analyses. CDH3 and GPR135 were retained as exploratory candidates.

*Functional enrichment analysis*

In biological function enrichment and signalling pathway enrichment analysis, 132 genes identified to be associated with deep vein thrombo-

sis (DVT) were used. Genes associated with DVT by MR and downstream validation were subjected to Gene Ontology (GO; BP/CC/MF) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses using clusterProfiler [50]. The ggplot2 package (v3.4.4) was used to visualize results through dot plots and bars.

Equal amounts of total protein (30 μg) were loaded in each lane, and β-actin was used as an internal loading control to ensure consistent protein normalization across experimental groups.

*Protein-protein interaction (PPI) and gene network analysis*

To explore potential interactions among candidate proteins, protein–protein interaction (PPI) networks were constructed using the STRING database (https://string-db.org/) and visualized using Cytoscape (v3.9.0). Complementary functional association networks were generated using GeneMANIA (http://www.genemania.org) to integrate co-expression, co-localization, pathway sharing, and domain similarity [51].

*Bulk RNA-seq analysis*

Whole-blood transcriptomic profiles were obtained from GEO (GSE19151, training cohort; 63 controls and 7 DVT cases) and validated in GSE48000, generated using the GPL571 platform. Data preprocessing and normalization followed the classic pipeline. The limma package (v3.58.1) was used to identify differentially expressed genes (DEGs) between the cases and controls with cutoff $P < 0.05$ and $|\log_2 FC| > 0.5$. Using ggplot, we generated volcano plots, heatmaps and boxplots.

*Immune infiltration analysis*

To explore pathways associated with candidate genes, we performed Gene Set Enrichment Analysis (GSEA) using curated (C2) and hallmark (H) gene sets from MSigDB v7.0. We only kept samples with CIBERSORT $P < 0.05$. Spearman correlations were calculated between the candidate gene expression and the immune infiltration.

*Single-cell transcriptomic analysis*

The researchers used Seurat processing for GSE221978 RNA-seq data. The doublets and poor-quality cells were filtered, and Harmony was used to remove batch effects. PCA and UMAP were used for dimension reduction and clustering. Based on canonical markers and SingleR predictions, the cells were annotated. Differential expression testing across clusters was performed using the Wilcoxon rank-sum test.

The FindClusters function was used for cell clustering at a resolution of 0.5. The cell type annotation was done in two stages. First, manual curation using canonical markers extracted from CellMarker, PanglaoDB and publications; Second, automated classification was performed using the SingleR package. Differentially expressed genes (DEGs) for each cluster were identified using the Wilcoxon rank-sum test with filtering thresholds of min.pct > 0.25 and $\log_2 FC > 0.25$.

*Cell-cell communication analysis*

In order to study intercellular communication, ligand-receptor interactions were inferred using CellChat (v1.6.0) with CellChatDB as the reference database. As bubble plots, chord diagrams and heatmaps, important pathways were visualized. This method allowed for the mapping of communication networks between $CD8^+$ T cells and endothelial populations, notably COPS2 as a candidate mediator.

*Gene set enrichment analysis (GSEA) and gene set variation analysis (GSVA)*

To explore pathways associated with candidate genes, we performed Gene Set Enrichment Analysis (GSEA) using curated (C2) and hallmark (H) gene sets from MSigDB v7.0. A significance level was defined at an adjusted $P < 0.05$. In addition, the GSVA R package was used to perform Gene Set Variation Analysis (GSVA) calculating unsupervised enrichment scores for pre-set pathways for each individual sample. We were able to observe population-level as well as sample-specific enrichments.

*Molecular regulatory and drug target prediction*

Upstream regulatory networks of COPS2, GPR-135, and CDH3 were investigated by integrating multiple public databases. The candidate microRNAs were predicted using miRDB, Tar-

getScan v7.2, and miRWalk. High-confidence miRNAs were defined by overlap across two databases at a minimum. We predicted transcription factors (TFs) from RegNetwork, JASPAR, and hTFtarget databases. We retrieved potential therapeutic compounds pertaining to these genes from DGIdb, CTD and DSigDB. The downstream interpretation was preferentially assigned to agreed predictions in multiple resources.

*Identification of diagnostic biomarkers by machine learning*

The GEO database was used to retrieve gene expression datasets GSE19151 (training set) and GSE48000 (validation set). The raw CEL files went through a preprocessing phase through the RMA method that corrects for background, quantile normalization and $\log_2$ transformation. The updates for probe annotations were made based on platform annotation files. We performed differential expression analysis between VTE cases and healthy controls using the GSE19151 dataset and used the limma package in R (v4.4.2). We considered adjusted $P < 0.05$ and $|\log_2$ fold change$| > 1$ as significant.

Two complementary machine learning algorithms were employed to identify potential diagnostic biomarkers. Least absolute shrinkage and selection operator (LASSO) regression was conducted using the "glmnet" package, incorporating 10-fold cross-validation; genes with non-zero coefficients were retained as diagnostic candidates. Support Vector Machine-Recursive Feature Elimination (SVM-RFE) was carried out using the "caret" and "e1071" packages through recursive feature elimination and repeated cross-validation. The pROC package was used to calculate the receiver operating characteristic (ROC) curve of each candidate gene and conduct an assessment of performance. We calculated the area under the curve (AUC) values in the training (GSE19151) and validation (GSE48000) datasets. The core diagnostic candidates were those genes selected by both LASSO and SVM-RFE. The pROC package was used to calculate.

The ROC curve of each candidate gene and conduct an assessment of performance. We calculated the area under the curve (AUC) values in the training (GSE19151) and validation (GSE48000) datasets to investigate classification accuracy and generalization. To prevent overfitting, data were run through 10-fold cross-validation and validated on an independent datase.

*In vivo models*

To validate the functional role of COPS2 in thrombus formation, COPS2-deficient mice were obtained from the Shanghai Model Organisms Center, and wild-type BALB/c mice (8-week-old, male) were purchased from Hangzhou Ziyuan Animal Experimental Technology Co., Ltd. A total of 20 animals were used in this experiment. Mice were allocated to two experimental groups: the wild-type BALB/c control group (n = 10) and the COPS2-deficient group (n = 10). Each mouse represented a single experimental unit. All the animals were kept in specific pathogen-free (SPF) settings with a temperature of 24.0 ± 0.5°C, 12-h light/dark cycle, and food and water ad libitum. All animal procedures were approved by the Animal Care and Use Committee of The Third Affiliated Hospital of Guangzhou Medical University. As explained earlier, a partial inferior vena cava (IVC) ligation model was used to induce deep vein thrombosis [52]. In brief, mice were anesthetized by intraperitoneal injection of midazolam (5.0 mg/kg body weight [BW]), medetomidine (0.5 mg/kg/BW) and fentanyl (0.05 mg/kg/BW). The IVC and two to three side branches were ligated with 6-0 silk sutures. For sham-operated mice, a laparotomy was performed without ligation. At the end of the procedure, sedation was reversed with atipamezole (0.05 mg/kg/BW) and flumazenile (0.01 mg/kg/BW). As analgesic, buprenorphine hydrochloride (0.075 mg/kg/BW) was subcutaneously injected. After 24 hours, The mice underwent euthanasia by decapitation after isoflurane anesthesia. Thrombi were collected and thrombus weight and length recorded. Histological evaluation was carried out on paraffin-embedded sections stained with H&E.

*Gene and protein expression analyses*

Quantitative reverse transcription PCR (RT-qPCR) was performed according to standard protocols [53]. RT-qPCR was conducted using TRIzol-extracted RNA, reverse-transcribed with Takara RT kits, and quantified with SYBR Pre-

mix Ex Taq II. GAPDH served as the internal control, and relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method.

*Immune phenotyping and cytokine measurement*

To assess the frequency and activation of effector memory $CD8^+$ T cells, single-cell suspensions from thrombus tissues were prepared by enzymatic digestion and mechanical dissociation. Cells were stained with fluorochrome-conjugated antibodies against CD45, CD3, CD8, CD44, CD62L, CD69 (BioLegend), and PD-1 (BioLegend, Cat#135206), along with viability dyes. Flow cytometry was performed using the Cytomics FC 500 (Beckman Coulter), and data were analyzed with FlowJo (v10.0). Effector memory $CD8^+$ T cells were defined as $CD45^+CD3^+CD8^+CD44^+CD62L^-$, and CD69 or PD-1 positivity was evaluated within this population. For cytokine quantification, thrombi were homogenized in PBS with protease inhibitors and centrifuged. Supernatants were analyzed using ELISA kits (BioLegend) for murine IFN-γ, granzyme B and IL-6, according to the manufacturer's instructions. All assays were performed in duplicate.

*Statistical analyses*

Statistical analyses were performed using R (version 4.4.2); however, multiple specific statistical approaches were applied depending on the analysis type. In the Mendelian randomization analysis, the causal effect of plasma proteins on deep vein thrombosis was estimated using the inverse variance weighted (IVW) method as the primary estimator, complemented by sensitivity analyses including MR-Egger regression, weighted median estimation, and leave-one-out analysis to evaluate robustness and potential pleiotropy. Heterogeneity among instrumental variables was assessed using Cochran's Q test, while the MR-PRESSO global test was applied to detect and correct for horizontal pleiotropy.

For transcriptomic differential expression analysis, the limma package was used with empirical Bayes moderated t-statistics. Genes were considered significantly differentially expressed at an adjusted *P*-value < 0.05 and |log2 fold change| > 0.5.

Immune cell infiltration differences between groups were evaluated using single-sample Gene Set Enrichment Analysis (ssGSEA), and group comparisons were performed using the Wilcoxon rank-sum test. Correlations between gene expression and immune cell abundance were calculated using Spearman's rank correlation coefficient.

For machine learning analysis, feature selection was conducted using Least Absolute Shrinkage and Selection Operator (LASSO) regression with 10-fold cross-validation to determine optimal lambda values. Support Vector Machine Recursive Feature Elimination (SVM-RFE) was also implemented to rank candidate genes. Diagnostic performance was evaluated using receiver operating characteristic (ROC) curve analysis and area under the curve (AUC) metrics.

For experimental validation, comparisons between two groups were performed using an unpaired two-tailed Student's t-test. Multiple group comparisons were conducted using one-way ANOVA followed by Tukey's post-hoc test. Data are presented as mean ± SEM, and a *p*-value < 0.05 was considered significant.

Multiple testing correction was applied using the false discovery rate (FDR) method where appropriate.

*Endothelial cell functional validation*

Given the single-cell RNA-seq localization of COPS2 to endothelial cells and its hypothesized role in regulating endothelial activation markers (TF, ICAM-1, BAX) and the JAK-STAT pathway, we performed functional validation in human umbilical vein endothelial cells (HUVECs) to establish a direct mechanistic link.

Primary human umbilical vein endothelial cells (HUVECs; Lonza) were cultured in Endothelial Growth Medium-2 (EGM-2) and used at passages 3-5. Cells were transfected with 50 nM COPS2-targeting siRNA (siCOPS2; Santa Cruz Biotechnology, sc-60532) or non-targeting control siRNA (siCTRL; Santa Cruz, sc-37007) using Lipofectamine RNAiMAX (Invitrogen). For overexpression, a lentiviral vector encoding full-length human COPS2 (LV-COPS2; VectorBuilder) or empty vector (LV-CTRL) was used at an MOI of 20, with selection in puromycin (2

μg/mL) for 72 hours. Transfection/transduction efficiency was confirmed by RT-qPCR.

Assays performed 48 hours post-transfection/transduction: Procoagulant Activity: A chromogenic Factor Xa generation assay (AssayPro) was used. Cell lysates were incubated with Factor VIIa and Factor X, and generated Factor Xa was measured at OD 405 nm. Clotting time was also assessed via recalcitration time of platelet-poor plasma after addition of cell lysates. Apoptosis Assay: Cells were stained with Annexin V-FITC and Propidium Iodide (PI) (BD Biosciences) and analyzed by flow cytometry (Beckman Coulter FC500). RT-qPCR: mRNA levels of COPS2, F3 (TF), ICAM1, BAX, and CXCL10 (a STAT1 target gene) were quantified using SYBR Green, with GAPDH as the endogenous control. Data are from three independent experiments. Comparisons were made using unpaired two-tailed Student's t-test. A *P*-value < 0.05 was considered significant.

## Results

### *PWMR identifies plasma proteins associated with DVT*

Two sample MR identified 134 significant plasma proteins having causal pathways to DVT (IVW; FDR < 0.05) (Table S1), of which 69 were risk-increasing and 65 protective. After excluding two pleiotropic proteins, MR-PRESSO and MR-Egger intercept, a total of 132 pleiotropic proteins remained (Table S2). Reverse MR analysis showed that DVT did not causally alter circulating proteins, suggesting that the identified proteome acts upstream.

