Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Integrative analysis of single-cell sequencing, bulk transcriptomics and experimental verification reveals key molecular mechanisms of vascular smooth muscle cells involved in intracranial aneurysm progression

**Yuhao He**
The Third People's Hospital of Chengdu
**Sunfu Zhang**
sunfuzh@163.com
The Third People's Hospital of Chengdu
**Shengming Liu**
The Third People's Hospital of Chengdu
**Qiang Li**
The Third People's Hospital of Chengdu
**Chunmiao Wu**
The Third People's Hospital of Chengdu

**Research Article**

**Keywords:** intracranial aneurysm, vascular smooth muscle cells, single-cell sequencing, immune infiltration, machine learning, molecular docking

**Posted Date:** April 8th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9028687/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

# Background and Purpose

Intracranial aneurysm (IA) represents a prevalent cerebrovascular disorder. Although VSMC dysfunction has been implicated as a central contributor to IA pathogenesis, the precise molecular underpinnings remain incompletely elucidated. This study sought to leverage multi-omics integration for characterizing VSMC-associated pivotal genes and their upstream regulatory architectures during IA progression.

# Methods

We obtained IA-related single-cell sequencing dataset GSE193533 and transcriptomic microarray datasets GSE75436 and GSE122897 from the GEO database. Seurat package was used for single-cell analysis. Limma package and WGCNA were used to obtain DEGs and IA-related hub genes. The intersection of the three gene sets was taken to obtain candidate genes, followed by GO and KEGG enrichment analysis. Machine learning methods were applied to screen key genes and construct a diagnostic prediction model. Immune infiltration analysis and ceRNA/transcriptional regulatory networks were performed. CMap and molecular docking predicted therapeutic drugs. Key genes were validated using qRT-PCR and Western blot analysis.

# Results

Compared to the normal group, the proportion of VSMCs gradually decreased in IA tissues. The intersection yielded 113 candidate genes, mainly enriched in neutrophil degranulation, lysosome, and other pathways. Machine learning screened out four key genes: COL5A1, IGFBP2, RASL12, and PLCB4. RT-qPCR and Western blot validation confirmed that COL5A1 and IGFBP2 were significantly upregulated while RASL12 and PLCB4 were significantly downregulated in IA samples at both mRNA and protein levels ($P < 0.001$). Immune analysis suggested that M0 macrophages and gamma delta T cells were significantly upregulated in the IA group, and key genes were significantly correlated with the infiltration of M2 macrophages and other immune cells. Furthermore, we constructed a ceRNA network centered on KCNQ1OT1 and identified key transcription factors such as SREBF1. Drug prediction yielded five candidate drugs, including carteolol. Key genes were validated in an independent dataset and at the single-cell level.

# Conclusion

This study constructed a multi-omics integrative analysis strategy, revealing the important role of VSMC dysfunction and related molecular events in IA development and progression, and discovered some potential markers and therapeutic targets, providing new insights for the diagnosis and treatment of IA.

# 1 Introduction

Intracranial aneurysms (IAs) are characterized by localized bulging or dilation of the intracranial arterial wall, which can be observed in approximately 2–5% of the general population and are recognized as a major cause of hemorrhagic stroke[1]. Although the majority of IAs remain asymptomatic, rupture of an aneurysm can lead to subarachnoid hemorrhage, resulting in mortality rates of approximately 50%, and imposing substantial economic and health burdens on patients and their families[2]. Therefore, comprehensive investigation of the pathogenesis of IAs and elucidation of the molecular events associated with their formation and progression are of significant importance for guiding early diagnosis and treatment strategies.

The development and progression of IAs are driven by a complex interplay of genetic and environmental factors, involving dysregulation of multiple cell types and signaling pathways[2]. Among these, dysfunction of vascular smooth muscle cells (VSMCs) is considered a critical pathological foundation for IA formation[3]. In normal arterial walls, VSMCs are distributed within the medial layer, where they maintain arterial wall integrity and structural stability through the synthesis and secretion of extracellular matrix proteins[3]. However, during IA pathogenesis, VSMCs undergo phenotypic transformation from a contractile to a synthetic phenotype, which results in extracellular matrix degradation, thinning of the medial layer, and destruction of the arterial wall structure, ultimately leading to aneurysm formation[3]. Previous studies have indicated that multiple signaling pathways and transcription factors are involved in the regulation of VSMC phenotypic switching and dysfunction, including the NF-κB signaling pathway, Smad pathway, and KLF4 transcription factor[4–6]. Nevertheless, due to the involvement of multiple cell types and complex molecular regulatory networks in IA pathogenesis, a comprehensive understanding of the precise molecular mechanisms by which VSMCs contribute to IA progression remains limited.

