

Multi-omics Identifies Hub Genes and Immune Networks in Stroke-CAD Comorbidity

Yingna Tian

XuChang Center Hospital

Qingbin Xu

Henan Provincial Health Commission Key Laboratory of Precision Medicine

Shuling Guo

Henan Provincial Health Commission Key Laboratory of Precision Medicine

Shuhao Mei

Henan Provincial Health Commission Key Laboratory of Precision Medicine

Huaqiang Wang

Henan Provincial Health Commission Key Laboratory of Precision Medicine

Weinan Zhao

Henan Provincial Health Commission Key Laboratory of Precision Medicine

Yan Zhao

XuChang Center Hospital

Hanjing Wei

XuChang Center Hospital

Wei Liu

XuChang Center Hospital

Lulu Li

XuChang Center Hospital

Tiantian Zhang

XuChang Center Hospital

Wei Li

XuChang Center Hospital

Yongsheng Liu

Chongqing Medical and Pharmaceutical College

Li Zhang

Chongqing Medical and Pharmaceutical College

Bingxin Xu

xuxinlzu@126.com

XuChang Center Hospital

Article

Keywords: Cerebral ischemic stroke, Coronary artery disease, Comorbidity, Biomarkers, Immune infiltration, VEGF signaling pathway

Posted Date: August 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9361787/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Cerebral ischemic stroke (CIS) and coronary artery disease (CAD) frequently co-occur, yet their shared molecular basis remains unclear. Using multi-omics analysis of transcriptomic data, we identified 96 common differentially expressed genes enriched in immune-metabolic pathways. PPI network analysis prioritized hub genes (ARG1, CLEC7A, PECAM1, TRIB1), which showed consistent dysregulation in independent cohorts and clinical samples, strong diagnostic accuracy (AUC > 0.7), and correlation with altered immune infiltration (M0 macrophages/neutrophils). A regulatory network involving TFs (CREB1, FOS) and miRNAs (hsa-miR-340-3p, hsa-miR-186-5p) was constructed. Our study reveals immunometabolic crosstalk in CIS-CAD comorbidity, providing novel biomarkers and therapeutic targets for vascular comorbidity management.

Introduction

Cerebral ischemic stroke (CIS) and coronary artery disease (CAD) collectively represent leading global causes of mortality, accounting for over 15 million deaths annually[1]. Epidemiological studies reveal a profound clinical synergy between these conditions: 20–40% of CIS patients exhibit concurrent CAD, while 10–15% of CAD cases develop stroke within 5 years[2–4]. This interconnection is further underscored by studies on transient ischemic attack (TIA), which shares etiological mechanisms with CIS. For instance, the Oxfordshire Community Stroke Project (OCSP) reported that TIA patients face a 27.8% 10-year risk of developing CAD[5], and individuals with TIA have a five-fold increased likelihood of coronary plaque presence[6]. Multivariate Mendelian randomization analyses confirm that the genetic predisposition to TIA is associated with a higher CAD risk, independent of conventional risk factors[7]. In the short term, CAD is a leading cause of 30-day readmission after acute stroke or TIA (prevalence 17.8%)[8], while long-term studies show that 32% of TIA patients develop CAD within 5 years[9]. The REACH registry further highlights an elevated rate of recurrent cardiovascular events in patients with a history of TIA or stroke (HR 1.52, 95% CI 1.40–1.65)[10]. These findings collectively emphasize that cerebral ischemic events (both TIA and CIS) are potent markers of widespread vascular vulnerability and significantly elevate the risk of CAD. Despite this substantial comorbidity burden, therapeutic strategies remain siloed due to poorly defined shared pathomechanisms at the molecular level. Current guidelines lack comorbidity-specific interventions, highlighting an urgent need to dissect the convergent molecular drivers underlying this clinical synergy.

Extensive transcriptomic studies have delineated distinct pathogenic pathways for CIS and CAD in isolation. In CIS, thrombotic inflammation driven by platelet-derived microvesicles and fibrinogen deposition is amplified by SH2B3-mediated JAK-STAT signaling, exacerbating neuroinflammation[11–12]. For CAD, lipid-driven atherosclerosis centers on PCSK9-dependent low-density lipoprotein(LDL) receptor degradation and foam cell formation[13], while hyperglycemia-induced H3K4me3/H3K27ac modifications in hematopoietic stem cells generate pro-inflammatory myeloid cells that promote plaque rupture [14]. Concurrently, immunometabolic dysregulation via the CD27/CD70-TRAF2 -SHP1-Lck axis impairs regulatory T cell(Treg) function, accelerating CAD progression [15].

Critically, three gaps impede progress in understanding comorbidity mechanisms. First, comorbidity-specific gene signatures are neglected. Existing gene expression profiles, such as GSE102558, predominantly focus on acute phase responses, overlooking chronic adaptations, such as ALDH2 dysregulation in lipid-stroke comorbidity[12]. Second, Multidimensional mechanisms remain unintegrated. No comprehensive framework that integrates differential gene expression—like the neuroprotective role of BTN3A2—with immune dysregulation, such as CD11c + B-cell expansion in CD27-deficient coronary artery disease, and transcriptional control mechanisms, such as FOS/MAX regulation of TRIB1 [15, 16]. Finally, Validated pan-disease biomarkers/therapeutics are absent. As current biomarkers like high-sensitivity C-Reactive protein (hsCRP) lack comorbidity specificity, while promising candidates such as ARG1/CLEC7A and shared therapeutic targets, like CD36 palmitoylation, await validation in comorbidity cohorts [17, 18].

To address the current gaps in understanding the molecular underpinnings of CIS-CAD comorbidity, we propose a novel hypothesis that this comorbid relationship originates from dysregulated immunometabolic hub genes governed by shared transcriptional networks. In pursuit of this hypothesis, we employ an integrated multi-omics approach to systematically interrogate the transcriptomic landscape associated with CIS-CAD comorbidity. Specifically, this study aims to: identify and map comorbidity-associated transcriptomic signatures and enriched biological pathways; delineate core regulatory hub genes through functional convergence analysis across multiple omics layers; characterize the dynamic alterations in the immune microenvironment and elucidate upstream transcriptional regulatory mechanisms; and assess the translational relevance of identified molecular features through biomarker discovery and validation. This integrative strategy is expected to yield novel insights into the shared pathophysiological mechanisms underlying CIS and CAD, potentially informing future targeted therapeutic and diagnostic developments.

Results

Analysis of Differentially Expressed Genes in Stroke and Coronary Artery Disease

Analysis of transcriptomic profiles identified distinct sets of differentially expressed genes (DEGs) in cerebral ischemic stroke (CIS) and coronary artery disease (CAD) patients compared to controls (Fig. 1). Applying consistent filtering criteria ($P < 0.05$ and $\log_2FC > 0.585$, corresponding to $FC > 1.5$), the CIS cohort (Fig. 1A) exhibited 1,280 significantly upregulated and 640 significantly downregulated genes. Similarly, the CAD cohort (Fig. 1B) showed 608 upregulated and 471 downregulated genes. Hierarchical clustering analysis further visualized these DEG expression patterns. Heatmaps contrasting control (Control) versus stroke (Stroke) groups (Fig. 1C) and control versus CAD groups (Fig. 1D) demonstrated clear separation of disease states, with genes (rows) color-coded from red (upregulation) to blue (downregulation) across samples (columns). These results reveal significant transcriptomic alterations underlying both CIS and CAD pathogenesis.

Gene Set Enrichment Analysis (GSEA) of Differential Gene Expression

Pathway enrichment analysis revealed distinct yet convergent biological mechanisms underlying CIS and CAD pathogenesis (Fig. 2). For stroke-associated DEGs (Fig. 2A), GSEA identified significant enrichment ($P < 0.05$) in five core pathways: lipid and atherosclerosis (hsa05417), Fluid Shear Stress and atherosclerosis (hsa05418), VEGF signaling pathway (hsa04370), Fc gamma R-mediated phagocytosis (hsa04666), and lysosome (hsa04141). The characteristic enrichment plots collectively implicated dysregulated lipid metabolism, platelet hyperactivity, angiogenesis modulation, and innate immune responses in stroke pathophysiology. Conversely in CAD (Fig. 2B), DEGs showed prominent enrichment in cholesterol metabolism (hsa04979), Type 2 Diabetes Mellitus (hsa04930), VEGF signaling pathway (hsa04370), non-alcoholic fatty liver disease (NAFLD; hsa04932), and IL-17 signaling pathway (hsa04657), highlighting impaired cholesterol homeostasis, ectopic lipid accumulation, vascular dysfunction, and IL-17-driven inflammation as central drivers of CAD progression. Critically, VEGF signaling (hsa04370) emerged as a shared node of dysregulation in both diseases ($P < 0.05$), suggesting aberrant angiogenesis regulation represents a common pathogenic mechanism in CIS-CAD comorbidity.