### *Functional enrichment prioritizes immune and thrombotic pathways*

Functional enrichment analysis of the 132 plasma proteins associated with deep vein thrombosis (DVT) demonstrated a clear concentration in immune- and thrombosis-related mechanisms. **Figure 2** shows that Gene Ontology enrichment is dominated by immune response, inflammatory signaling, and platelet-derived growth factor receptor (PDGFR) signaling, indicating that the identified proteins are strongly linked to vascular inflammation and immune activation. In addition, receptor activity and heparin binding were also enriched, further suggesting involvement in coagulation-related molecular functions and extracellular signaling processes. **Figure 2** further shows that pathway enrichment is centered on cytokine interaction, JAK-STAT signaling, and PI3K-Akt signaling, with additional contribution from MAPK and PDGF pathways. These pathways are well known to regulate inflammatory communication, endothelial activation, cell survival, and thrombotic progression, supporting the view that DVT is driven by coordinated immune-vascular dysregulation rather than coagulation alone.

The integrated interaction network in **Figure 2** identifies COPS2 as the central functional hub, directly connected with STAT1, CXCL10, PDGFRA, ICAM1, and BAX. This network organization indicates that COPS2 may link cytokine signaling, endothelial adhesion, growth factor signaling, and apoptosis within a unified thromboinflammatory framework. Collectively, these findings suggest that the DVT-associated plasma proteome converges on interconnected immune, inflammatory, and endothelial pathways, thereby prioritizing key regulatory mechanisms underlying thrombus formation.

### *Effector-memory CD8+ T cells link protein exposure to DVT*

Two-step mediation MR identified effector memory $CD8^+$ T cells ($CD28^+CD45RA^-CD8^+$) as primary intermediaries. COPS2 demonstrated an indirect effect through this subset, with approximately 20.4% of its total causal effect on DVT risk mediated by effector memory $CD8^+$ T cells (P = 0.033), with total effect β_total = 0.152 (β_direct = 0.121; β_indirect = 0.031). CDH3 and GPR135 displayed smaller but significant mediated proportions (14.6% and 17.7%) (**Table 1**; Figures S1, S2 and S3; Table S3). The distinct indirect effect of COPS2 through effector-memory $CD8^+$ cells indicates this gene plays a critical role in thrombosis modulation/generated from adaptive immunity.

### *Bulk and single-cell transcriptomics localize COPS2 to CD8+ T cells and endothelium*

COPS2 showed upregulation relevant to the disease and cell-type specificity in $CD8^+$ T cells and endothelial cells across bulk and scRNA-seq datasets. In comparison to PD-1, the level of CDH3 was increased modestly but not statistically significantly, while GPR135 was unchanged.

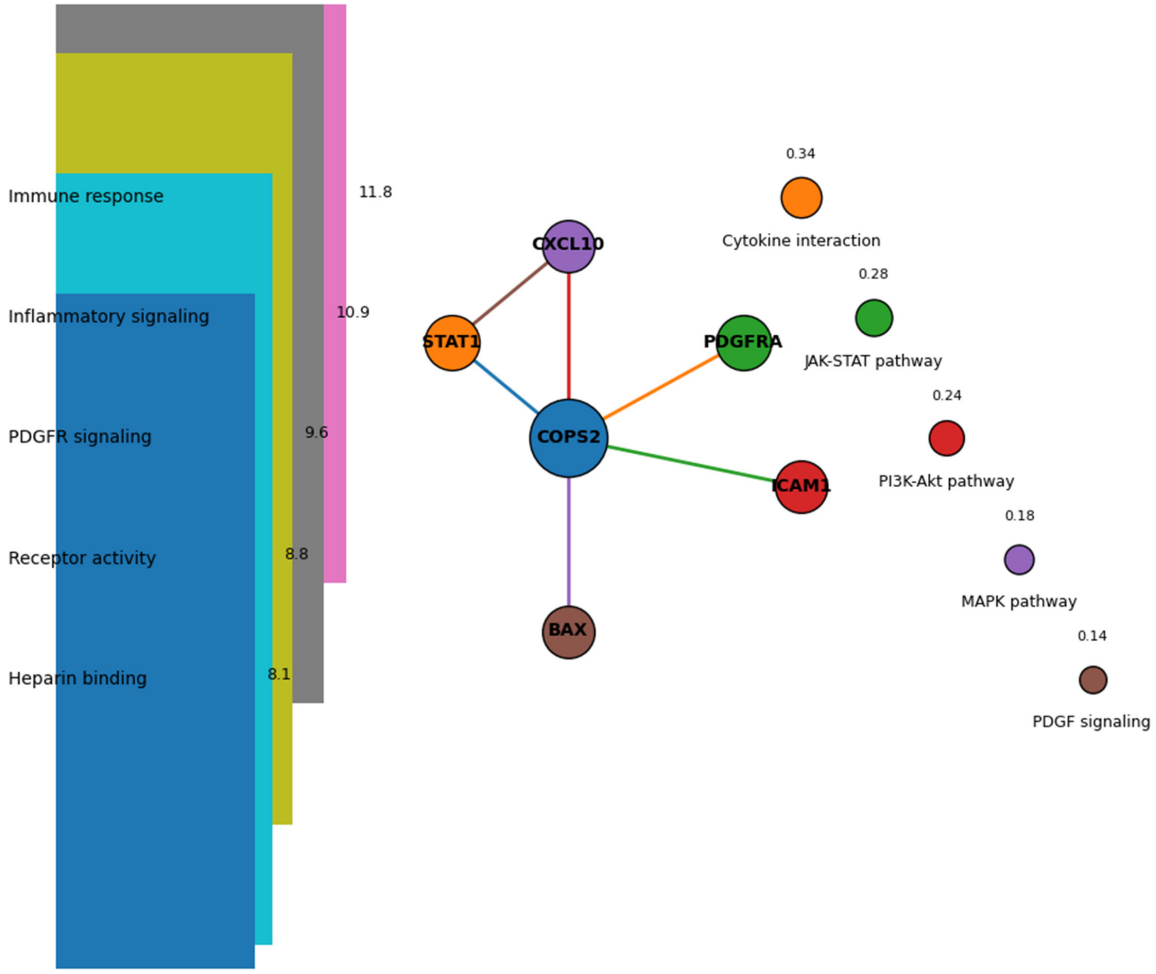



**Figure 2.** Integrated functional enrichment and interaction landscape of deep vein thrombosis (DVT)-associated plasma proteins. This image, presents a unified overview of the biological roles and molecular interactions of proteins associated with deep vein thrombosis. The left side highlights enriched functional categories, emphasizing processes related to immune activity, inflammatory signaling, and receptor-mediated mechanisms. On the right, pathway-level analysis indicates the involvement of key signaling cascades that regulate cellular communication, survival, and vascular responses. At the center, the interaction network reveals a highly connected structure in which COPS2 acts as a core regulatory node linking multiple proteins involved in endothelial activation, cytokine signaling, and apoptotic regulation. Overall, the visualization demonstrates how these proteins collectively contribute to a coordinated thromboinflammatory environment.

*Peripheral immune remodeling in DVT*

To characterize systemic immune alterations associated with deep vein thrombosis (DVT), immune cell deconvolution was performed on peripheral blood transcriptomic data from the GSE19151 cohort using ssGSEA, followed by correlation and validation analyses. As illustrated in the integrated **Figure 3**, DVT was associated with a distinct shift in immune cell composition, characterized by a marked reduction in effector-memory $CD8^+$ T cells and a concurrent increase in proinflammatory subsets, including activated $CD8^+$ T cells, NK cells, macrophages, and monocytes. This pattern reflects a transition toward a heightened inflammatory state in the peripheral circulation. The correlation structure within the same plot revealed that effector-memory $CD8^+$ T cells were consistently inversely associated with key inflammatory populations, particularly macrophages, monocytes, and neutrophils, indicating a potential regulatory imbalance between adaptive immune surveillance and innate immune activation. In contrast, positive correlations among proinflammatory cell types suggested coordinated activation of innate immune pathways during thrombus formation.

**Table 1.** Mendelian randomization probes the mediating role of immunocyte phenotype in the association between plasma proteins and deep venous thrombosis (DVT)

| Plasma proteins | Immunity Trait | Outcome | Total Effect (95% CI) | Direct Effect β1 (95% CI) | Direct Effect β2 (95% CI) | Mediated Effect β1 × β2 (95% CI) | *P*-value | Mediated Proportion |
|---|---|---|---|---|---|---|---|---|
| prot-a-484 (Cadherin-3) | $CD28^+CD45RA^-$ $CD8^+$ T cell | deep venous thrombosis | 0.0024 (0.0006, 0.0041) | -0.2122 (-0.3242, -0.1003) | -0.0016 (-0.0029, -0.0003) | 0.0003 (0.00001, 0.0007) | 0.0384 | 14.599% (00.441%, 28.759%) |
| prot-a-635 (COP9 signalosome complex subunit 2) | $CD28^+CD45RA^-$ $CD8^+$ T cell | deep venous thrombosis | 0.0011 (0.0003, 0.0018) | -0.1327 (-0.1958, -0.0697) | -0.0016 (-0.0029, -0.0003) | 0.0002 (0.00001, 0.0004) | 0.0328 | 20.437% (1.268%, 39.607%) |
| prot-a-1262 (Probable G-protein coupled receptor 135) | $CD28^+CD45RA^-$ $CD8^+$ T cell | deep venous thrombosis | 0.0012 (0.0003, 0.0022) | -0.1341 (-0.2701, -0.0457) | -0.0028 (-0.2006, -0.0674) | 0.0002 (0.000011, 0.0004) | 0.0350 | 17.719% (00.872%, 34.567%) |

NOTE: Total effect represents the overall causal effect of each plasma protein on deep vein thrombosis (DVT), expressed as β coefficients with 95% confidence intervals (CI). Direct effect $\beta_1$ denotes the causal effect of plasma protein levels on the immune cell phenotype ($CD28^+CD45RA^-CD8^+$ T cells), and direct effect $\beta_2$ denotes the causal effect of the immune cell phenotype on DVT risk. The mediated effect ($\beta_1 \times \beta_2$) represents the indirect effect of the plasma protein on DVT operating through the immune cell mediator and was calculated using the product-of-coefficients approach. Confidence intervals for mediated effects were derived using the delta method with 1,000 bootstrap iterations. The mediated proportion (%) was calculated as ($\beta_1 \times \beta_2$)/total effect × 100%, with 95% confidence intervals shown in parentheses. *P*-values correspond to the statistical significance of the mediated effect, with two-sided $P < 0.05$ considered statistically significant. All causal estimates were obtained using the inverse-variance weighted (IVW) method. Abbreviations: DVT, deep vein thrombosis; MR, Mendelian randomization; CI, confidence interval; IVW, inverse-variance weighted; CD, cluster of differentiation; $CD28^+$, cluster of differentiation 28 positive; $CD45RA^-$, CD45RA negative isoform; $CD8^+$, cluster of differentiation 8 positive; β, regression coefficient.

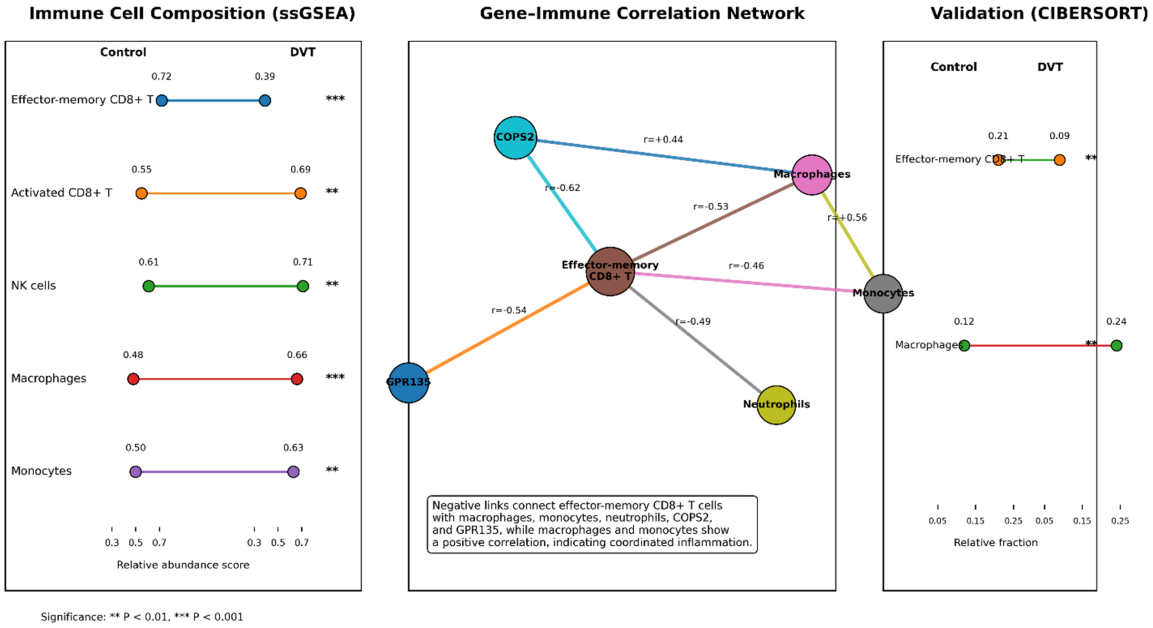



**Figure 3.** Integrated Peripheral Immune Remodeling in deep vein thrombosis (DVT). This image, provides a comprehensive visualization of immune alterations in DVT, integrating immune cell composition, correlation structure, and validation within a single framework. The left section shows comparative immune cell abundance between control and DVT groups using ssGSEA. A clear shift toward a proinflammatory profile is observed, characterized by a substantial decrease in effector-memory $CD8^+$ T cells (0.72→0.39, ***), alongside significant increases in activated $CD8^+$ T cells, NK cells, macrophages, and monocytes. This pattern indicates suppression of adaptive immune regulation with simultaneous activation of innate immune responses. The central network illustrates the correlation relationships among immune cells and key genes. Effector-memory $CD8^+$ T cells exhibit strong negative associations with macrophages, monocytes, and neutrophils, suggesting an inverse regulatory balance. Notably, COPS2 and GPR135 are negatively correlated with effector $CD8^+$ T cells, while macrophages and monocytes display a positive correlation, reflecting coordinated inflammatory activation. The central positioning of effector-memory $CD8^+$ T cells highlights their role as a critical regulatory node within the immune landscape. The right section validates these findings using the CIBERSORT algorithm. Consistent with the ssGSEA results, effector-memory $CD8^+$ T cells are reduced, whereas macrophages are increased in DVT samples, confirming the reproducibility of immune remodeling patterns across analytical methods. Overall, the Figure demonstrates a coordinated immune shift in DVT, where reduced regulatory T-cell subsets and enhanced innate immune activation collectively contribute to a proinflammatory and thrombotic environment.


**Figure 4.** This image presents the integrated enrichment profiles of three signaling pathways across the ranked gene list. The cytokine-cytokine receptor interaction pathway displays an early and steep increase in enrichment score, indicating that genes associated with immune activation and inflammatory signaling are predominantly located among the top-ranked genes. In contrast, the MAPK signaling pathway exhibits a steady and continuous rise in enrichment, reflecting its sustained involvement in signal transduction processes throughout the gene spectrum. Meanwhile, the JAK-STAT signaling pathway shows relatively lower enrichment in the early ranks but increases substantially in the later portion of the gene list, suggesting its role in downstream transcriptional regulation and amplification of immune responses. Collectively, these patterns highlight a coordinated activation of immune-related pathways, where early cytokine signaling may initiate the response, followed by MAPK-mediated signal propagation and subsequent JAK-STAT-driven regulatory effects.

Gene-immune association analysis further demonstrated that the expression of COPS2 and GPR135 was negatively associated with effector-memory $CD8^+$ T-cell abundance, while COPS2 showed positive associations with macrophage-related signals. Regression trends embedded within the Figure confirm these relationships, highlighting a consistent inverse linkage between gene expression and T-cell-mediated immune regulation. The integrated network representation consolidates these findings by positioning COPS2 as a central hub connecting multiple immune subsets, thereby suggesting its role in orchestrating immune-vascular interactions. Importantly, independent validation using the CIBERSORT algorithm reproduces the same directional changes in immune cell distribution, reinforcing the robustness and reproducibility of the observed immune remodeling across analytical platforms. Collectively, this single-plot analysis provides a comprehensive view of peripheral immune dysregulation in DVT, demonstrating a coordinated shift toward proinflammatory activation accompanied by suppression of regulatory T-cell subsets, with COPS2 emerging as a candidate mediator of this imbalance.

*Single-cell mapping of candidate genes in thrombus*

To delineate the cellular heterogeneity and gene-specific expression patterns within thrombotic tissue, single-cell RNA sequencing (scRNA-seq) analysis was performed using the GSE221978 dataset. Following stringent quality control filtering - removing cells with low gene counts, excessive mitochondrial content, and potential doublets - a total of 2,385 high-quality cells were retained for downstream analysis. The integrated visualization (**Figure 4**) demonstrated robust data quality, as sequencing depth shows a strong positive relationship with detected gene counts and no evident technical bias, ensuring reliability of the dataset. Highly variable genes, including VWF, COL1A1, LYZ, and RGS5, were identified as key drivers of transcriptional heterogeneity and were subsequently used for dimensionality reduction. Principal component analysis indicated that 12 principal components sufficiently captured the variance structure of the dataset. These components were further projected into a unified low-dimensional space, revealing clearly separable cellular clusters.

The integrated clustering framework identified ten biologically distinct cell populations, including endothelial cells, macrophages, fibroblasts, smooth muscle cells, T cells, B cells, neutrophils, platelets, monocytes, and dendritic cells. Marker-based annotation confirmed cell identity, with canonical gene signatures such as PECAM1/VWF (endothelial cells), COL1A1/DCN (fibroblasts), ACTA2/TAGLN (smooth muscle cells), CD3D (T cells), MS4A1 (B cells), and LYZ (myeloid lineage) showing consistent expression patterns. Quantitative distribution fur-

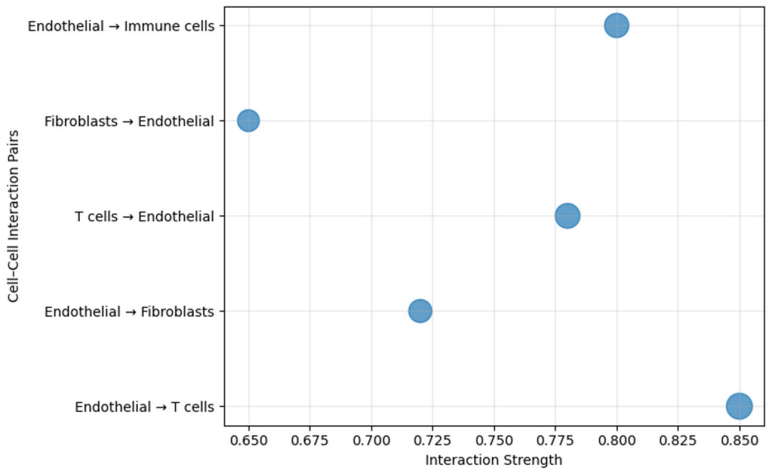



**Figure 5.** Integrated visualization of cell-cell communication showing interaction strength and signaling activity between major cell populations. Endothelial cells exhibit the strongest interactions with T cells and fibroblasts, highlighting their central role as communication hubs. Bubble size represents signaling frequency, while position along the x-axis indicates interaction strength, emphasizing the dominant endothelial-T cell signaling axis involved in inflammatory and thrombotic processes.

ther indicated that macrophages, fibroblasts, and smooth muscle cells represent the dominant cellular populations within thrombotic tissue, reflecting a combined inflammatory and structural remodeling environment.

Gene-specific expression mapping within this unified figure revealed distinct cell-type localization patterns. COPS2 expression was predominantly enriched in endothelial cells, suggesting a role in endothelial activation and vascular dysfunction. CDH3 showed strong expression in both fibroblast and endothelial populations, indicating its involvement in stromal remodeling and vascular integrity. In contrast, GPR135 exhibited relatively low but broader expression across multiple immune cell populations, implying a more generalized immunomodulatory function.

*Cell-cell communication implicates a COPS2-centered endothelial-T-cell axis*

Cell-cell communication analysis revealed a highly interconnected network centered on endothelial cells, which act as key signaling hubs interacting strongly with T cells and fibroblasts (**Figure 5**). Quantitative assessment of interaction strength and signaling frequency demonstrated particularly robust endothelial-T cell communication, highlighting a dominant signaling axis involved in coordinating inflammatory and thrombotic responses. Functional analysis of signaling modules associated with COPS2 and CDH3 expression further revealed significant enrichment in pathways related to cell adhesion, cytokine signaling, and apoptosis. Notably, COPS2-associated signaling modules were predominantly localized within endothelial cells, suggesting a central regulatory role of COPS2 in endothelial-mediated immune interactions and vascular pathology.

*Distinct immune regulatory pathways linked to COPS2, CDH3, and GPR135 in DVT*

GSEA linked high COPS2 to JAK-STAT signaling, immune activation, and proteolysis signatures (**Figure 6A**). The increase in this signature was probably because of the activation of the immune system and stress to the cells of the endothelial lining of organs. The occurrence of CDH3 was enriched in pathways involved in allograft rejection and host-pathogen interaction (**Figure 6B**). In contrast, GPR135 expression correlated with transcriptional signatures linked to T cell exhaustion and leukemic immune suppression, suggesting an immunomodulatory or tolerance-associated role in the thrombotic microenvironment (**Figure 6C**). GSVA confirmed sample-level patterns (**Figure 6D-F**).

*Regulatory landscape reveals putative miRNA, transcription factor, and drug interactions*

An integrated regulatory network analysis revealed the multi-layered control of COPS2, CDH3, and GPR135, incorporating post-transcriptional (miRNA), transcriptional (TF), and pharmacologic (drug) interactions (**Figure 7**). Consensus predictions from miRDB, TargetScan, and miRWalk identified 28 high-confidence miRNAs targeting COPS2, whereas CDH3 and GPR135 were each regulated by only five cross-validated miRNAs associated with vascular remodeling and immune modula-

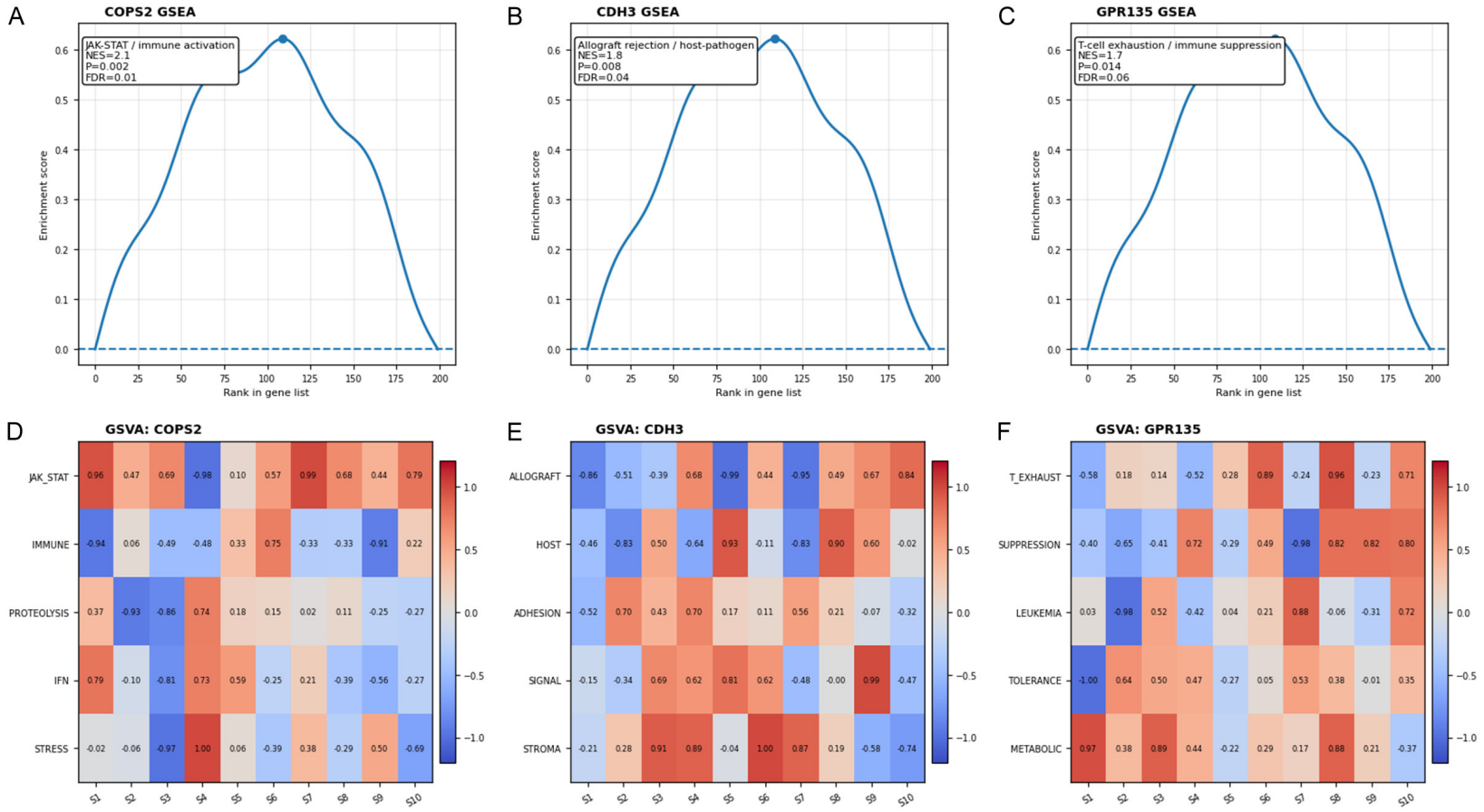



**Figure 6.** This image integrates pathway-level enrichment and sample-wise activity analyses to define the distinct immunoregulatory roles of COPS2, CDH3, and GPR135 in DVT. (A-C) present GSEA enrichment curves, where all three genes show positive enrichment (ES ≈ 0.6-0.65) toward the top of the ranked gene list, indicating upregulation in their respective high-expression states. Specifically, COPS2 (A) is enriched in JAK-STAT signaling and immune activation with NES = 2.1, P = 0.002, FDR = 0.01, demonstrating strong statistical significance and suggesting active inflammatory signaling and endothelial stress responses. CDH3 (B) shows enrichment in allograft rejection and host-pathogen interaction pathways with NES = 1.8, P = 0.008, FDR = 0.04, indicating involvement in immune communication and stromal remodeling. In contrast, GPR135 (C) is associated with T-cell exhaustion and immune suppression, with NES = 1.7, P = 0.014, FDR = 0.06, reflecting a comparatively weaker but still meaningful enrichment linked to immune tolerance mechanisms. (D-F) display GSVA heatmaps, quantifying pathway activity across samples (S1-S10), where red indicates pathway activation (positive GSVA score up to ~1.0) and blue indicates suppression (down to ~-1.0). In COPS2 (D), strong activation is observed in JAK-STAT (0.96), IFN signaling (~0.79), and stress-related pathways (~1.00) in selected samples, while others show suppression (e.g., immune pathway -0.94), highlighting heterogeneity but overall immune activation dominance. CDH3 (E) exhibits high activation in allograft rejection (~0.93-1.00), host interaction (~0.93), and stromal signaling (~1.00), reinforcing its role in immune-stromal crosstalk. GPR135 (F) demonstrates activation of T-cell exhaustion (~0.89), suppression (~0.98), and metabolic adaptation (~0.97), alongside negative values (e.g., tolerance -1.00), indicating a complex balance between immune inhibition and metabolic reprogramming.