Intersection Analysis and Functional Enrichment of Comorbidity-Associated Genes

Venn diagram analysis (Fig. 3A) identified 96 co-expressed genes within the stroke (1,824 DEGs) and CAD (983 DEGs) transcriptomes (3.3% intersection), suggesting their potential synergistic role in comorbid pathogenesis. KEGG pathway analysis (Fig. 3B) demonstrated significant enrichment of these shared genes in several key processes, most notably immune-related pathways such as neutrophil extracellular trap (NET) formation and efferocytosis, as well as the VEGF signaling pathway ($P < 0.05$) [19–21]. Enrichment was also observed in other processes including transcriptional misregulation in cancer and adipocyte lipolysis regulation. Complementarily, functional enrichment analysis of biological processes (Fig. 3C) revealed their significant participation in negative regulation of interleukin-12 production, cellular response to lipopolysaccharide, and monocyte migration ($P < 0.05$), highlighting central roles in immune regulation, inflammatory response, and leukocyte trafficking. Quantitative bar plots visualizing enriched gene counts and significance levels for both KEGG pathways and GO biological processes systematically validated these findings, providing robust functional annotation for comorbidity-associated hub genes.

Identification and Validation of Comorbidity Hub Genes

Protein-protein interaction (PPI) network analysis (Fig. 4A) systematically identified ten hub genes central to CIS-CAD comorbidity. Cross-disease validation using independent external cohorts demonstrated consistent dysregulation of these genes: In datasets (GSE58294, GSE16561), five hub genes (ARG1, CAMP, CLEC7A, PECAM1, and TRIB1) exhibited significantly differential expression ($P <$

0.05) with concordant directional changes between CIS patients versus controls (Fig. 4B). Similarly, analysis of two CAD cohorts (GSE42148, GSE23561) confirmed robust dysregulation of these key candidates in CAD patients compared to controls (Fig. 4C). To further strengthen these findings, we performed experimental validation using qPCR on human blood samples. Consistent with the bioinformatic predictions, the expression levels of ARG1, CAMP, CLEC7A, PECAM1, and TRIB1 were significantly upregulated in CIS patients compared to healthy controls (Fig. 4D, $P < 0.05$). Similarly, in CAD patients, these five hub genes also showed marked overexpression relative to controls (Fig. 4E, $P < 0.05$). This multi-level validation, combining computational and experimental approaches, establishes these hub genes as core regulatory elements consistently perturbed in both disease states, positioning them as high-priority targets for elucidating shared molecular mechanisms and developing comorbidity-specific therapeutic strategies.

Diagnostic Potential of Hub Genes Across CIS and CAD

The dual diagnostic utility of prioritized hub genes (ARG1, CAMP, CLEC7A, PECAM1, TRIB1) was systematically evaluated through receiver operating characteristic (ROC) curve analysis in independent validation cohorts (Fig. 5). For stroke (Fig. 5A-5B), analysis across two datasets (GSE58294, GSE16561) revealed that ARG1, CLEC7A, PECAM1, and TRIB1 consistently achieved area under the curve (AUC) values > 0.7 , demonstrating robust discriminative capacity for distinguishing stroke patients from controls. Similarly in CAD (Fig. 5C-5D), evaluation using two independent cohorts (GSE42148, GSE23561) confirmed clinically significant diagnostic performance ($AUC > 0.7$) for the same four genes. Notably, a logistic regression model combining these four genes into a multi-gene panel demonstrated enhanced diagnostic accuracy across both stroke and CAD cohorts, consistently outperforming the individual markers in their respective validation sets. The fifth candidate, CAMP, which did not meet the consistency threshold ($AUC < 0.7$ in one dataset), is presented in Supplementary Fig. 1. This multi-cohort validation establishes ARG1, CLEC7A, PECAM1, and TRIB1 as pan-comorbidity biomarkers with clinically actionable diagnostic potential for both CIS and CAD, supporting their dual utility as diagnostic indicators and therapeutic targets.

Comprehensive Profiling of the Comorbid Immune Microenvironment

The immune landscape and core gene-immunocyte relationships in CIS and CAD comorbidity were systematically delineated (Fig. 6). Compositional analysis of immune infiltration (Fig. 6A, stacked bar plots) revealed disease-specific alterations. CIS samples exhibited distinct shifts in macrophage subpopulations versus controls, while CAD showed unique distribution patterns across immune subsets, indicating pathological remodeling of the immune microenvironment. Quantitative assessment (Fig. 6B) confirmed significant infiltration differences ($P < 0.05$; $P < 0.01$) for key subsets—including elevated M0 macrophages and neutrophils in CIS, and altered gamma delta T cells and macrophage subsets in CAD. Correlation heatmaps (Fig. 6C, Supplementary Tables) further uncovered complex immunocyte

interaction networks, with significant positive/negative associations ($P < 0.05$) between T cell subsets and macrophages reflecting disease-state-specific crosstalk. Critically, hub gene-immunocyte correlation analysis (Fig. 6D) demonstrated: ARG1 expression positively correlated with M0 macrophages and neutrophil infiltration ($P < 0.05$); CLEC7A, PECAM1, and TRIB1 showed significant associations with specific macrophage and T cell subsets ($P < 0.05$), mechanistically linking these genes to immunocyte recruitment and polarization. Collectively, this multi-dimensional profiling establishes aberrant immune dynamics as a hallmark of CIS-CAD comorbidity and identifies actionable targets for immune-focused therapeutic strategies.

Transcriptomic Regulatory Architecture of Comorbidity Hub Genes

The integrated transcriptional regulatory network orchestrating CIS-CAD comorbidity pathogenesis, constructed through systematic integration of hub genes, transcription factors (TFs), and microRNAs (miRNAs), is delineated (Fig. 7). Within this interactome, core comorbidity-associated genes (ARG1, TRIB1, CLEC7A, PECAM1) are coordinately regulated by key TFs including NFIC, CREB1, and FOS, while simultaneously being targeted by disease-relevant miRNAs such as hsa-miR-340 and hsa-miR-186. The edge mappings explicitly capture: TF-mediated gene regulation (e.g., ARG1 control by NFIC and CREB1), and miRNA-directed post-transcriptional silencing (e.g., ARG1 targeting by hsa-miR-340/hsa-miR-186; TRIB1 repression by hsa-miR-7/hsa-miR-23a alongside FOS/MAX regulation). This comprehensive regulatory blueprint resolves the upstream molecular circuitry governing comorbidity pathogenesis, identifying master regulators (CREB1, FOS) and non-coding RNA mediators (hsa-miR-340, hsa-miR-7) as priority targets for therapeutic intervention in CIS-CAD comorbidity.

Discussion

The present study provides a comprehensive multi-omics characterization of the shared molecular architecture underlying cerebral ischemic stroke (CIS) and coronary artery disease (CAD) comorbidity. We identified 96 common differentially expressed genes (DEGs) between CIS and CAD, enriched in immune-metabolic pathways such as neutrophil extracellular trap formation, VEGF signaling, and adipocyte lipolysis regulation. Through integrative bioinformatics and experimental validation, we further pinpointed five hub genes—ARG1, CAMP, CLEC7A, PECAM1, and TRIB1—that exhibited consistent dysregulation across independent CIS and CAD cohorts. Notably, the differential expression of these hub genes was further confirmed through qPCR analysis in clinical blood samples from both CIS and CAD patients, strengthening the reliability of our findings. These genes (ARG1, CLEC7A, PECAM1, and TRIB1) also demonstrated strong diagnostic potential ($AUC > 0.7$) and are embedded within a broader regulatory network involving key transcription factors (e.g., CREB1, FOS) and miRNAs (e.g., hsa-miR-340-3p, hsa-miR-186-5p). Furthermore, they correlated significantly associations with altered immune cell infiltration profiles, notably enriched M0 macrophages and neutrophils.

Our findings reinforce and extend current understanding of the immunometabolic axis in cardiovascular-cerebrovascular disease crosstalk. The sustained activation of VEGF signaling in both pathological contexts aligns with its well-established roles in vascular remodeling and inflammation. This pathway potentially modulates post-ischemic tissue repair and plaque stability, further validating its therapeutic potential for patients with comorbid cardiovascular conditions[22, 23]. Similarly, the preponderance of innate immune pathways—particularly neutrophil extracellular trap formation (NETosis) and FcγR-mediated phagocytosis[24, 25]—highlights the underappreciated role of neutrophil- and macrophage-driven inflammation in mediating end-organ damage in both vascular beds. The significant enrichment in neutrophil extracellular trap (NET) formation underscores a key role for dysregulated neutrophil activity, which can drive thromboinflammation—a process central to both ischemic stroke and coronary events. Concurrent enrichment in efferocytosis (the clearance of apoptotic cells) points to a potential impairment in the resolution of inflammation and tissue repair; defective efferocytosis is a known contributor to atherosclerotic plaque progression[26]. Notably, the hub gene ARG1, a marker of alternative macrophage activation[27], was consistently upregulated and showed a significant association with pro-inflammatory immune phenotypes. This correlation is consistent with the hypothesis that maladaptive immunometabolic responses could contribute to a shared mechanism, a possibility that requires functional validation.

The identified hub genes associated with CIS-CAD comorbidity—including ARG1, CLEC7A, TRIB1, CAMP, and PECAM1—have each been previously implicated in distinct biological processes relevant to cardiovascular and cerebrovascular diseases, such as lipid metabolism, immune regulation, and endothelial function. However, their collective role in the comorbidity of CIS and CAD has not been established. For example, TRIB1 is known to regulate lipoprotein metabolism and has been genetically associated with CAD through genome-wide association studies[28, 29], while CLEC7A acts as a pattern recognition receptor that modulates inflammatory responses[30, 31]. In addition, CAMP, which encodes the antimicrobial peptide LL-37, may contribute to plaque destabilization by regulating neutrophil activity and apoptosis[32, 33]. Meanwhile, PECAM1 (CD31), a key component of endothelial junctions, is critical for maintaining vascular integrity and facilitating leukocyte transendothelial migration[34]. Its dysregulation suggests widespread endothelial dysfunction in comorbid patients. The concurrent dysregulation of these genes, now validated at both transcriptional and clinical levels, suggests that they may operate within shared pathways—potentially related to innate immunity, leukocyte recruitment, and endothelial damage—that may collectively promote inflammatory and thrombotic processes in both cerebral and coronary vasculature. Thus, these hub genes represent promising targets for dual-purpose therapeutic strategies aimed at mitigating the shared mechanisms of stroke-CAD comorbidity.

Building on their observed co-dysregulation, we propose a synergistic immunometabolic model as a framework to hypothesize their potential joint contribution to CIS-CAD pathophysiology. These hub genes are predicted to function as interconnected nodes within a dysregulated network. In this model, PECAM1 dysregulation may be associated with altered endothelial integrity and leukocyte trafficking. Within the resultant inflammatory milieu, infiltrating myeloid cells could be influenced by pattern recognition via CLEC7A, with ARG1 potentially modulating local immune responses. Concurrently, TRIB1-

associated lipid metabolic dysregulation might provide a permissive metabolic environment. This potential reinforcing circuit—linking metabolic disturbance (TRIB1), endothelial alterations (PECAM1), and innate immune signaling (CLEC7A/ARG1)—presents a testable hypothesis for how these associations may collectively contribute to vascular inflammation and dysfunction across both vascular beds.

From a translational perspective, the diagnostic performance of the identified biomarkers (AUC > 0.7 in independent datasets) supports their potential utility in risk stratification and early detection of comorbidity. Importantly, a combined diagnostic model integrating these four hub genes showed superior accuracy compared to any single gene, underscoring the additive value of a multi-marker approach for comorbidity risk assessment. However, their clinical significance must be considered in the context of established biomarkers. Current clinical practice for cardiovascular risk stratification often relies on measures of systemic inflammation, such as high-sensitivity C-reactive protein (hs-CRP) or interleukin-6 (IL-6). The present discovery-phase study was not designed as a direct, head-to-head comparative diagnostic trial against these canonical markers. Instead, we propose ARG1, CLEC7A, PECAM1, and TRIB1 as novel, comorbidity-derived candidates that may capture distinct aspects of the shared immunometabolic dysregulation in CIS-CAD, potentially offering mechanistic specificity beyond general inflammation. Future studies in larger, prospective cohorts that concurrently measure these hub genes alongside established biomarkers are warranted to determine their independent, additive, or superior diagnostic and prognostic value for managing patients with vascular comorbidity. Moreover, the regulatory network constructed here—incorporating TFs, miRNAs, and hub genes—generates a testable hypothesis framework for future mechanistic validation and could inform therapeutic intervention strategies. For example, miRNAs such as hsa-miR-340-3p, which targets ARG1, could be explored for their modulating effects in experimental models of combined neurovascular and cardiovascular injury.

Limitations of the study

Several limitations should be acknowledged. Although we integrated multiple public datasets and performed external validation (including qPCR in clinical samples), the sample sizes of some cohorts were modest, and prospective clinical validation is required to establish generalizability. The integration of data from independent cohorts, while powerful for discovery, introduces inherent challenges related to dataset heterogeneity and potential batch effects, which may influence cross-study comparability despite our use of standardized analytical pipelines. The functional roles of the proposed hub genes and regulatory elements remain largely inferred from bioinformatics; future studies using animal models or cellular assays are necessary to establish causal relationships. Furthermore, while our integrative analysis identifies strong associations among hub genes, immune infiltration, and regulatory elements, the causal relationships implied within these networks remain to be established through functional gain- or loss-of-function experiments. Additionally, the interplay between specific immune subtypes and gene expression deserves deeper single-cell resolution.

Conclusion

In conclusion, our study elucidates a core set of hub genes and regulatory networks that underlie the comorbidity between cerebral ischemic stroke and coronary artery disease, highlighting convergent immunometabolic pathways as key mechanistic drivers. The robust validation of key hub genes through qPCR in human clinical samples enhances the translational relevance of our findings. These results provide a foundational resource for future research aimed at co-diagnosis and co-therapy in patients with overlapping cerebrovascular and cardiovascular diseases. Further functional studies and clinical validation efforts will be essential to translate these insights into precise diagnostic tools and targeted therapies for high-comorbidity populations.

STAR Methods

Experimental model and study participant details

Study Population and Clinical Sample Collection

For the subsequent qPCR validation, an independent set of clinical blood samples were collected. Peripheral blood samples were obtained from three distinct groups: patients with cerebral ischemic stroke (CIS), patients with coronary artery disease (CAD), and age- and sex-matched healthy controls. All participants were recruited at Xuchang Central Hospital between January 2023 and August 2023. The study protocol was approved by the Institutional Review Board of Xuchang Central Hospital (Approval No. 2022-11-001), and written informed consent was obtained from all participants. For the CIS group (n = 5): Diagnosis was confirmed based on the World Health Organization (WHO) criteria and neuroimaging findings (computed tomography or magnetic resonance imaging). Inclusion criteria comprised: (1) first-ever acute ischemic stroke within 24 hours of symptom onset; (2) age between 18 and 80 years. Exclusion criteria included: (1) history of other neurological diseases; (2) presence of malignant tumors; (3) severe hepatic or renal dysfunction. For the CAD group (n = 9): Diagnosis was confirmed by coronary angiography showing >50% stenosis in at least one major coronary artery. Inclusion criteria were: (1) stable CAD or acute coronary syndrome; (2) age between 18 and 80 years. Patients were excluded if they had: (1) a prior history of stroke; (2) concomitant severe inflammatory or autoimmune diseases. For the healthy control group (n = 5): Participants were recruited from individuals undergoing routine health check-ups. They had no history of cardiovascular disease, stroke, or other serious chronic illnesses, and were matched for age and sex with the patient groups.

This validation cohort consisted of: 5 patients with CIS (4 males, 1 female; mean age 60.6 years, range 30-84), 9 patients with CAD (7 males, 2 females; mean age 63.6 years, range 36-80), and 5 healthy controls (3 males, 2 females; mean age 63.2 years, range 52-79). These samples were independent of the public datasets used for bioinformatic discovery and analysis.

Ethics approval and consent to participate

All procedures involving human participants in this study were performed in compliance with the ethical standards of the Institutional Review Board of Xuchang Central Hospital (Approval No. 2022-11-001) and

adhered to the principles outlined in the Declaration of Helsinki. Written informed consent was obtained from all participants prior to their enrollment in the study. The consent process included a detailed explanation of the research objectives, procedures, potential risks and benefits, and the right to withdraw at any time without penalty.

Blood samples collected from patients with cerebral ischemic stroke (CIS) and coronary artery disease (CAD) were procured in strict accordance with the above-mentioned ethical guidelines. All samples and associated data were anonymized and handled with strict confidentiality to ensure the protection of participants' privacy.

Method details

Data Collection and Processing

Raw gene expression datasets for CIS (GSE58294 and GSE16561) and CAD (GSE42148 and GSE23561) were downloaded from GEO. To ensure a valid comparison of systemic molecular profiles, we exclusively selected datasets derived from human peripheral blood samples. The datasets and their sample sizes were as follows: for CIS, the discovery set GSE58294 comprised 69 patients and 23 controls (total n=92), and the validation set GSE16561 comprised 39 patients and 24 controls (total n=63); for CAD, the discovery set GSE42148 comprised 13 patients and 11 controls (total n=24), and the validation set GSE23561 comprised 6 patients and 9 controls (total n=15).

All datasets were independently subjected to a standardized preprocessing pipeline using R. First, quality control (QC) was performed to filter out low-quality samples and probes. This included assessing sample clustering, inter-array correlation, and the presence of outlier samples using principal component analysis (PCA). Probe-level data were then normalized using the robust multi-array average (RMA) method (for microarray platforms) / the regularized-log transformation (for RNA-seq data) to ensure comparability across samples. The resulting harmonized and normalized expression matrices were used for all downstream analyses.

GSE58294 and GSE42148 were used as discovery sets to identify differentially expressed genes (DEGs) common to CIS and CAD. GSE16561 and GSE23561 served as external validation sets to confirm the expression patterns and diagnostic utility of overlapping genes. Protein–protein interaction (PPI) networks were employed to screen candidate biomarkers. Additionally, blood samples from patients with CIS and CAD were collected, and the expression levels of key hub genes were validated using quantitative polymerase chain reaction (qPCR).

Identification and Functional Enrichment Analysis of Differentially Expressed Genes (DEGs)

Visualization of DEGs was performed using volcano plots and heatmaps generated with the 'ggplot2' and 'pheatmap' packages. Gene Set Enrichment Analysis (GSEA) was conducted to explore enriched biological functions and signaling pathways in CIS and CAD. The analysis was performed using the 'clusterProfiler' package with the Kyoto Encyclopedia of Genes and Genomes (KEGG) database as the

reference gene set. A pre-ranked analysis strategy was employed, using the log2 fold change values of all genes for ranking. Gene sets with a size smaller than 10 or larger than 500 genes were excluded from the analysis. Shared DEGs between the two diseases were illustrated using the 'ggvenn' package. Functional annotation of DEGs was further performed through Gene Ontology (GO) and KEGG analyses using 'clusterProfiler', with results visualized via 'ggplot2'.