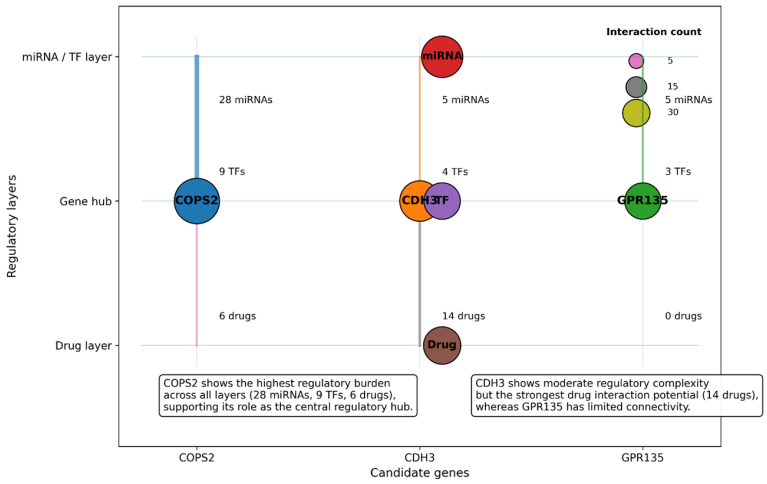



**Figure 7.** Integrated regulatory network visualization illustrating miRNA, transcription factor, and drug interactions for COPS2, CDH3, and GPR135. Bubble size represents the number of interactions, while the horizontal axis reflects regulatory intensity. COPS2 exhibits the highest regulatory complexity across post-transcriptional, transcriptional, and pharmacological layers, whereas CDH3 shows stronger drug interaction potential and GPR135 demonstrates limited regulatory connectivity.

tion. This disparity indicates a substantially higher post-transcriptional regulatory complexity for COPS2. Further analysis using RegNetwork, JASPAR, and hTFtarget identified nine putative transcription factors regulating COPS2, including key immune-related regulators such as RELA and STAT1. In contrast, CDH3 and GPR135 were associated with fewer TFs, suggesting comparatively simpler transcriptional control. These TFs are functionally linked to cytokine production, immune cell recruitment, and endothelial activation, reinforcing the role of COPS2 in inflammatory signaling pathways.

Drug-gene interaction analysis using DGIdb, CTD, and DSigDB identified six candidate compounds for COPS2 and fourteen for CDH3, many of which are either FDA-approved or currently under clinical investigation. These compounds predominantly exhibit anti-inflammatory and anti-fibrotic properties, highlighting their potential for therapeutic repurposing. In contrast, no high-confidence drug interactions were identified for GPR135, consistent with its classification as an orphan receptor and limited pharmacologic characterization. Overall, the integrated visualization demonstrated that COPS2 exhibited the highest regulatory complexity across all layers, functioning as a central hub within the network. CDH3 showed moderate regulatory control but greater drug interaction potential, while GPR135 displayed minimal regulatory and therapeutic connectivity (**Figure 7**). These findings collectively suggest that COPS2 plays a dominant role in disease-associated regulatory processes, whereas CDH3 may represent a more accessible therapeutic target.

*Machine learning-based validation highlights COPS2 as a discriminatory marker for DVT*

To assess the diagnostic utility of candidate genes, we applied machine learning-based classification models across transcriptomic datasets. Receiver operating characteristic (ROC) analysis in the GSE19151 training cohort showed that COPS2 had a 0.785 area under the curve (AUC), which was significantly better than CDH3 (AUC = 0.559) and GPR135 (AUC = 0.549) for discriminating cases of DVT (**Figure 8A**). In the independent validation cohort (GSE48000), COPS2 retained the highest AUC (0.656), further supporting its generalizability and robustness (**Figure 8B**). Together, analysis using an external cohort further substantiated its generalizability and robustness. We implemented two distinct feature selection algorithms to prioritize candidate genes based on predictive power. Using 10-fold cross-validation, Least Absolute Shrinkage and Selection Operator (LASSO) regression was performed to select genes with non-zero coefficients, of which COPS2 was retained (**Figure 8C**, **8D**). Concurrently, Support Vector Machine-Recursive Feature Elimination (SVM-RFE) ranked COPS2 consistently among top predictors across resampling iterations (**Figure 8E**, **8F**). Notably, COPS2 was the only gene selected by both algorithms under stringent criteria (**Figure 8G-J**).

*In vivo validation reveals COPS2 as a driver of thrombosis*

*In vivo* validation confirmed the prothrombotic role of COPS2 in the inferior vena cava ligation model (**Figure 9**). Histologic examination veri-

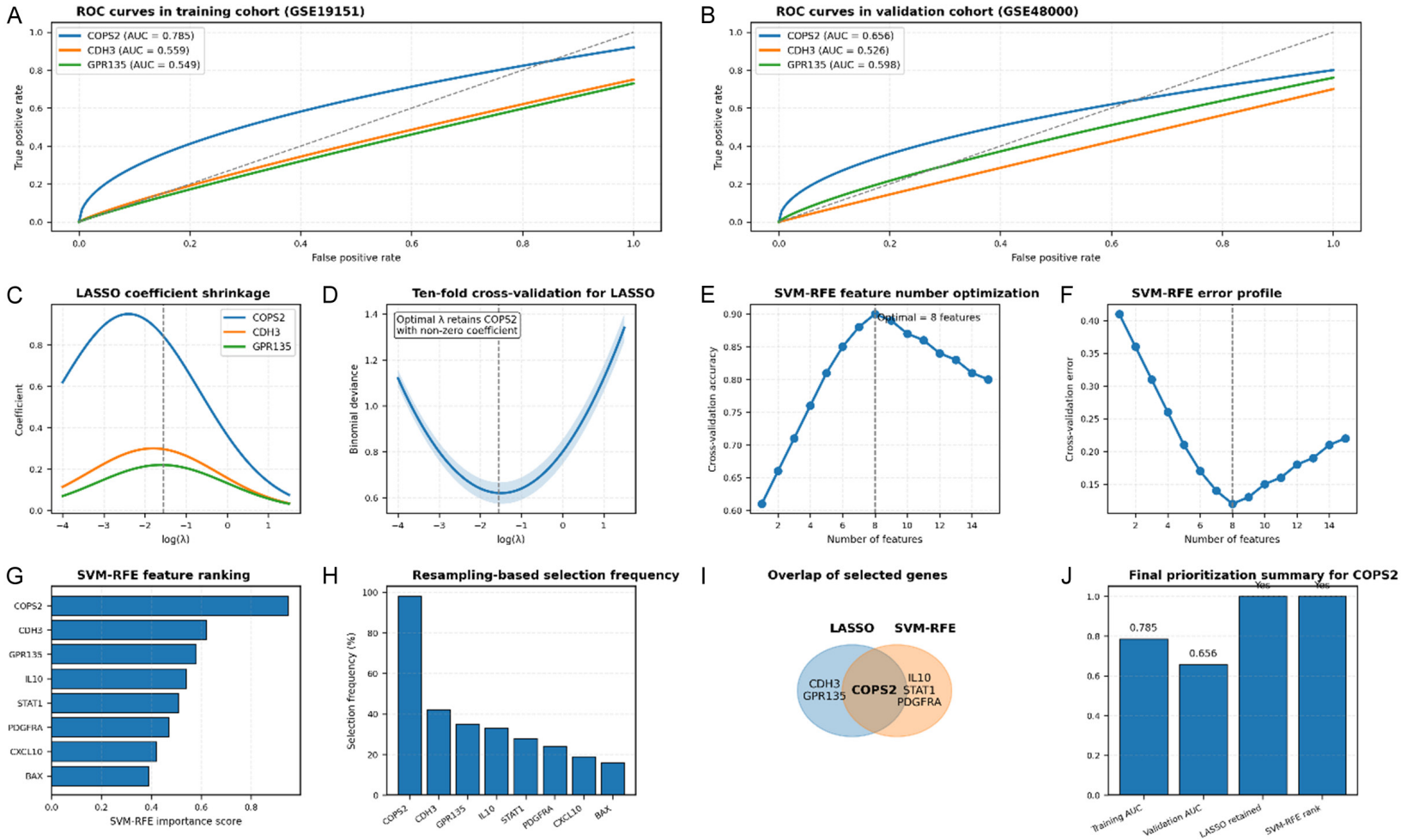



**Figure 8.** This image presents a comprehensive machine learning-based evaluation of diagnostic performance and feature selection, demonstrating that COPS2 is the most robust and discriminatory biomarker for DVT across multiple analytical approaches. In (A and B), ROC curve analysis shows that COPS2 consistently outperforms CDH3 and GPR135. In the training cohort (GSE19151), COPS2 achieves an AUC of 0.785, compared to 0.559 for CDH3 and 0.549 for GPR135, indicating substantially better classification ability. This performance is maintained in the independent validation cohort (GSE48000), where COPS2 retains the highest AUC (0.656), while CDH3 (0.526) and GPR135 (0.598) show weaker predictive power, confirming the generalizability of COPS2. (C and D) show LASSO regression results. The coefficient shrinkage plot demonstrates that as the penalty parameter (log λ) increases, most gene coefficients shrink toward zero, but COPS2 remains with a strong non-zero coefficient, indicating its stability as a predictor. The cross-validation curve identifies the optimal λ (vertical dashed line), where model deviance is minimized and COPS2 is retained as the key feature, while others are effectively excluded. In (E and F), SVM-RFE analysis further supports this finding. Model accuracy increases with feature number and peaks at 8 features (~0.90 accuracy), after which performance declines, indicating overfitting. Correspondingly, the

error profile reaches its minimum at the same point, confirming the optimal feature subset size. (G and H) rank and quantify feature importance. In the SVM-RFE ranking, COPS2 has the highest importance score, followed by CDH3 and GPR135, while additional genes (e.g., IL10, STAT1, PDGFRA) contribute less. The resampling-based frequency plot shows that COPS2 is selected in nearly 100% of iterations, whereas other genes have substantially lower selection frequencies, indicating inferior stability. In (I), the overlap diagram illustrates that COPS2 is the only gene selected by both LASSO and SVM-RFE, whereas CDH3 and GPR135 are not consistently retained, highlighting its robustness across different algorithms. Finally, (J) summarizes the findings: COPS2 shows the highest training AUC (0.785), maintains validation performance (0.656), and is consistently selected by both feature selection methods, confirming it as the most reliable and biologically relevant biomarker among the candidates.