Screening of Hub Genes Using PPI Interaction Networks

Protein–protein interaction (PPI) networks for shared DEGs were constructed using the STRING database and visualized in Cytoscape. The top 10 hub genes were identified using the CytoHubba plugin based on degree of connectivity, and these were selected as candidate diagnostic biomarkers.

Expression Analysis and Diagnostic Evaluation of Candidate Biomarkers

The expression levels of core genes were compared between disease and control groups in the validation datasets (GSE16561 and GSE23561) using the 'ggplot2' package. Group differences were assessed with the Wilcoxon test ($p < 0.05$). The diagnostic performance of candidate biomarkers was evaluated by constructing receiver operating characteristic (ROC) curves and calculating the area under the curve (AUC) along with its 95% confidence interval (CI) using the 'pROC' package. The 95% CIs were computed using the DeLong test. An AUC value > 0.7 was considered indicative of good diagnostic accuracy. Additionally, logistic regression analysis was used to develop a multi-gene diagnostic model, and ROC curves were plotted to compare the performance of single-gene and multi-gene diagnostic models.

Quantitative Real-Time PCR (qPCR) Validation

Blood samples were collected from patients with CIS, CAD, and healthy controls. For each subject, approximately 2 ml of peripheral whole blood was collected. Total RNA was extracted from whole-blood-derived leukocytes. Briefly, red blood cells were lysed with ammonium chloride (NH_4Cl) solution for 3 minutes, followed by neutralization with phosphate-buffered saline (PBS) and centrifugation at 1,500 rpm for 5 minutes. The resulting leukocyte pellet was washed once with PBS. Total RNA was then extracted using TRIzol reagent (Invitrogen, USA) and reverse transcribed into cDNA using HiScript III RT SuperMix (Vazyme, China). qPCR was performed on a qTOWER3/G Real-Time PCR System (Jena Analytics, Germany) with Taq Pro Universal SYBR qPCR Master Mix (Vazyme, China). The $2^{-\Delta\Delta\text{Ct}}$ method was applied to quantify relative gene expression, with β -actin serving as the internal control. The primer sequences used were as follows: ARG1 (forward: CCCTTTGCTGACATCC; reverse: TTCCGTTCTTCTTGACTT); CAMP (forward: CTCGGATGCTAACCTCT; reverse: CATAACCGCTTCACC); CLEC7A (forward: TCTTTCCAGCCCTTGT; reverse: TGCCAGCATTGTCTTT); PECAM1 (forward: GATAGCCTCAAAGTCG; reverse: TGGGCATCATAAGAAA); TRIB1 (forward: CATTACTGGCATTGTGGAA; reverse: ACGATGGCTGACTGGTG); β -actin (forward: ACACTGTGCCCATCTACG; reverse: TGTCACGCACGATTTC).

Immune Cell Infiltration Analysis

The CIBERSORT algorithm with the LM22 signature matrix was used to estimate immune cell infiltration levels in CIS and CAD samples. Differences in immune cell abundances between disease and control groups were assessed using the Wilcoxon test. Spearman correlation analysis was applied to examine relationships between immune cell subtypes and core immune-related DEGs.

Construction of miRNA–mRNA–Transcription Factor (TF) Regulatory Networks

Experimentally supported miRNA–mRNA interactions were retrieved from the starBase database, while TF–mRNA interactions were obtained from JASPAR using the NetworkAnalyst 3.0 platform. Integrated regulatory networks comprising miRNAs, TFs, and hub genes were constructed and visualized using Cytoscape.

Quantification and statistical analysis

Statistical Analysis

All statistical analyses were performed using R software (version 4.4.1). Differential gene expression between disease and control groups was identified using the limma R package, with a significance threshold set at adjusted P-value (False Discovery Rate, FDR) < 0.05 and an absolute $\log_2$ fold change $|\log_2\text{FC}| > 0.585$. Comparisons of hub gene expression and immune cell infiltration levels between two groups were conducted using the non-parametric Wilcoxon rank-sum test. Spearman's correlation analysis was applied to assess the relationships between hub gene expression and immune cell abundances. The diagnostic performance of candidate biomarkers was evaluated by generating receiver operating characteristic (ROC) curves and calculating the area under the curve (AUC) using the pROC package. A two-tailed p-value of less than 0.05 was considered statistically significant for all analyses, unless otherwise specified.

Declarations

Resource availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Acknowledgments

The authors would like to thank all the participants and clinical staff who contributed to this study. We also extend our gratitude to the research teams and institutions that provided data and technical support. Special thanks are due to our colleagues for their insightful discussions and constructive suggestions throughout this project.

Author's contributions

Yingna Tian: Conceptualization, Methodology, Investigation, Formal Analysis, Writing-Original Draft. Qingbin Xu: Supervision, Methodology, Writing-Review & Editing, Funding Acquisition. Shuling Guo: Investigation, Resources. Shuhao Mei: Validation, Visualization. Huaqiang Wang: Resources, Project Administration. Weinan Zhao: Data Curation, Visualization. Yan Zhao: Formal Analysis, Visualization. Hanjing Wei: Investigation, Data Curation. Wei Liu: Methodology, Resources. Lulu Li: Resources, Project Administration. Tiantian Zhang: Investigation. Wei Li: Validation, Formal Analysis. Yongsheng Liu: Supervision, Writing-Review & Editing. Li Zhang: Supervision, Funding Acquisition. Bingxin Xu: Conceptualization, Supervision, Writing-Review & Editing, Funding Acquisition.

Consent to Participate declaration

Not applicable.

Declaration of interests

No potential conflict of interest was reported by the authors.

Clinical trial number

Not applicable.

Consent to Participate declaration:

Not applicable.

Funding

This work was supported by the Key Research and Development Program of Henan Province (Grant No. 232102311112) and the Chongqing Medical Scientific Research Project (Joint Project of Chongqing Health Commission and Science and Technology Bureau, China) (Grant Nos. 2023GGXM006, 2024ZDXM026, and 2024MSXM115).

References

1. Hata, J. & Kiyohara, Y. Epidemiology of stroke and coronary artery disease in Asia. *Circ. J.* **77** (8), 1923–1932 (2013).
2. Lackland, D. T. et al. Inclusion of stroke in cardiovascular risk prediction instruments: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* **43** (7), 1998–2027 (2012).
3. Steg, P. G. et al. One-year cardiovascular event rates in outpatients with atherothrombosis. *Jama* **297** (11), 1197–1206 (2007).
4. Calvet, D. et al. Prevalence of asymptomatic coronary artery disease in ischemic stroke patients: the PRECORIS study. *Circulation* **121** (14), 1623–1629 (2010).

5. Clark, T. G., Murphy, M. F. G. & Rothwell, P. M. Long term risks of stroke, myocardial infarction, and vascular death in low risk patients with a non-recent transient ischaemic attack. *J. Neurol. Neurosurg. Psychiatry*. **74** (5), 577–580 (2003).
6. Jensen, J. K. et al. Association of ischemic stroke to coronary artery disease using computed tomography coronary angiography. *Int. J. Cardiol.* **160** (3), 171–174 (2012).
7. Qi, X. et al. Transient ischemic attack and coronary artery disease: a two-sample Mendelian randomization analysis. *Front. Cardiovasc. Med.* **10**, 1192664 (2023).
8. Zhong, W. et al. Prevalence, causes and risk factors of hospital readmissions after acute stroke and transient ischemic attack: a systematic review and meta-analysis. *Neurol. Sci.* **37** (8), 1195–1202 (2016).
9. Yousufuddin, M. et al. Five-year mortality after transient ischemic attack: focus on cardiometabolic comorbidity and hospital readmission. *Stroke* **49** (3), 730–733 (2018).
10. Ducrocq, G. et al. A history of stroke/transient ischemic attack indicates high risks of cardiovascular event and hemorrhagic stroke in patients with coronary artery disease. *Circulation* **127** (6), 730–738 (2013).
11. Bai, W. et al. Shared genetic architecture between stroke and blood lipids: a large-scale genome-wide cross-trait analysis. *Hum. Genomics*. **19** (1), 75 (2025).
12. Zheng, J. et al. Immunological characterization of stroke-heart syndrome and identification of inflammatory therapeutic targets. *Front. Immunol.* **14**, 1227104 (2023).
13. Choudhury, R. P. et al. Inflammation reprogramming and immunomodulation: Next-generation treatments for atherosclerosis. *Cell. Rep. Med.* (2025).
14. García-Calzón, S. et al. *Epigenetic biomarkers predict macrovascular events in individuals with type 2 diabetes* (Cell Reports Medicine, 2025).
15. Song, X. F. et al. The Relationship and Mechanism of CD27 with Coronary Atherosclerotic Heart Disease. *Cardiovasc. Drugs Ther.* Aug 8. (2025).
16. Li, C. et al. Proteome-Wide Mendelian Randomization Identifies Candidate Causal Proteins for Cardiovascular Diseases. *Adv. Genet. (Hoboken)*. **6** (2), 2500003 (2025).
17. Short, S. A. P. et al. Improving stroke risk prediction in atrial fibrillation with circulating biomarkers: the CHA2DS2-VASc–Biomarkers model. *J. Thromb. Haemost.*, (2025).
18. Zhang, Q. et al. Inhibiting CD36 palmitoylation improves cardiac function post-infarction by regulating lipid metabolic homeostasis and autophagy. *Nat. Commun.* **16** (1), 6602 (2025).
19. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res.* **53**, D672–D677 (2025).
20. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).
21. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

22. Hu, Y., Zheng, Y., Wang, T., Jiao, L. & Luo, Y. VEGF, a Key Factor for Blood Brain Barrier Injury After Cerebral Ischemic Stroke. *Aging Dis.* **13** (3), 647–654 (2022).
23. Lemström, K. et al. Role of angiogenic growth factors in transplant coronary artery disease. *Ann. Med.* **36** (3), 184–193 (2004).
24. Morris, A. B. et al. Signaling through the inhibitory Fc receptor FcγRIIB induces CD8 + T cell apoptosis to limit T cell immunity. *Immunity* **52** (1), 136–150 (2020). e6.
25. Alemán, O. R. & Rosales, C. Human neutrophil Fc gamma receptors: different buttons for different responses. *J. Leukoc. Biol.* **114** (6), 571–584 (2023).
26. Mehrotra, P. & Ravichandran, K. S. Drugging the efferocytosis process: concepts and opportunities. *Nat. Rev. Drug Discovery.* **21** (8), 601–620 (2022).
27. Martinenaite, E. et al. Arginase-1-specific T cells target and modulate tumor-associated macrophages. *J. Immunother. Cancer.* **13** (1), e009930 (2025).
28. Prete, V. et al. TRIB1: a multifaceted regulator of cardiometabolic health. *American Journal of Physiology-Cell Physiology*, 328(6): C1973-C1981. (2025).
29. Jiang, J. et al. Polymorphisms of TRIB1 genes for coronary artery disease and stroke risk: A systematic review and meta-analysis. *Gene* **880**, 147613 (2023).
30. Li, W. et al. CLEC7A Knockdown Alleviates Ischemic Stroke by Inhibiting Pyroptosis and Microglia Activation. *J. Integr. Neurosci.* **23** (12), 219 (2024).
31. Ye, X. C. et al. Dectin-1/Syk signaling triggers neuroinflammation after ischemic stroke in mice. *J. Neuroinflamm.* **17** (1), 17 (2020).
32. Wang, Z. et al. PGE2 inhibits neutrophil phagocytosis through the EP2R-cAMP-PTEN pathway. *Immun. Inflamm. Dis.* **10** (7), e662 (2022).
33. Fantidis, P. The role of intracellular 3'5'-cyclic adenosine monophosphate (cAMP) in atherosclerosis. *Curr. Vasc. Pharmacol.* **8** (4), 464–472 (2010).
34. Chuntharpursat-Bon, E. et al. PIEZO1 and PECAM1 interact at cell-cell junctions and partner in endothelial force sensing. *Commun. Biol.* **6** (1), 358 (2023).

Figures

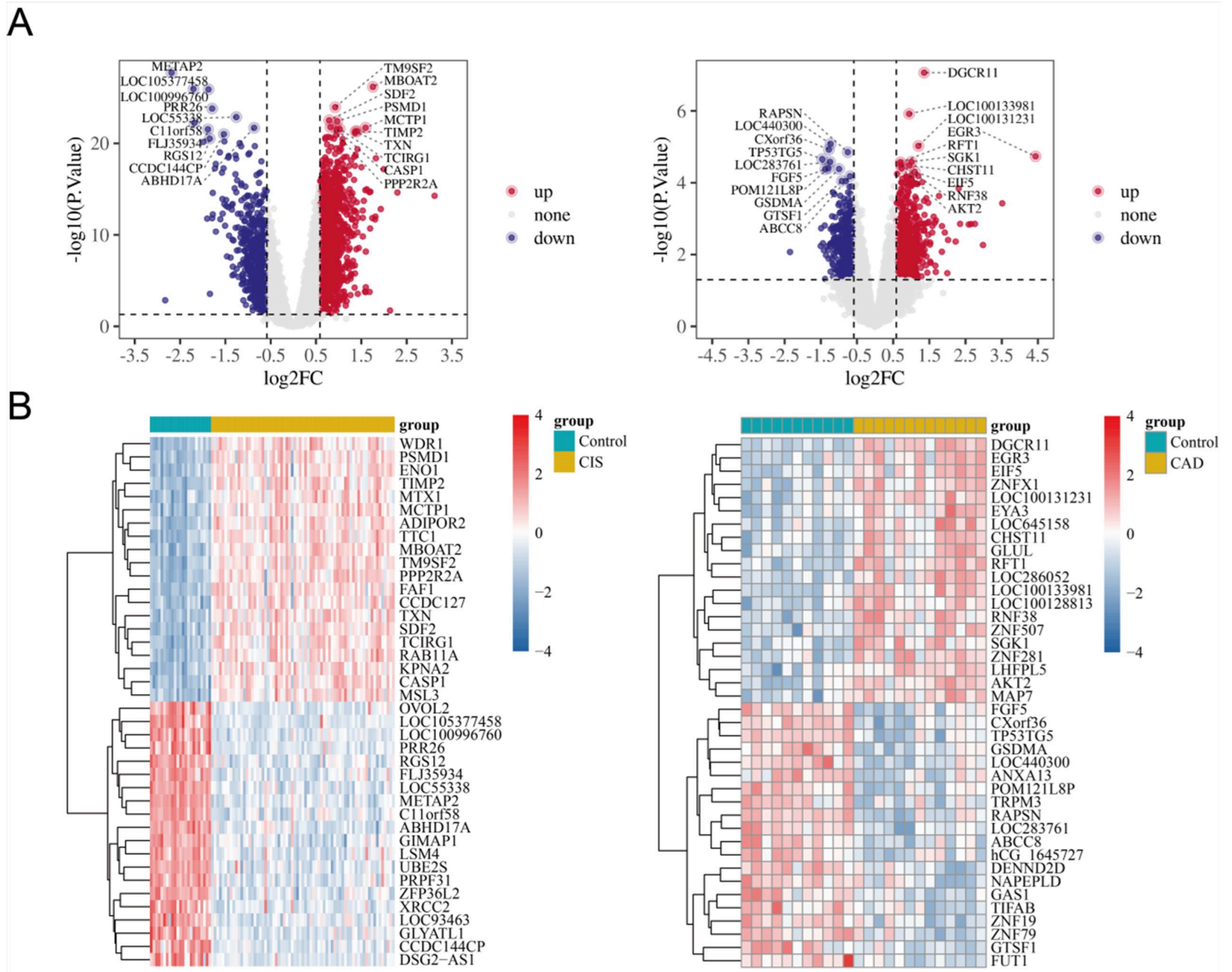


Figure 1

Transcriptomic profiling identifies differentially expressed genes in cerebral ischemic stroke and coronary artery disease.

(A) Volcano plot displaying differentially expressed genes (DEGs) in the cerebral ischemic stroke (CIS) cohort. Red points represent significantly upregulated genes (1,280), blue points indicate downregulated genes (640), and grey points denote non-significant genes(left). Volcano plot of DEGs in the coronary artery disease (CAD) cohort, with 608 upregulated (red) and 471 downregulated (blue) genes(right).(filter criteria: $P < 0.05$ and $|\log_2FC| > 0.585$).

(B) Heatmap of DEG expression in CIS compared to control samples. Rows represent genes, columns represent samples. Red indicates upregulation and blue indicates downregulation(left). Heatmap of DEG expression in CAD compared to control samples, showing clear separation between disease and control groups.

Hierarchical clustering was performed to visualize distinct transcriptomic profiles between disease states and controls(right).