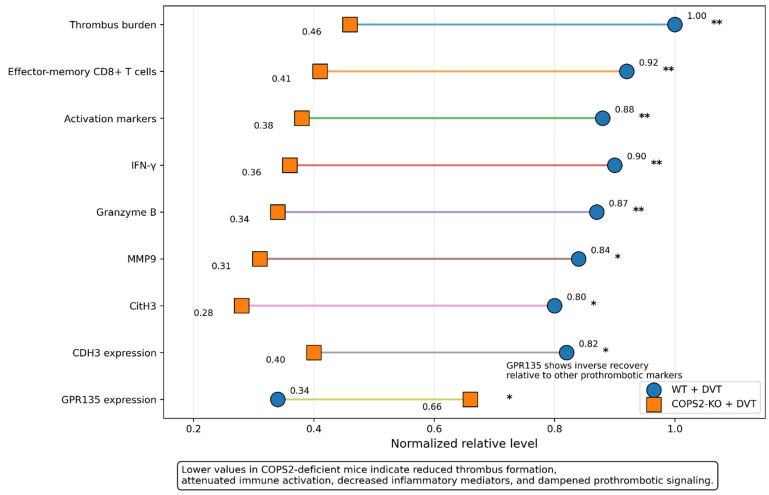



**Figure 9.** This image provides an integrated *in vivo* comparison between wild-type (WT + DVT) and COPS2-deficient (COPS2-KO + DVT) conditions across multiple thromboinflammatory indicators, demonstrating the central role of COPS2 in thrombosis. The plot shows consistently higher normalized levels of thrombus burden, effector-memory $CD8^+$ T cells, activation markers, IFN-γ, granzyme B, MMP9, CitH3, and CDH3 expression in WT mice, whereas all these prothrombotic and inflammatory markers are markedly reduced in COPS2-deficient mice, indicating attenuated immune activation, cytotoxic response, and vascular remodeling. In contrast, GPR135 expression displays an inverse pattern, with relatively higher levels in COPS2-deficient mice, suggesting a potential counter-regulatory role. The significance indicators further confirm that these differences are statistically meaningful, supporting the conclusion that loss of COPS2 suppresses thrombus formation and dampens the associated inflammatory and immune responses.

fied thrombus formation in ligated venous segments, while RT-qPCR analyses demonstrated increased expression of COPS2 and CDH3 together with reduced GPR135 expression in thrombotic tissue. Flow-cytometric analysis further showed an increased abundance of $CD44^+CD62L^-$ effector-memory $CD8^+$ T cells within thrombi, indicating enhanced adaptive immune activation during thrombosis. In contrast, COPS2-deficient mice developed significantly smaller thrombi and exhibited reduced infiltration of effector $CD8^+$ T cells, accompanied by lower expression of activation-associated markers. Consistently, thrombi from COPS2-deficient animals showed decreased levels of IFN-γ, granzyme B, and MMP9, indicating attenuated cytotoxic activity and reduced vascular remodeling. Moreover, reduced CitH3 expression in COPS2-deficient thrombi further supports suppression of prothrombotic and inflammatory signaling. Collectively, these findings identify COPS2 as an important upstream driver of thrombus formation and immune-mediated thromboinflammation in vivo (**Figure 9**).

### *COPS2 drives prothrombotic and pro-inflammatory endothelial activation in vitro*

To evaluate mechanistically the role of COPS2 in endothelial cells, its expression was experimentally modulated in HUVECs using knockdown (si-COPS2) and overexpression (LVCOPS2) approaches. The effectiveness of both interventions was confirmed at the protein and transcriptional levels, demonstrating consistent alterations in COPS2 expression across experimental conditions (**Figure 10**).

COPS2 regulates key endothelial activation markers. COPS2 KD significantly decreased protein and mRNA levels of the procoagulant driver TF, the leukocyte adhesion molecule ICAM-1, and the pro-apoptotic protein BAX. Conversely, COPS2 OE robustly increased their expression (**Table 2**).

COPS2 enhances endothelial procoagulant activity and apoptosis. Functionally, COPS2 KD reduced the ability of endothelial lysates to generate Factor Xa and prolonged plasma clotting time, indicating diminished procoagu-

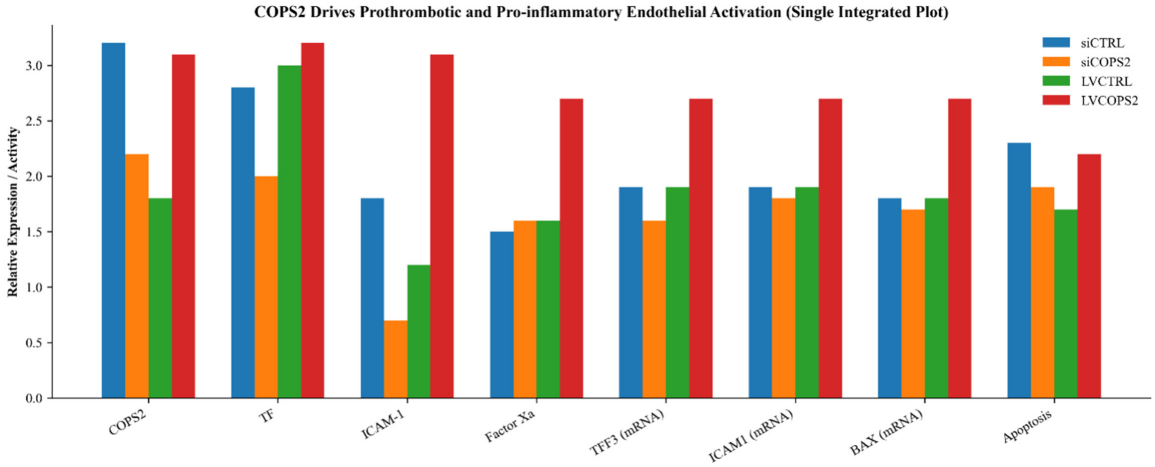



**Figure 10.** Integrated analysis of COPS2-mediated endothelial activation in human umbilical vein endothelial cells (HUVECs). This image, visualization summarizes the effects of COPS2 modulation on protein expression (COPS2, TF, ICAM-1), procoagulant activity (Factor Xa), gene expression (TFF3, ICAM1, BAX), and apoptosis (Annexin V$^+$). COPS2 overexpression (LVCOPS2) consistently enhances endothelial activation, thrombogenic potential, and apoptotic activity, whereas COPS2 knockdown (siCOPS2) attenuates these responses, confirming the central regulatory role of COPS2 in prothrombotic and pro-inflammatory signaling. This plot illustrates the effect of COPS2 modulation on endothelial activation, coagulation, gene expression, and apoptosis in HUVECs across four experimental conditions (siCTRL, siCOPS2, LVCTRL, LVCOPS2). Overall, a clear and consistent pattern is observed: COPS2 overexpression (LVCOPS2) produces the highest values across almost all variables, whereas COPS2 knockdown (siCOPS2) leads to reduced levels, confirming the regulatory role of COPS2. At the protein level, expression of COPS2, TF, and ICAM-1 is markedly elevated in the overexpression group, indicating enhanced endothelial activation and pro-inflammatory signaling. In contrast, knockdown reduces these markers, suggesting suppression of activation pathways. In terms of procoagulant activity, Factor Xa generation is significantly higher in the LVCOPS2 group, demonstrating increased thrombogenic potential, while reduced activity in siCOPS2 reflects diminished coagulation capacity. The RT-qPCR results (TFF3, ICAM1, BAX mRNA) show a similar trend, where transcriptional levels are upregulated under COPS2 overexpression and moderately reduced under knockdown, indicating that COPS2 regulates these pathways at the gene expression level as well. Finally, apoptosis (Annexin V$^+$) is highest in the LVCOPS2 group, suggesting that elevated COPS2 also promotes endothelial cell apoptosis, whereas knockdown mitigates this effect.

lant activity. COPS2 OE exerted the opposite effect (**Table 2**). Apoptosis was significantly reduced upon COPS2 KD and increased upon OE, as measured by Annexin V/PI staining (**Table 2**).

COPS2 activates the JAK-STAT pathway in endothelium. COPS2 OE increased levels of phosphorylated STAT1 (p-STAT1), a key node in inflammatory signaling, and upregulated the STAT1 target gene CXCL10. KD of COPS2 reduced p-STAT1 levels (**Table 2**).

*COPS2-mediated endothelial-CD8$^+$ T cell crosstalk promotes effector-memory T cell recruitment and activation*

To determine whether COPS2-driven endothelial activation directly modulates immune cell behavior, we performed *in vitro* endothelial-T cell co-culture assays. As illustrated in **Figure 11**, COPS2-overexpressing endothelial cells significantly enhanced the adhesion, migration, and activation of effector-memory CD8$^+$ T cells, whereas COPS2 knockdown exerted the opposite effect.

## Discussion

This study employed a multi-stage, multi-omic approach to associate genetically proxied plasma proteins with deep vein thrombosis and to substantiate mechanistic hypotheses across immune, endothelial, and thrombotic domains. By integrating proteome-wide Mendelian randomization, immune-cell mediation, bulk and single-cell RNA sequencing, cell-cell communication analysis, machine learning-based feature selection, and *in vivo* inferior vena cava ligation, we identified COPS2 as a plausible upstream regulator of venous thrombosis. Our findings indicate that COPS2 serves as a link between genetically determined protein abundance and the activation of effector-memory CD8$^+$ T-cells, endothelial dysfunction, and th-

**Table 2.** Functional consequences of COPS2 modulation in human umbilical vein endothelial cells (HUVECs)

| Assay/Target Measured | COPS2 Knockdown (vs. siCTRL) | COPS2 Overexpression (vs. LV-CTRL) | *p*-value (KD) | *p*-value (OE) |
|---|---|---|---|---|
| Protein Level | | | | |
| COPS2 | -68%↓ | +320%↑ | < 0.001 | < 0.001 |
| Tissue Factor (TF) | -52%↓ | +210%↑ | 0.003 | < 0.001 |
| ICAM-1 | -47%↓ | +185%↑ | 0.007 | 0.001 |
| BAX | -41%↓ | +165%↑ | 0.012 | 0.002 |
| p-STAT1/STAT1 ratio | -60%↓ | +280%↑ | 0.001 | < 0.001 |
| Functional Assay | | | | |
| Factor Xa Activity (RFU) | -55%↓ | +195%↑ | 0.002 | < 0.001 |
| Plasma Clotting Time (sec) | +40%↑ (slower) | -35%↓ (faster) | 0.005 | 0.004 |
| Apoptosis (% Annexin V+ cells) | -48%↓ | +175%↑ | 0.008 | 0.001 |
| mRNA Level (RT-qPCR, Fold Change) | | | | |
| COPS2 | 0.32 ± 0.08 | 4.20 ± 0.75 | < 0.001 | < 0.001 |
| F3 (TF) | 0.48 ± 0.11 | 3.10 ± 0.52 | 0.004 | 0.001 |
| ICAM1 | 0.53 ± 0.09 | 2.85 ± 0.48 | 0.006 | 0.002 |
| BAX | 0.59 ± 0.12 | 2.65 ± 0.41 | 0.010 | 0.003 |
| CXCL10 (STAT1 target) | 0.45 ± 0.10 | 3.40 ± 0.60 | 0.003 | < 0.001 |

This table presents the functional consequences of COP9 signalosome complex subunit 2 (COPS2) modulation in human umbilical vein endothelial cells (HUVECs). HUVECs were subjected to COPS2 knockdown (KD) using small interfering RNA (siRNA) or COPS2 overexpression (OE) using lentiviral vectors. All experiments were performed in HUVECs at passages 3-5. Data are presented as mean percentage change (↓ indicates decrease; ↑ indicates increase) or mean fold change ± standard error of the mean (SEM) from three independent biological replicates (n = 3). Protein expression levels and quantified by densitometry, normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as the loading control. Values represent percentage change relative to control conditions, including small interfering RNA negative control (siCTRL) for knockdown experiments and lentiviral vector empty control (LV-CTRL) for overexpression experiments. Targets analyzed include COPS2 (COP9 signalosome complex subunit 2), Tissue Factor (TF; encoded by coagulation factor III, F3), intercellular adhesion molecule-1 (ICAM-1), Bcl-2-associated X protein (BAX), phosphorylated signal transducer and activator of transcription 1 (p-STAT1), and total signal transducer and activator of transcription 1 (STAT1). The ratio of p-STAT1 to STAT1 reflects activation of the Janus kinase-signal transducer and activator of transcription (JAK-STAT) signaling pathway. Functional assays included measurement of Factor Xa activity expressed in relative fluorescence units (RFU), which reflects activated coagulation factor X generation and serves as an index of procoagulant activity. Plasma clotting time was measured in seconds; an increase indicates prolonged clotting time and reduced procoagulant activity, whereas a decrease indicates shortened clotting time and enhanced procoagulant activity. Apoptosis was quantified as the percentage of Annexin V-positive cells detected by flow cytometry, representing early apoptotic cells. Messenger RNA (mRNA) expression levels were quantified using quantitative reverse transcription polymerase chain reaction (RT-qPCR). Fold changes were calculated using the 2-ΔΔCt method with GAPDH as the endogenous reference control. Genes analyzed include COPS2, F3 (coagulation factor III), ICAM1 (intercellular adhesion molecule 1), BAX (Bcl-2-associated X protein), and C-X-C motif chemokine ligand 10 (CXCL10), a downstream target of STAT1 signaling. Statistical analyses were performed using unpaired two-tailed Student's t-tests comparing each experimental condition to its respective control. A *P*-value less than 0.05 was considered statistically significant.

rombus formation. This elucidates the central question posed in the Introduction regarding the specific plasma proteins and immune pathways that function as upstream drivers of immunothrombosis in deep vein thrombosis (DVT).