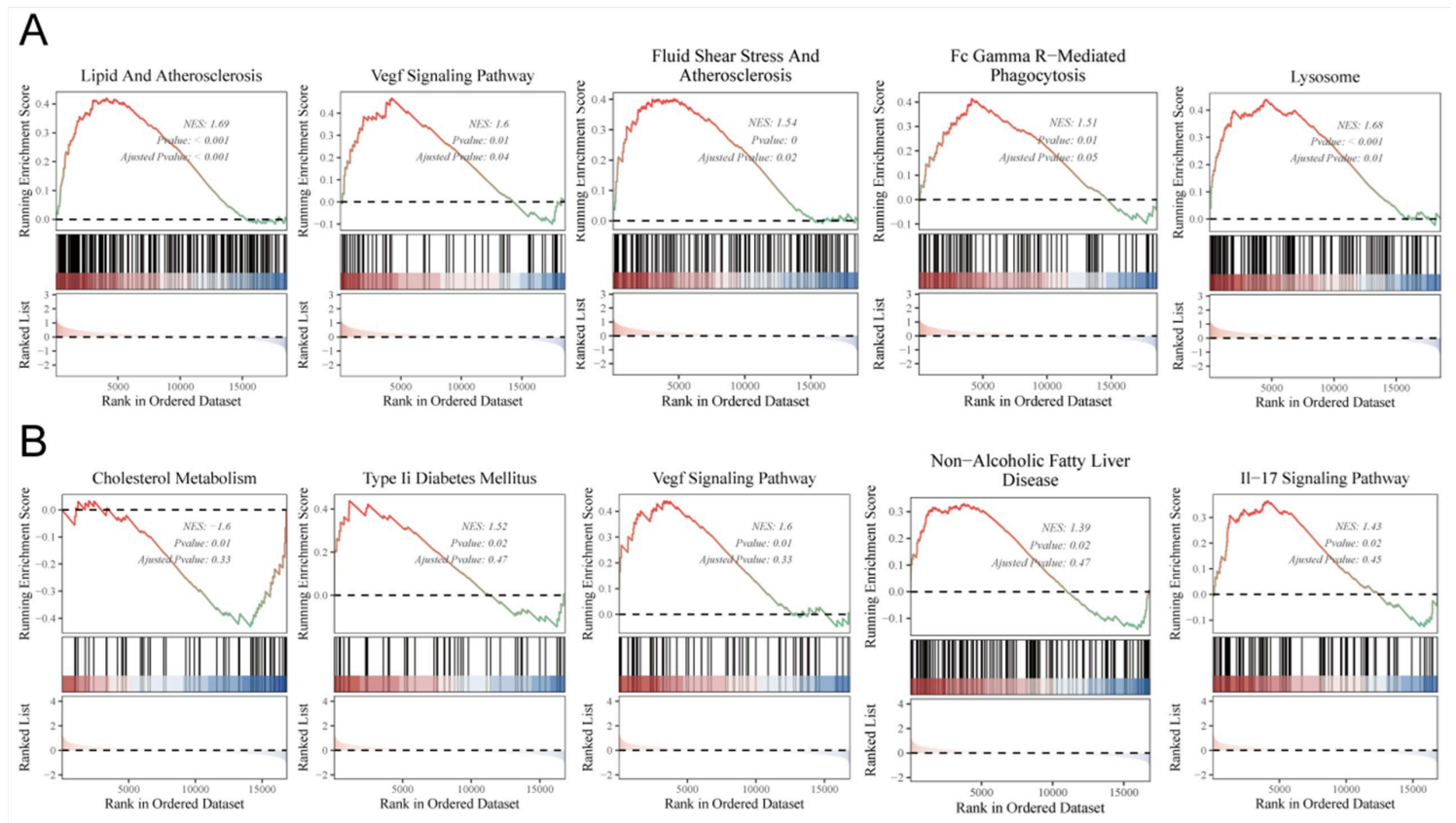


Figure 2

Gene set enrichment analysis reveals shared and distinct pathways in cerebral ischemic stroke and coronary artery disease.

(A) GSEA enrichment plots for cerebral ischemic stroke (CIS) showing significantly enriched pathways (FDR < 0.25): lipid and atherosclerosis (hsa05417), fluid shear stress and atherosclerosis (hsa05418), VEGF signaling pathway (hsa04370), Fc gamma R-mediated phagocytosis (hsa04666), and lysosome (hsa04142).

(B) GSEA enrichment plots for coronary artery disease (CAD) indicating significantly enriched pathways: cholesterol metabolism (hsa04979), type II diabetes mellitus (hsa04930), VEGF signaling pathway (hsa04370), non-alcoholic fatty liver disease (NAFLD, hsa04932), and IL-17 signaling pathway (hsa04657).

The VEGF signaling pathway (hsa04370) was significantly enriched in both diseases, suggesting a common angiogenic regulatory mechanism in CIS-CAD comorbidity.

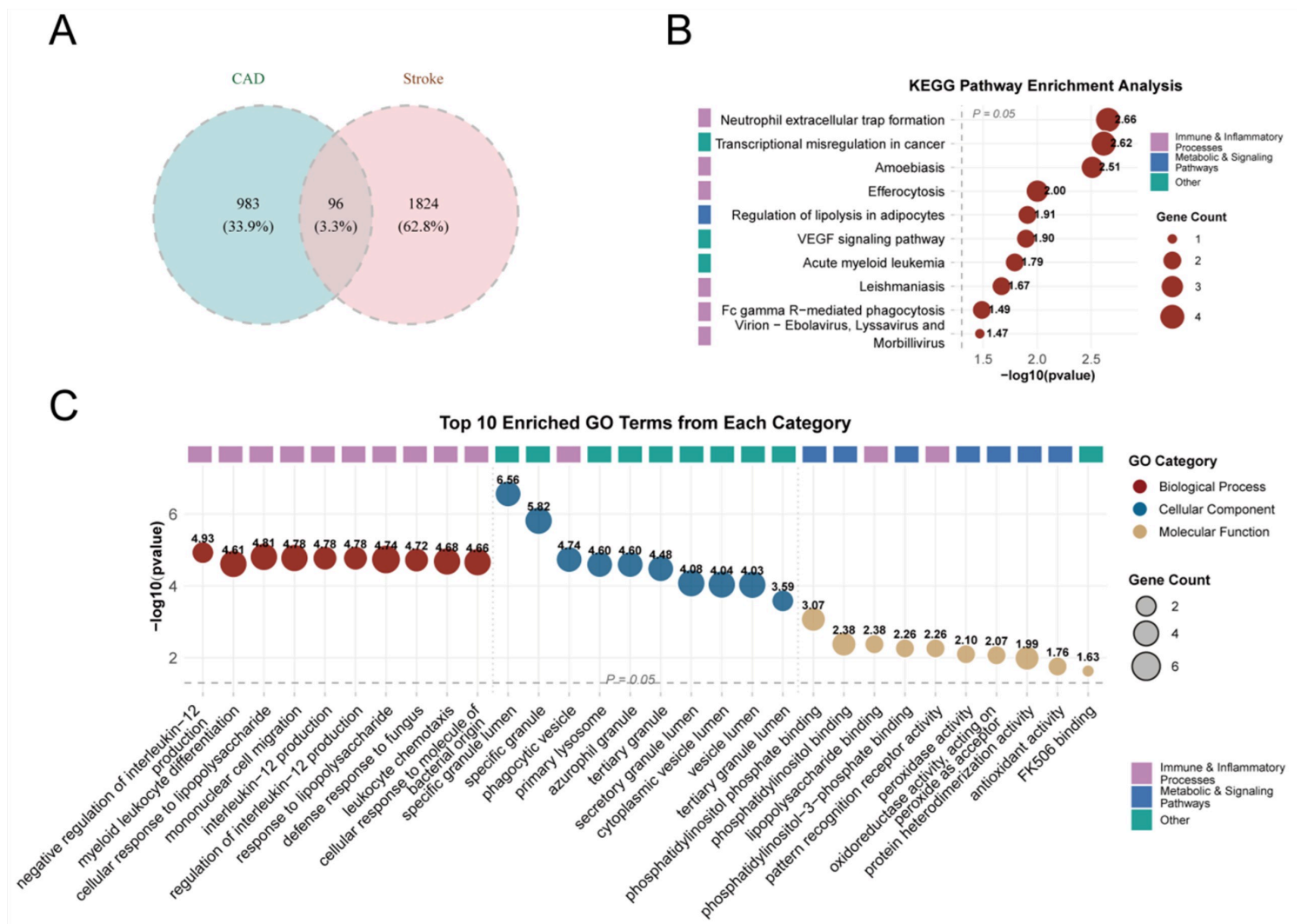


Figure 3

Intersection analysis and functional enrichment of comorbidity-associated genes in cerebral ischemic stroke and coronary artery disease.

(A) Venn diagram illustrating the overlap of differentially expressed genes between cerebral ischemic stroke (1,824 DEGs) and coronary artery disease (983 DEGs), identifying 96 common genes (3.3% overlap).

(B) KEGG pathway enrichment analysis of the shared genes. Significantly enriched pathways include neutrophil extracellular trap (NET) formation, transcriptional misregulation in cancer, efferocytosis, adipocyte lipolysis regulation, and VEGF signaling ($P < 0.05$).

(C) Gene Ontology (GO) biological process enrichment analysis of the common genes, showing significant terms including negative regulation of interleukin-12 production, cellular response to lipopolysaccharide, and monocyte migration ($P < 0.05$). Bar plots represent the number of enriched genes and the corresponding $-\log_{10}(P\text{-value})$ for each term.