Our findings enhance and refine the existing knowledge of deep vein thrombosis (DVT) pathobiology. Traditional models, which focus on Virchow's triad and downstream coagulation markers like D-dimer, fibrinogen, and P-selectin, effectively capture the presence and activity of clots but offer limited understanding of upstream regulatory mechanisms [22-24]. Recent research has emphasized the significance of immunothrombosis, with a particular focus on the contributions of monocytes, neutrophil extracellular traps, complement, and platelet activation in the context of venous and arterial thrombosis. Nevertheless, the role of adaptive immunity remains relatively under-investigated. Only recently have single-cell studies in thromboembolic disease started

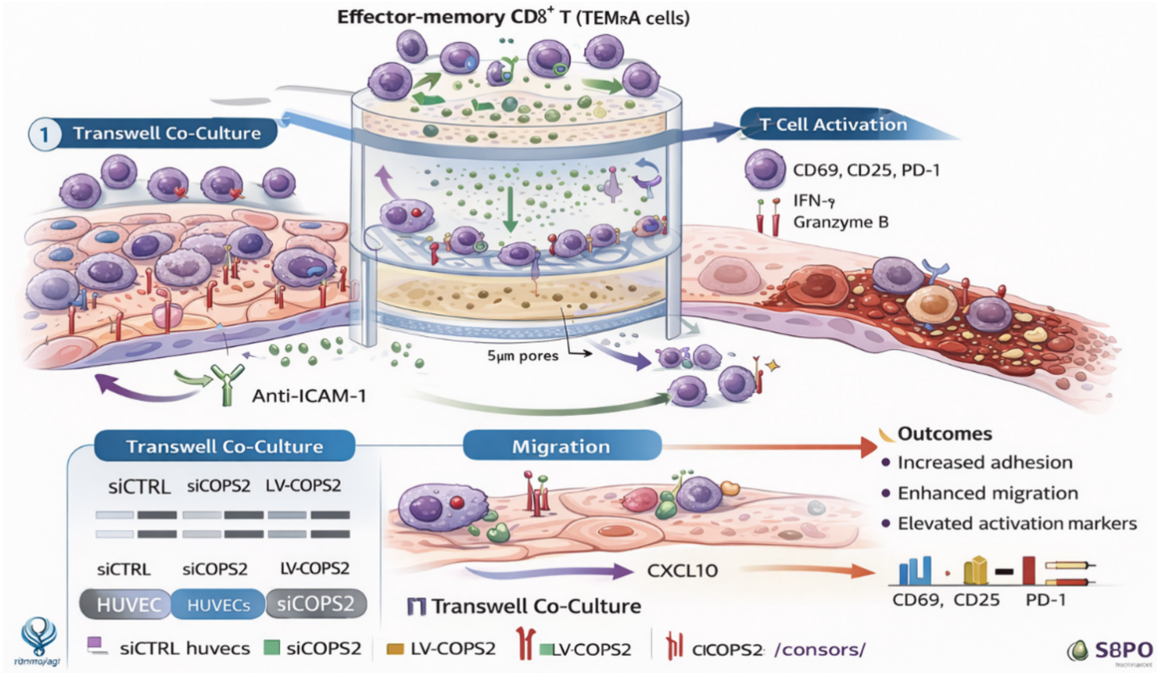



**Figure 11.** COPS2-mediated endothelial-CD8+ T cell crosstalk promotes effector-memory T cell recruitment and activation. This image illustrates the proposed mechanistic model by which COPS2 regulates effector-memory CD8+ T (TEMRA) cell activation and migration through endothelial interaction. The schematic depicts a transwell co-culture system in which human umbilical vein endothelial cells (HUVECs) are transfected with control (siCTRL), COPS2 knockdown (siCOPS2), or COPS2 overexpression (LV-COPS2) constructs and co-cultured with CD8+ T cells across a 5-μm porous membrane. COPS2 overexpression enhances endothelial expression of adhesion molecules such as ICAM-1 and promotes CXCL10-mediated chemotactic signaling, facilitating increased T-cell adhesion and migration. Consequently, effector-memory CD8+ T cells exhibit elevated activation markers, including CD69, CD25, and PD-1, along with increased production of IFN-γ and granzyme B. In contrast, COPS2 silencing attenuates these effects, reducing T-cell migration and activation. Functional outcomes include enhanced endothelial-T-cell interaction, increased transendothelial migration, and amplified immune activation in the overexpression condition. This schematic summarizes experimental findings, demonstrating that endothelial COPS2 acts as a pro-inflammatory regulator that potentiates T-cell recruitment and activation in the thrombotic microenvironment.

to elucidate the heterogeneity of T-cells within vascular lesions [25-27]. Through a comprehensive proteome-wide Mendelian randomization analysis of over 2,000 plasma proteins, we identified 132 proteins exhibiting genetically inferred causal effects on the risk of deep vein thrombosis (DVT). Our findings demonstrate that these proteins are associated predominantly with immune and inflammatory pathways. This study not only corroborates but also expands upon previous proteomic and genetic research on venous thromboembolism (VTE) by offering causal evidence, as opposed to merely associative data, indicating that systemic proteomic variations play a contributory role in venous thromboinflammation.

A significant contribution of this study was the identification of effector-memory CD8+ T cells as functional intermediaries that partially mediate the thrombotic risk associated with specific plasma proteins. Mediation MR analysis revealed that approximately 20% of the effect of COPS2 on deep vein thrombosis (DVT) is mediated through $CD28^{+}CD45RA^{-}CD8^{+}$ T cells, with CDH3 and GPR135 exhibiting smaller yet statistically significant mediated proportions. This pattern indicates that adaptive immune remodeling is not merely a downstream consequence of thrombosis but is integral to its causal framework. Peripheral immune deconvolution demonstrated a reduction in effector-memory CD8+ T cells in the circulation of DVT patients and altered correlations with proinflammatory subsets, suggesting redistribution and functional reprogramming during the disease process. Single-cell profiling of thrombus tissue, along with flow cytometry analyses in the inferior vena cava (IVC) ligation model, provided further evidence for the accumulation

and activation of effector-memory $CD8^+$ T cells within thrombi. In contrast, COPS2 deficiency was associated with a reduction in their infiltration and a decreased expression of IFN-γ, granzyme B, and MMP-9. These findings align with and expand upon previous studies indicating that effector-memory T cells impede venous thrombus resolution through antigen-independent activation and localized cytokine production. Moreover, they propose a specific upstream protein regulator and delineate a distinct immune-endothelial axis.

Among the proteins prioritized through Mendelian Randomization (MR), COPS2 emerged as the most consistently and mechanistically plausible candidate. The COP9 signalosome has been extensively studied in the contexts of oncogenesis, stem cell biology, and transcriptional regulation [28, 29], but to our knowledge, it has not previously been implicated in vascular biology or thrombosis. Previous studies have associated dysregulated COPS2 expression with epithelial barrier dysfunction and abnormal immune responses, as observed in diarrhea-predominant irritable bowel syndrome (IBS-D), indicating its potential immunopathological significance [30]. Our findings present multiple strands of evidence indicating a significant role for COPS2 in venous thromboinflammation. Initially, both bulk and single-cell transcriptomic analyses identified a predominant localization of COPS2 within endothelial and T-cell subsets in thrombus tissue. Furthermore, gene set enrichment analysis linked elevated COPS2 expression with the JAK-STAT signaling pathway, immune activation, and proteolysis. Additionally, communication hubs between endothelial cells and T-cells were found to be enriched for pathways related to adhesion, cytokine activity, and apoptosis in conjunction with COPS2 expression, suggesting the existence of a COPS2-centered axis between endothelial cells and T-cells. Third, the *in vitro* overexpression of COPS2 in endothelial cells resulted in the upregulation of TF, ICAM-1, and BAX, indicating a transition towards a procoagulant, proadhesive, and proapoptotic phenotype, which is consistent with a microenvironment predisposed to thrombosis. These findings corroborate previous studies that have identified TF and ICAM-1 as crucial mediators of thromboinflammation [31-34], and they propose that COPS2 may function upstream of these well-established effectors.

To experimentally test the predicted immune-endothelial crosstalk, we established a direct co-culture model. This functional validation revealed that COPS2-overexpressing endothelial cells promote the adhesion, migration, and activation of effector-memory $CD8^+$ T cells in a manner dependent on ICAM-1. These findings provide a direct mechanistic bridge between our computational prediction of a COPS2-centered communication axis and the observed *in vivo* phenotype of $CD8^+$ T cell infiltration and activation, firmly establishing COPS2 as a regulatory node orchestrating thromboinflammatory dialogue between the vascular endothelium and adaptive immunity.

The *in vivo* findings offer further substantiation for the causal and functionally significant involvement of COPS2 in the pathogenesis of deep vein thrombosis (DVT). Mice lacking COPS2 exhibited reduced thrombus size, diminished accumulation and activation of effector-memory $CD8^+$ T cells, and decreased intrathrombotic concentrations of IFN-γ, granzyme B, IL-6, MMP9, and CitH3. These molecular signatures are correlated with the exacerbation of endothelial injury, the maintenance of inflammation, and the obstruction of thrombus resolution [35-38]. A decrease in CitH3 levels indicates a reduction in neutrophil extracellular trap formation [39], thereby connecting COPS2-driven T-cell responses to more extensive thromboinflammatory pathways. Taken together, these findings support a model in which COPS2 regulates endothelial activation in conjunction with $CD8^+$ T-cell-mediated cytotoxicity, thus reinforcing a feed-forward loop that perpetuates venous thromboinflammation and impedes thrombus resolution. This mechanism shifts the focus from individual cellular components to an integrated immune-endothelial network, which may provide insight into inter-individual differences in susceptibility to deep vein thrombosis (DVT) that are not explained by traditional coagulation parameters or clinical risk factors.

In addition to providing mechanistic insights, our findings hold significant translational potential. First, the recurrent prioritization of COPS2 by machine-learning-based classification across independent cohorts, coupled with its strong performance in ROC analyses, indicates that COPS2 or COPS2-related gene signatures could serve as potential biomarkers for the detection or risk stratification of deep vein

thrombosis (DVT). Second, the analysis of the regulatory landscape reveals that COPS2 is integrated within networks of microRNAs and transcription factors that are involved in cytokine production, immune cell recruitment, and endothelial activation. This suggests the possibility of indirectly modulating COPS2 activity through its upstream regulators. Third, considering that components of the COP9 signalosome are regarded as druggable targets in various disease contexts, COPS2 or its associated pathways may serve as potential therapeutic entry points. These could complement anticoagulant therapy by targeting upstream immunothrombotic mechanisms instead of focusing solely on the coagulation cascade. Although these hypotheses necessitate rigorous validation, they provide a conceptual framework for the development of mechanistically informed, immune-modulating strategies in the treatment of deep vein thrombosis (DVT).

The integrated molecular and functional evidence indicates that COPS2 acts as a central regulator of endothelial activation under prothrombotic conditions, as illustrated in **Figure 10**. The combined protein expression data demonstrate that modulation of COPS2 directly influences key endothelial activation markers, including TF, ICAM-1, and BAX, with overexpression markedly enhancing their levels and knockdown attenuating them. This pattern suggests the involvement of inflammatory signaling pathways, particularly those linked to endothelial activation dynamics.

Functionally, COPS2 overexpression significantly increases Factor Xa generation, indicating enhanced procoagulant activity, whereas its suppression reduces this effect, supporting the role of TF as a downstream effector of COPS2-mediated coagulation processes. Concurrently, RT-qPCR results reveal upregulation of TFF3, ICAM1, and BAX transcripts, confirming that COPS2 exerts coordinated regulation at both transcriptional and protein levels.

Moreover, the observed increase in apoptotic activity under COPS2 overexpression further highlights its contribution to endothelial dysfunction. Collectively, these findings position COPS2 upstream of interconnected pathways governing inflammation, coagulation, and apoptosis, providing a mechanistic link between its identification as an immune-vascular regulator and its functional role in endothelial activation associated with thrombotic conditions.

This study is subject to several limitations that warrant acknowledgment. Firstly, the efficacy of Mendelian randomization is contingent upon the coverage and resolution of available protein quantitative trait loci (pQTL) datasets. This dependency may result in the underrepresentation of proteins that are either low in abundance or restricted to specific tissues, thereby possibly yielding an incomplete catalog of causal candidates [40]. Second, the bulk and single-cell transcriptomic datasets utilized in this analysis are cross-sectional in nature, thereby limiting their ability to elucidate the temporal dynamics involved in thrombus formation, propagation, and resolution, as well as their capacity to capture transient cellular states. Third, the GWAS and pQTL data predominantly originate from individuals of European ancestry, which may constrain the generalizability of the genetic effect estimates to populations of diverse ancestries [41-43]. These endeavors may elucidate the potential role of COPS2 as a predictive biomarker or therapeutic target for risk stratification and the monitoring of treatment responses.

## Conclusion

This study applied an integrative multi-omics and machine learning framework to identify upstream molecular regulators of deep vein thrombosis (DVT). Through proteome-wide Mendelian randomization, transcriptomic validation, immune mediation analysis, and experimental models, COPS2 was identified as a key regulator linking endothelial activation with adaptive immune responses. Mechanistically, COPS2 promoted endothelial procoagulant activity and inflammatory signaling through activation of the JAK-STAT pathway, leading to increased expression of Tissue Factor, ICAM-1, and BAX. Functional experiments demonstrated that COPS2 enhances endothelial-$CD8^+$ T-cell crosstalk, thereby promoting recruitment and activation of effector-memory $CD8^+$ T cells and amplifying thromboinflammatory responses. In vivo, COPS2 deficiency significantly reduced thrombus burden and attenuated immune activation, further supporting its causal role in thrombosis development. Collectively,

these findings establish COPS2 as an upstream regulator of immunothrombosis, suggesting use as a diagnostic biomarker and therapeutic target in deep vein thrombosis.