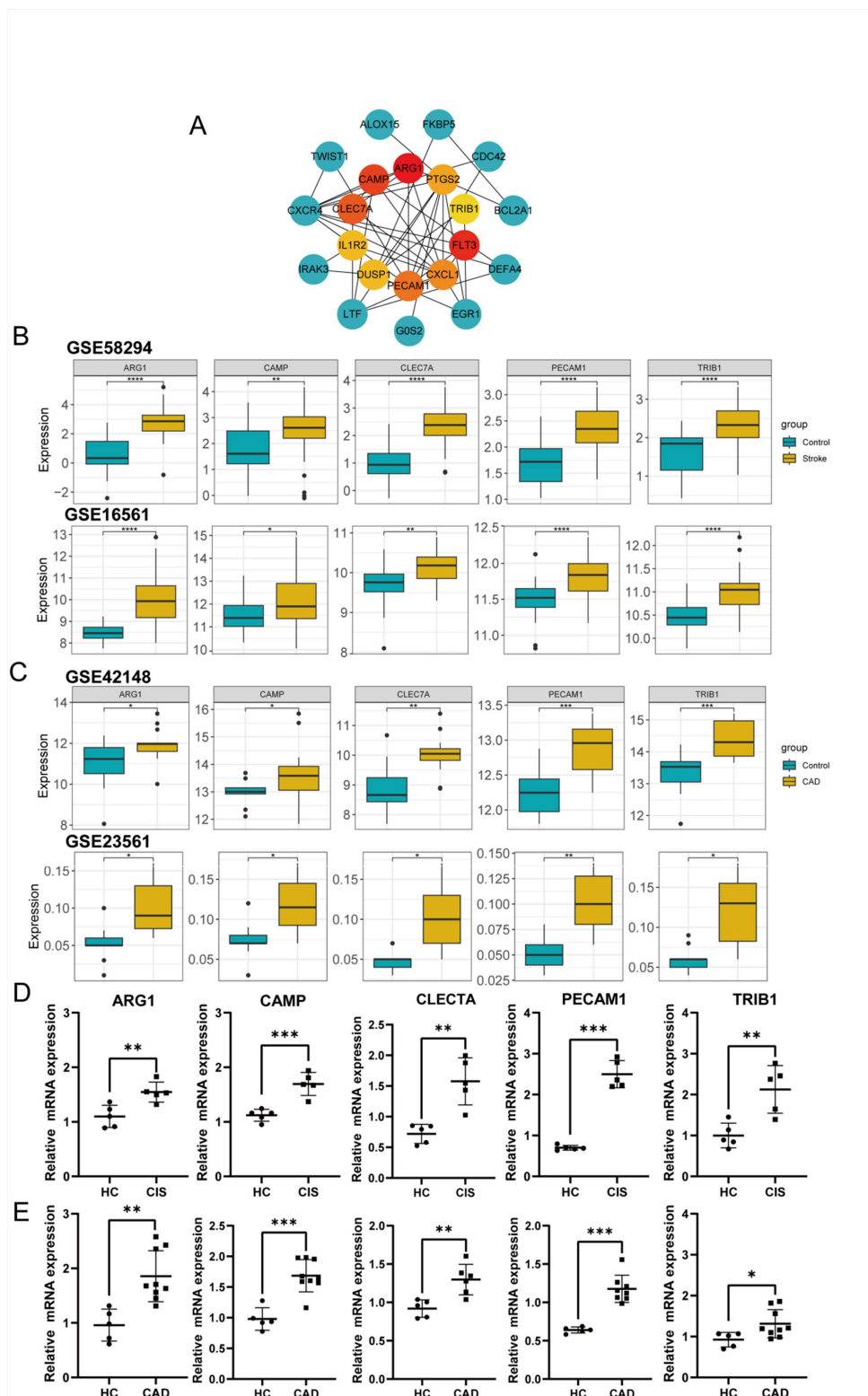


Figure 4

Identification and multi-level validation of hub genes in cerebral ischemic stroke and coronary artery disease comorbidity.

(A) Protein-protein interaction (PPI) network of shared differentially expressed genes. The top 10 hub genes were identified using the CytoHubba plugin in Cytoscape.

- (B) Expression of five key hub genes (ARG1, CAMP, CLEC7A, PECAM1, and TRIB1) in two independent cerebral ischemic stroke (CIS) datasets (GSE58294 and GSE16561). Box plots show significant differential expression in CIS patients compared to controls ($P < 0.05$).
- (C) Expression of the same five hub genes in two independent coronary artery disease (CAD) datasets (GSE42148 and GSE23561), confirming significant dysregulation in CAD patients ($P < 0.05$).
- (D) Experimental validation by qPCR showing upregulated expression of the five hub genes in blood samples from CIS patients compared to healthy controls ($P < 0.05$).
- (E) qPCR validation confirming significant overexpression of the five hub genes in blood samples from CAD patients compared to healthy controls ($P < 0.05$).

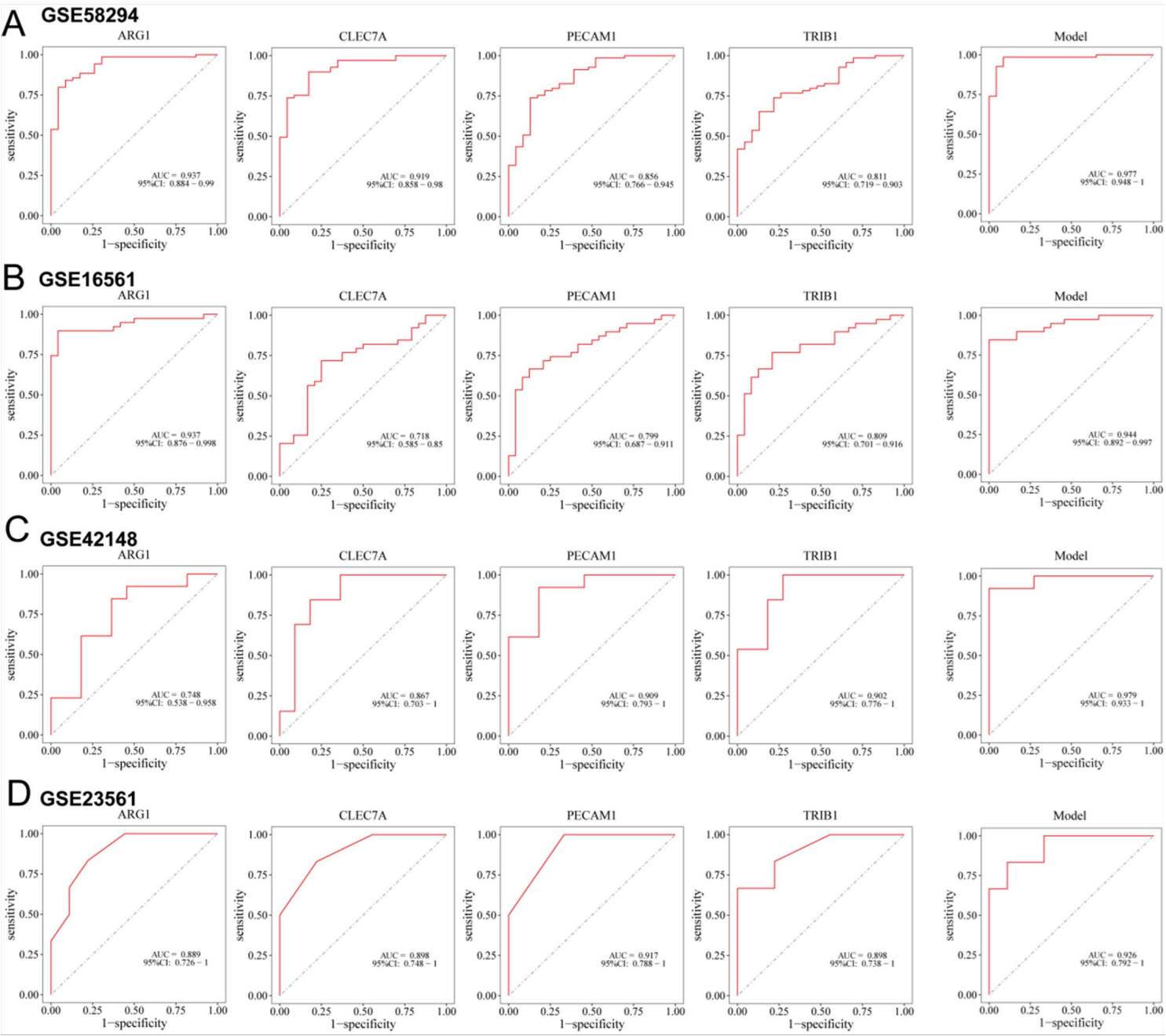


Figure 5

Diagnostic evaluation of individual hub genes and a combined multi-gene model for cerebral ischemic stroke and coronary artery disease.

(A) Receiver operating characteristic (ROC) curves demonstrating the diagnostic performance of hub genes (ARG1, CLEC7A, PECAM1, and TRIB1) and a combined multi-gene model in two independent cerebral ischemic stroke (CIS) datasets (GSE58294). The four hub genes (ARG1, CLEC7A, PECAM1, TRIB1) and the model exhibited AUC values > 0.7 .

(B) ROC curves demonstrating the diagnostic performance of hub genes and a combined multi-gene model in two independent cerebral ischemic stroke (CIS) datasets (GSE16561). The four hub genes (ARG1, CLEC7A, PECAM1, TRIB1) and the model exhibited AUC values > 0.7 .

(C) ROC curves validating the diagnostic utility of the same four hub genes and the combined model in two independent coronary artery disease (CAD) cohorts (GSE42148), with AUC values > 0.7 , confirming their discriminative capacity for CAD detection.

(D) ROC curves validating the diagnostic utility of the same four hub genes and the combined model in two independent coronary artery disease (CAD) cohorts (GSE23561), with AUC values > 0.7 , confirming their discriminative capacity for CAD detection.