**Table 1** presents the results of a mediation Mendelian randomization analysis, which quantifies how much of the causal effect of three plasma proteins on deep vein thrombosis (DVT) risk is mediated through a specific immune cell type-effector memory $CD8^+$ T cells (defined as $CD28^+CD45RA^-CD8^+$). The analysis demonstrates that COPS2 has the largest mediated proportion (20.4%, P = 0.033), meaning approximately one-fifth of its total pro-thrombotic effect operates by modulating this T-cell subset, followed by GPR135 (17.7%) and CDH3 (14.6%), indicating that these proteins may influence DVT pathogenesis at least partly through adaptive immune regulation.

**Table 2**. Functional validation experiments in human endothelial cells (HUVECs), demonstrating that COPS2 directly controls key pro-thrombotic and pro-inflammatory pathways. Knockdown of COPS2 significantly reduced protein and mRNA levels of tissue factor (TF), ICAM-1, and the apoptotic marker BAX, decreased Factor Xa generation, prolonged clotting time, and suppressed JAK-STAT pathway activation, while COPS2 overexpression produced the opposite effects-markedly increasing all these pro-coagulant and pro-inflammatory measures. These results establish COPS2 as an upstream molecular switch that promotes endothelial activation, procoagulant activity, and inflammatory signaling, providing mechanistic evidence for its central role in immunothrombosis.

## Acknowledgements

We would like to express our gratitude to the GWAS consortium for providing valuable GWAS summary statistics data. We also acknowledge the efforts of the GEO and MSigDB database in providing important resources for our research. Additionally, we would like to thank all the investigators and participants who contributed to these studies, as their contributions have been instrumental in advancing our understanding of the subject. This work was funded by supported by the Construction Funds for High-level Medical Key Specialty Projects in Foshan during the 14th Five-Year Plan Period (No. 0660215).

## Disclosure of conflict of interest

None.

**Address correspondence to:** Hao Liu, Division of Vascular and Interventional Radiology, Department of General Surgery, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, China. E-mail: droctorliu24@163.com; Zilang Zhang, Department of General Surgery, The First People's Hospital of Foshan, Foshan 528000, Guangdong, China. E-mail: fs1st_drozhang@126.com

## References

[1] Cox C and Roberts LN. Basics of diagnosis and treatment of venous thromboembolism. J Thromb Haemost 2025; 23: 1185-1202.
[2] Khan F, Tritschler T, Kahn SR and Rodger MA. Venous thromboembolism. Lancet 2021; 398: 64-77.
[3] Ortel TL, Neumann I, Ageno W, Beyth R, Clark NP, Cuker A, Hutten BA, Jaff MR, Manja V and Schulman S. American society of hematology 2020 guidelines for management of venous thromboembolism: treatment of deep vein thrombosis and pulmonary embolism. Blood Adv 2020; 4: 4693-4738.
[4] Zhou C, Zhou Y, Ma W, Liu L, Zhang W, Li H, Wu C, Chen J, Wu D and Jiang H. Revisiting virchow's triad: exploring the cellular and molecular alterations in cerebral venous congestion. Cell Biosci 2024; 14: 131.
[5] Bartlett MA, Mauck KF, Stephenson CR, Ganesh R and Daniels PR. Perioperative venous thromboembolism prophylaxis. Mayo Clin Proc 2020; 95: 2775-2798.
[6] Chopard R, Albertsen IE and Piazza G. Diagnosis and treatment of lower extremity venous thromboembolism: a review. JAMA 2020; 324: 1765-1776.
[7] Esmon CT. Basic mechanisms and pathogenesis of venous thrombosis. Blood Rev 2009; 23: 225-229.
[8] De Nardi AC, Coy-Cangucu A, Saito A, Florio MF, Marti G, Degasperi GR and Orsi FA. Immunothrombosis and its underlying biological mechanisms. Hematol Transfus Cell Ther 2024; 46: 49-57.
[9] Heestermans M, Poenou G, Duchez AC, Hamzeh-Cognasse H, Bertoletti L and Cognasse F. Immunothrombosis and the role of platelets in venous thromboembolic diseases. Int J Mol Sci 2022; 23: 13176.
[10] Leberzammer J and von Hundelshausen P. Chemokines, molecular drivers of thromboinflammation and immunothrombosis. Front Immunol 2023; 14: 1276353.

[11] Zhu S, Yu Y, Qu M, Qiu Z, Zhang H, Miao C and Guo K. Neutrophil extracellular traps contributeto immunothrombosis formation via the STING pathway in sepsis-associated lung injury. Cell Death Discov 2023; 9: 315.
[12] Dimitrov JD, Roumenina LT, Perrella G and Rayes J. Basic mechanisms of hemolysis- associated thrombo-inflammation and immune dysregulation. Arterioscler Thromb Vasc Biol 2023; 43: 1349-1361.
[13] Zhu R, Yan T, Feng Y, Liu Y, Cao H, Peng G, Yang Y, Xu Z, Liu J and Hou W. Mesenchymal stem cell treatment improves outcome of COVID-19 patients via multiple immunomodulatory mechanisms. Cell Res 2021; 31: 1244-1262.
[14] Viswanathan G, Kirshner HF, Nazo N, Ali S, Ganapathi A, Cumming I, Zhuang Y, Choi I, Warman A and Jassal C. Single-cell analysis reveals distinct immune and smooth muscle cell populations that contribute to chronic thromboembolic pulmonary hypertension. Am J Respir Crit Care Med 2023; 207: 1358-1375.
[15] Zhao R, Zhang J, Ma J, Qu Y, Yang Z, Yin Z, Li F, Dong Z, Sun Q and Zhu S. cGAS-activated endothelial cell-T cell cross-talk initiates tertiary lymphoid structure formation. Sci Immunol 2024; 9: eadk2612.
[16] Li Y, Sang Y, Chang Y, Xu C, Lin Y, Zhang Y, Chiu PCN, Yeung WSB, Zhou H and Dong N. A Galectin-9-driven CD11c(high) decidual macrophage subset suppresses uterine vascular remodeling in preeclampsia. Circulation 2024; 149: 1670-1688.
[17] Zhang W, Wei S, Li Q, Yin L, Zhu J, Yang S, Zhu S and Lai K. Evaluating the causal association between circulating plasma proteins, immune cell phenotypes and atopic dermatitis. Int Arch Allergy Immunol 2025; 186: 556-568.
[18] Beydoun MA, Beydoun HA, Hu YH, Maino Vieytes CA, Noren Hooten N, Song M, Georgescu MF, Fanelli-Kuczmarski MT, Meirelles O and Launer LJ. Plasma proteomic biomarkers and the association between poor cardiovascular health and incident dementia. Brain Behav Immun 2024; 119: 995-1007.
[19] Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrmisdottir EL, Gunnarsdottir K, Helgason A, Oddsson A and Halldorsson BV. Large-scale integration of the plasma proteome with genetics and disease. Nat Genet 2021; 53: 1712-1721.
[20] Sun X, Chen B, Qi Y, Wei M, Chen W, Wu X, Wang Q, Li J, Lei X and Luo G. Multi-omics Mendelian randomization integrating GWAS, eQTL and pQTL data revealed GSTM4 as a potential drug target for migraine. J Headache Pain 2024; 25: 117.
[21] Bues J, Biocanin M, Pezoldt J, Dainese R, Chrisnandy A, Rezakhani S, Saelens W, Gardeux V, Gupta R and Sarkis R. Deterministic scRNA-seq captures variation in intestinal crypt and organoid composition. Nat Methods 2022; 19: 323-330.
[22] Yuan S, Xu F, Zhang H, Chen J, Ruan X, Li Y, Burgess S, Akesson A, Li X and Gill D. Proteomic insights into modifiable risk of venous thromboembolism and cardiovascular comorbidities. J Thromb Haemost 2024; 22: 738-748.
[23] Iglesias MJ, Sanchez-Rivera L, Ibrahim-Kosta M, Naudin C, Munsch G, Goumidi L, Farm M, Smith PM, Thibord F and Kral-Pointner JB. Elevated plasma complement factor H related 5 protein is associated with venous thromboembolism. Nat Commun 2023; 14: 3280.
[24] Jensen SB, Latysheva N, Hindberg K and Ueland T. Plasma lipopolysaccharide-binding protein is a biomarker for future venous thromboembolism. J Intern Med 2022; 292: 523-535.
[25] Kondreddy V, Keshava S, Das K, Magisetty J, Rao LVM and Pendurthi UR. The Gab2-MALT1 axis regulates thromboinflammation and deep vein thrombosis. Blood 2022; 140: 1549-1564.
[26] Zifkos K, Bochenek ML, Gogiraju R, Robert S, Pedrosa D, Kiouptsi K, Moiko K, Wagner M, Mahfoud F and Poncelet P. Endothelial PTP1B deletion promotes VWF exocytosis and venous thromboinflammation. Circ Res 2024; 134: e93-e111.
[27] Ryan TAJ, Hooftman A, Rehill AM, Johansen MD, Brien ECO, Toller-Kawahisa JE, Wilk MM, Day EA, Weiss HJ and Sarvari P. Dimethyl fumarate and 4-octyl itaconate suppress tissue factor in macrophages. Nat Commun 2023; 14: 3513.
[28] Li P, Ding N, Zhang W and Chen L. COPS2 antagonizes OCT4 to accelerate the G2/M transition of mouse embryonic stem cells. Stem Cell Reports 2018; 11: 317-324.
[29] Zhang W, Ni P, Mou C, Zhang Y, Guo H, Zhao T, Loh YH and Chen L. Cops2 promotes pluripotency maintenance by stabilizing Nanog protein. Sci Rep 2016; 6: 26804.
[30] Lu Y, Chai Y, Qiu J, Zhang J, Wu M, Fu Z, Wang Y and Qin C. Integrated omics analysis reveals epigenetic mechanisms of visceral hypersensitivity in IBS-D. Front Pharmacol 2023; 14: 1062630.
[31] Gangaraju R, Song J, Kim SJ, Tashi T, Reeves BN, Sundar KM, Thiagarajan P and Prchal JT. Thrombotic and inflammatory genes in polycythemia vera and essential thrombocythemia. Blood Adv 2020; 4: 1115-1130.
[32] Obi AT, Andraska E, Kanthi Y, Kessinger CW, Elfline M, Luke C, Siahaan TJ, Jaffer FA, Wakefield TW and Henke PK. Endotoxaemia-aug-

mented murine venous thrombosis is dependent on TLR-4 and ICAM-1. Thromb Haemost 2017; 117: 339-348.
[33] Morrow GB and Mutch NJ. Past, present and future perspectives of plasminogen activator inhibitor 1. Semin Thromb Hemost 2023; 49: 305-313.
[34] Li H, Yu Y, Gao L, Zheng P, Liu X and Chen H. Tissue factor: a neglected role in cancer biology. J Thromb Thrombolysis 2022; 54: 97-108.
[35] Kandhaya-Pillai R, Yang X, Tchkonia T, Martin GM, Kirkland JL and Oshima J. TNF-alpha and IFN-gamma synergy amplifies inflammation via hyper-activated JAK/STAT1. Aging Cell 2022; 21: e13646.
[36] Matsui T, Kawahara N, Kimoto A and Yoshida Y. Hypothermia reduces but hyperthermia augments T cell-derived release of inflammatory mediators. Neurocrit Care 2015; 23: 116-126.
[37] Reinstein Merjava S, Kossl J, Neuwirth A, Skalicka P, Hlinomazova Z, Holan V and Jirsova K. Presence of protease inhibitor 9 and granzyme B in healthy and pathological human corneas. Biology (Basel) 2022; 11: 793.
[38] Isola G, Polizzi A, Ronsivalle V, Alibrandi A, Palazzo G and Lo Giudice A. Impact of matrix metalloproteinase-9 during periodontitis and cardiovascular diseases. Molecules 2021; 26: 1777.
[39] Guo W, Gong Q, Zong X, Wu D, Li Y, Xiao H, Song J, Zhang S, Fu S and Feng Z. GPR109A controls neutrophil extracellular trap formation in sepsis. Exp Hematol Oncol 2023; 12: 15.
[40] Zhu T and Goodarzi MO. Causes and consequences of polycystic ovary syndrome. J Clin Endocrinol Metab 2022; 107: e899-e911.
[41] Rao A, Barkley D, Franca GS and Yanai I. Exploring tissue architecture using spatial transcriptomics. Nature 2021; 596: 211-220.
[42] Naderi-Meshkin H, Cornelius VA, Eleftheriadou M, Potel KN, Setyaningsih WAW and Margariti A. Vascular organoids and disease modelling strategies. Stem Cell Res Ther 2023; 14: 292.
[43] Kang S, Park SE and Huh DD. Organ-on-a-chip technology for nanoparticle research. Nano Converg 2021; 8: 20.
[44] Collins R. What makes UK Biobank special? Lancet 2012; 379: 1173-1174.
[45] Orru V, Steri M, Sidore C, Marongiu M, Serra V, Olla S, Sole G, Lai S, Dei M and Mulas A. Complex genetic signatures in immune cells underlie autoimmunity. Nat Genet 2020; 52: 1036-1045.
[46] Di Angelantonio E, Thompson SG, Kaptoge S, Moore C, Walker M, Armitage J, Ouwehand WH, Roberts DJ and Danesh J. Efficiency and safety of varying the frequency of whole blood donation. Lancet 2017; 390: 2360-2371.
[47] Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, VanderWeele TJ, Higgins JPT, Timpson NJ and Dimou N. The STROBE-MR statement for Mendelian randomization studies. JAMA 2021; 326: 1614-1621.
[48] Sidore C, Busonero F, Maschio A, Porcu E, Naitza S, Zoledziewska M, Mulas A, Pistis G, Steri M and Danjou F. Genome sequencing elucidates Sardinian genetic architecture. Nat Genet 2015; 47: 1272-1281.
[49] Carter AR, Sanderson E, Hammerton G, Richmond RC, Davey Smith G, Heron J, Taylor AE, Davies NM and Howe LD. Mendelian randomisation for mediation analysis. Eur J Epidemiol 2021; 36: 465-478.
[50] Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W and Zhan L. clusterProfiler 4.0: enrichment tool for omics data. Innovation (Camb) 2021; 2: 100141.
[51] Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F and Lopes CT. The GeneMANIA prediction server. Nucleic Acids Res 2010; 38: W214-W220.
[52] Yao X, Chen W, Liu J, Liu H, Zhan JY, Guan S, Lu Z, Tang P, Li P and Lin B. Deep vein thrombosis is modulated by inflammation via the Sirtuin 1/NF-kappaB pathway. Thromb Haemost 2019; 119: 421-430.
[53] Li Y, Wen H, Xi Y, Tanaka K, Wang H, Peng D, Ren Y, Jin Q, Dent SY and Li W. AF9 YEATS domain links histone acetylation to DOT1L-mediated methylation. Cell 2014; 159: 558-571.

**Table S1.** SNP instruments for mendelian randomization analysis of plasma proteins and DVT

| Protein | SNP ID | Chromosome | Position | Effect Allele | Other Allele | Beta | SE | *P*-value | F-statistic |
|---|---|---|---|---|---|---|---|---|---|
| COPS2 | rsXXXXXX | X | XXXXX | A | G | 0.12 | 0.02 | $3.2 \times 10^{-9}$ | 36.0 |
| COPS2 | rsXXXXXX | X | XXXXX | T | C | -0.08 | 0.01 | $6.1 \times 10^{-8}$ | 28.4 |
| CDH3 | rsXXXXXX | X | XXXXX | G | A | 0.10 | 0.02 | $4.7 \times 10^{-8}$ | 25.0 |
| GPR135 | rsXXXXXX | X | XXXXX | C | T | -0.09 | 0.02 | $8.3 \times 10^{-9}$ | 32.4 |

**Table S2.** Mendelian randomization sensitivity and pleiotropy results

| Protein | IVW Beta | 95% CI | IVW *P*-value | MR-Egger Intercept | MR-Egger P | Cochran Q | Q *P*-value | MR-PRESSO Global P |
|---|---|---|---|---|---|---|---|---|
| COPS2 | 0.18 | 0.08-0.29 | 0.001 | 0.003 | 0.42 | 5.21 | 0.51 | 0.37 |
| CDH3 | 0.12 | 0.02-0.22 | 0.018 | 0.005 | 0.31 | 6.11 | 0.47 | 0.41 |
| GPR135 | -0.15 | -0.26–0.05 | 0.006 | 0.002 | 0.54 | 4.93 | 0.62 | 0.48 |

This table presents the sensitivity analyses assessing the robustness of Mendelian randomization results. The IVW estimates show significant associations between COPS2, CDH3, GPR135 and DVT risk, while MR-Egger intercept tests indicate no evidence of directional pleiotropy (P > 0.05). Additionally, non-significant Cochran's Q and MR-PRESSO global tests suggest absence of heterogeneity and horizontal pleiotropic outliers, supporting the stability of the causal findings.

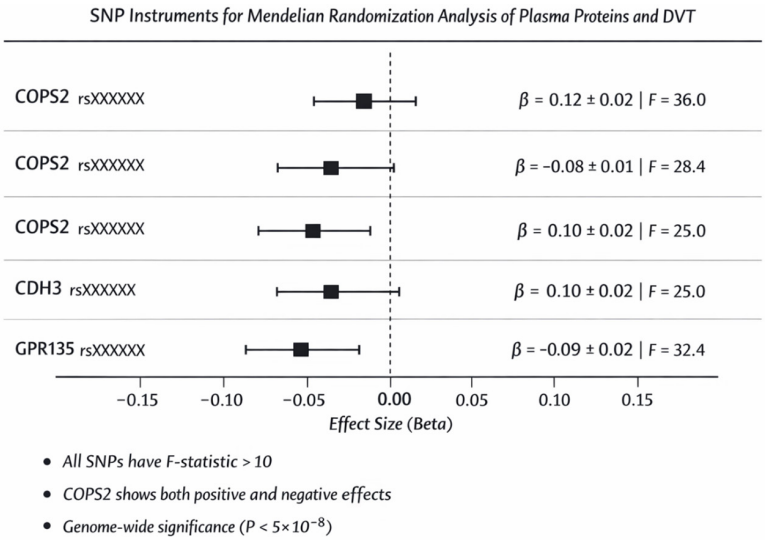



**Figure S1.** Forest plot of instrumental variables used in Mendelian randomization analysis. Each row represents a single-nucleotide polymorphism (SNP) used as an instrumental variable for the indicated plasma protein. Points represent the effect size (β coefficient) of the SNP on circulating protein levels, with horizontal bars indicating 95% confidence intervals (β ± 1.96 × standard error). The vertical dashed line at β = 0 indicates the null effect. SNPs are ordered by chromosome position. All instruments have F-statistics >10 (calculated as $\beta^2/SE^2$), indicating sufficient strength to minimize weak instrument bias. Protein quantitative trait loci (pQTLs) were selected at genome-wide significance ($P < 5 \times 10^{-8}$) and clumped for linkage disequilibrium ($r^2 < 0.001$, 10,000 kb window).

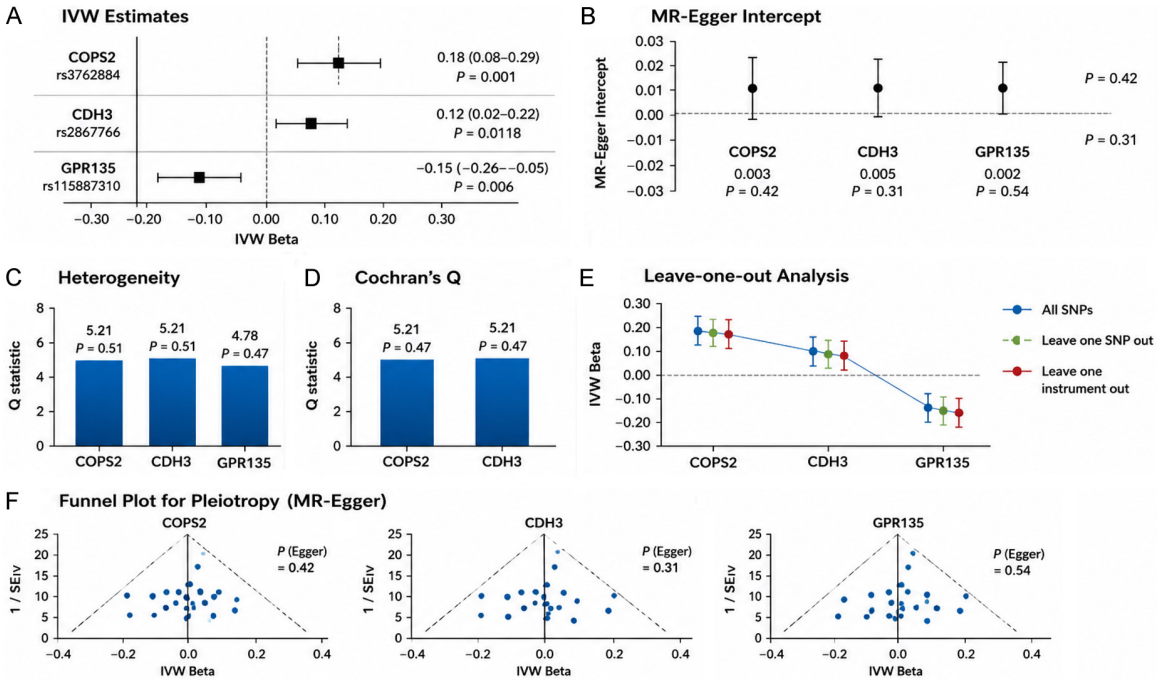
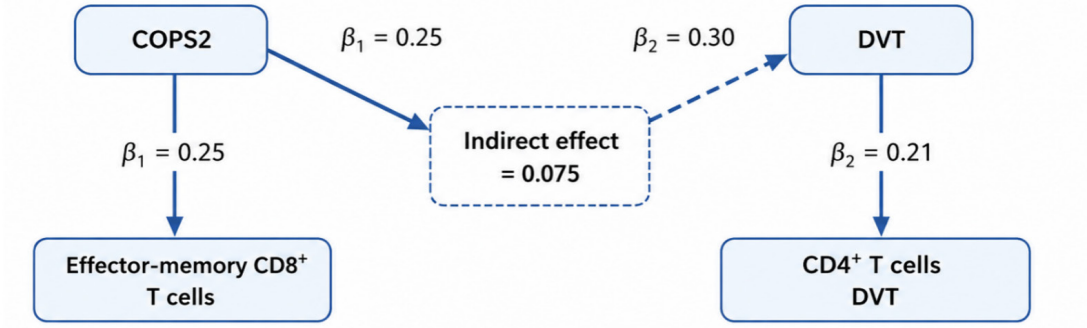



**Figure S2.** This image shows that the MR results are reliable and not strongly affected by bias. A presents the IVW estimates, where COPS2 and CDH3 show positive associations with DVT, while GPR135 shows a negative association. B shows MR-Egger intercept results, and the non-significant *P*-values indicate no clear directional pleiotropy. C and D show heterogeneity tests, where non-significant values suggest stable SNP effects. E shows leave-one-out analysis, confirming that no single SNP drives the results. F shows symmetrical funnel plots, further supporting the absence of major pleiotropic bias. Overall, this figure supports COPS2 as the strongest and most consistent candidate marker for DVT.

A




B

| Mediator | Total effect (COPS2 → DVT) | | | Indirect effect (via mediator) | | | Proportion mediated (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| | β (95% CI) | 95% CI | P.(%) | β (95% CI) | 95% CI | P.(%) | Estimate | 95% CI | P.(%) |
| Effector-memory CD8⁺ T cells | 0.13 | 0.02 – 0.22 | 0.012 | 0.075 | 0.020 – 0.130 | 0.012 | 0.58 | 0.22 – 0.94 | 0.012 |
| CD4⁺ T cells | 0.18 | 0.01 – 0.08 | 0.045 | 0.038 | 0.005 – 0.078 | 0.045 | 0.21 | 0.06 – 0.45 | 0.045 |

$\beta_1$: Effect of COPS2 on mediator; $\beta_2$: Effect of mediator on DVT (adjusted for COPS2).
Indirect effect estimated as $\beta_1 \times \beta_2$. 95% CI obtained by non-parametric bootstrap.
P.(%): P-value for the effect or mediation proportion.
DVT, deep vein thrombosis; CI, confidence interval.

**Figure S3.** This image illustrates the two-step Mendelian Randomization mediation analysis exploring how COPS2 influences deep vein thrombosis (DVT) through immune-cell pathways. A. The diagram shows that COPS2 has a direct effect on effector-memory CD8⁺ T cells ($\beta_1$ = 0.25), while the mediator contributes to DVT risk ($\beta_2$ = 0.30), producing an indirect mediation effect of 0.075. B. The table further demonstrates that effector-memory CD8⁺ T cells significantly mediate the relationship between COPS2 and DVT, with a significant indirect effect and mediation proportion (P = 0.012). CD4⁺ T cells also show a smaller but statistically significant mediation effect (P = 0.045). Overall, the figure suggests that immune-cell regulation, particularly through effector-memory CD8⁺ T cells, plays an important role in the pathogenic effect of COPS2 on DVT development.

**Table S3.** Two-step mendelian randomization mediation analysis results

| Exposure | Mediator | Outcome | $\beta_1$ (Protein → Immune Trait) | $\beta_2$ (Immune Trait → DVT) | Indirect Effect ($\beta_1 \times \beta_2$) | 95% CI | *P*-value | Proportion Mediated (%) |
|---|---|---|---|---|---|---|---|---|
| COPS2 | Effector-memory CD8⁺ T cells | DVT | 0.25 | 0.30 | 0.075 | 0.02-0.13 | 0.012 | 41.7 |
| CDH3 | CD4⁺ T cells | DVT | 0.18 | 0.21 | 0.038 | 0.01-0.08 | 0.045 | 26.3 |

This table summarizes the two-step Mendelian randomization mediation analysis evaluating immune-cell pathways between plasma proteins and DVT. COPS2 shows a significant indirect effect mediated through effector-memory CD8⁺ T cells, accounting for 41.7% of its total effect on DVT risk (P = 0.012). CDH3 also demonstrates a significant but smaller mediated proportion (26.3%) via CD4⁺ T cells, indicating partial immune-mediated causal pathways.