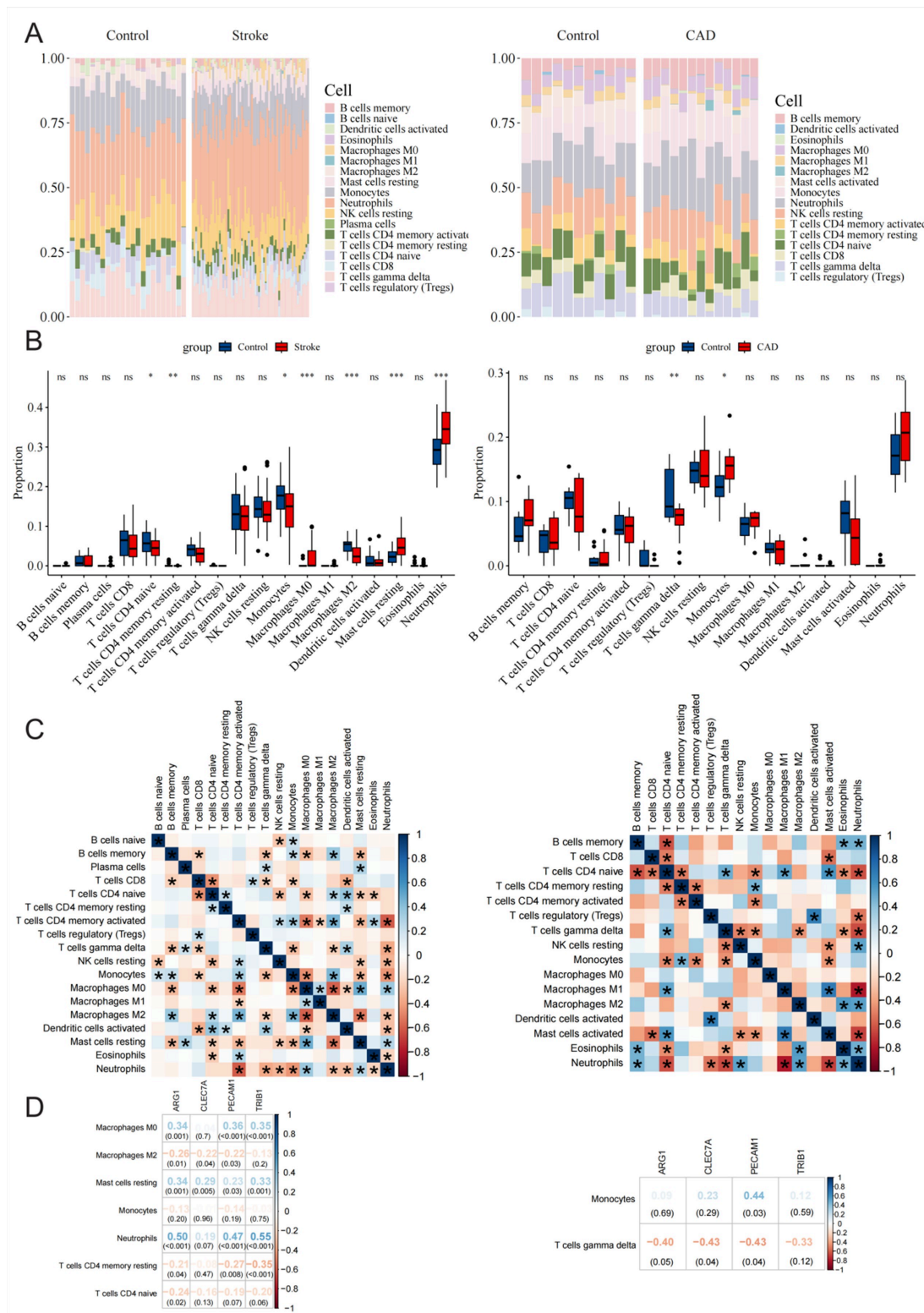


Figure 6

Comprehensive characterization of the immune microenvironment in cerebral ischemic stroke and coronary artery disease comorbidity.

(A) Stacked bar plots showing the composition of immune cell infiltration in cerebral ischemic stroke (CIS), coronary artery disease (CAD), and control samples.

(B) Box plots quantifying significant differences in infiltration levels of key immune cell subsets between disease and control groups ($P < 0.05$, $P < 0.01$). CIS samples showed elevated M0 macrophages and neutrophils, while CAD samples exhibited altered gamma delta T cells and macrophage subpopulations.

(C) Correlation heatmap illustrating significant positive and negative interactions ($P < 0.05$) among immune cell types in the comorbid microenvironment.

(D) Scatter plots or heatmap demonstrating positive or negative correlations between hub gene expression (ARG1, CLEC7A, PECAM1, TRIB1) and infiltration levels of specific immune cells ($P < 0.05$), revealing mechanistic links between gene dysregulation and immune recruitment. Significant correlations are annotated with the Spearman correlation coefficient (r) and P-value (P).

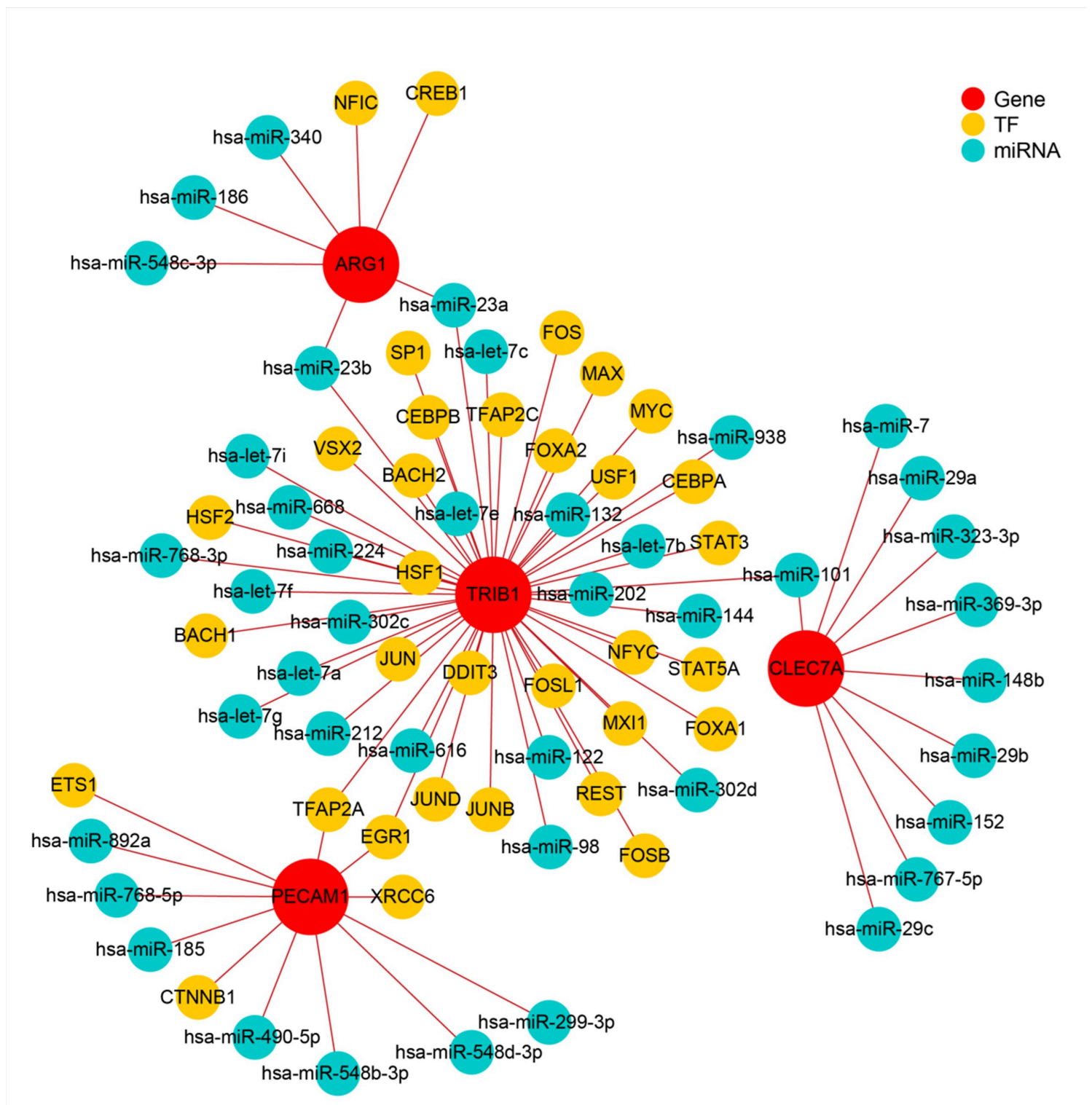


Figure 7

Transcriptional regulatory network of hub genes in cerebral ischemic stroke and coronary artery disease comorbidity.

The integrated miRNA–TF–mRNA regulatory network was constructed using Cytoscape. Red nodes represent core comorbidity-associated hub genes (ARG1, TRIB1, CLEC7A, PECAM1); yellow nodes denote transcription factors (e.g., NFIC, CREB1, FOS); cyan nodes indicate miRNAs (e.g., hsa-miR-340,

hsa-miR-186). Edges represent regulatory interactions: solid lines for TF-mediated gene regulation (e.g., NFIC and CREB1 regulating ARG1) and dashed lines for miRNA-directed post-transcriptional silencing (e.g., hsa-miR-340 and hsa-miR-186 targeting ARG1; hsa-miR-7 and hsa-miR-23a repressing TRIB1). The network identifies master regulators (CREB1, FOS) and non-coding RNA mediators as potential therapeutic targets for comorbidity intervention.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementalfiles.zip](#)