In recent years, the emergence of single-cell RNA sequencing (scRNA-seq) has revolutionized our understanding of disease mechanisms by enabling unprecedented resolution of cellular heterogeneity and dynamic states[7]. While conventional bulk transcriptomic studies in IA research have identified differentially expressed genes such as inflammatory factors and extracellular matrix degrading enzymes[8, 9], these approaches provide averaged expression profiles across mixed cell populations, potentially masking critical cell-type-specific changes and overlooking the functional heterogeneity within key cell populations like VSMCs. The transformative power of scRNA-seq technology lies in its ability to dissect gene expression profiles of individual cells and distinct subpopulations at single-cell resolution, enabling the identification of rare or transitional cell states, disease-associated cell subtypes, and their specific marker genes that would be undetectable in bulk analyses[10]. More importantly, scRNA-seq facilitates the reconstruction of intercellular communication networks, reveals dynamic signaling pathway alterations across different cell states, and provides insights into how cellular microenvironments influence disease progression—capabilities that are fundamentally limited in conventional transcriptomic approaches.

Given this research background, we integrated scRNA-seq and bulk transcriptomic data to explore VSMC functional changes and molecular mechanisms during IA development. Using public scRNA-seq data, we identified VSMC subpopulations and their marker genes across normal arteries, unruptured and ruptured aneurysms. Bulk RNA-seq analysis identified DEGs and hub genes through WGCNA. Integrating VSMC markers with DEGs/hub genes yielded candidate genes for functional enrichment analysis. Machine learning algorithms identified key diagnostic genes, while immune infiltration and regulatory network analyses revealed upstream mechanisms. Potential therapeutic drugs were predicted via CMap and molecular docking, with validation in independent datasets. This multi-omics approach provides comprehensive insights into VSMC dysfunction in IA pathogenesis, identifying potential biomarkers and therapeutic targets for precision medicine strategies. In conclusion, through the integration of multi-omics data and the application of various bioinformatics approaches, our study provides a comprehensive analysis of the molecular mechanisms underlying VSMC dysfunction in IA pathogenesis. The identified key genes and pathways may serve as potential diagnostic biomarkers and therapeutic targets for IAs, contributing to the development of precision medicine strategies for this disease.

## 2 Materials and Methods

# 2.1 Data download and preprocessing

Two bulk transcriptomic datasets pertaining to IA, namely GSE75436 and GSE122897, were retrieved from the GEO database (https://www.ncbi.nlm.nih.gov/geoprofiles/). GSE75436, profiled on the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array), comprised transcriptomic profiles of 15 IA specimens and 15 matched superficial temporal artery wall specimens, serving as the training cohort. GSE122897, generated on the GPL16791 platform (Illumina HiSeq 2500), encompassed 44 IA specimens alongside 16 intracranial dura mater artery specimens and was designated as the external validation cohort for evaluating reproducibility. Standard preprocessing procedures, including background correction, probe-to-gene symbol mapping, and quantile normalization, were applied to both datasets.

The scRNA-seq dataset GSE193533 was also obtained from the GEO repository. Three samples were included: a sham-operated control (GSM5813881), an unruptured IA sample (formed group, GSM5813883), and a ruptured IA sample (ruptured group, GSM5813885), all derived from murine Willis circle vasculature. Raw sequencing reads were processed through the 10x Genomics Cell Ranger pipeline for genome alignment, transcript quantification, and barcode demultiplexing. The resulting gene-cell count matrix was imported into R via the Seurat package, and only cells harboring 500–7,000 detected genes (nFeature_RNA) were retained for downstream analyses.

# 2.2 Single-cell analysis

Quality-controlled single-cell expression matrices were processed with the Seurat package (version 4.0.1). Library-size normalization was conducted via the "LogNormalize" approach, after which the "FindVariableFeatures" function was employed to rank genes by expression variability, retaining the top

2,000 most variable features. Following mean-centering and variance-scaling through "ScaleData", PCA was executed for linear dimensionality reduction.

Cell clustering was achieved by first constructing a shared nearest neighbor (SNN) graph via "FindNeighbors" and then partitioning cells into discrete subpopulations using the Louvain algorithm implemented in "FindClusters". Cell-type identity was assigned by cross-referencing cluster-specific expression profiles against established marker gene panels using the SingleR package. Non-linear embedding was generated with "RunUMAP" to produce two-dimensional UMAP representations for visualization of intercellular relationships.

To characterize cell-type compositional shifts across disease stages, the relative abundance of each annotated population was compared among the sham, formed, and ruptured groups. Subpopulations exhibiting notable alterations were subjected to differential expression testing via the "FindMarkers" function, where the Wilcoxon rank-sum test was applied with thresholds of adjusted $P < 0.05$ and $|avg_log_2FC| > 1$ to delineate marker genes..

# 2.3 Differential expression analysis

The limma package in R was applied to the training cohort (GSE75436) to detect transcriptomic differences between IA and control specimens. Genes meeting the dual criteria of $|log_2FC| > 1.5$ and $P < 0.05$ were classified as DEGs. Volcano plots and heatmaps were generated to provide a global overview of expression alterations.

# 2.4 Weighted gene co-expression network analysis (WGCNA)

WGCNA was implemented on the GSE75436 dataset to delineate co-expression modules linked to IA phenotype. The optimal soft-thresholding power β was determined via the pickSoftThreshold function by jointly evaluating the scale-free topology fit index and mean network connectivity. An adjacency matrix was derived using the selected β and subsequently converted into a TOM to capture pairwise co-expression proximity. Modules were delineated by performing hierarchical clustering on the TOM-based dissimilarity matrix, followed by application of the dynamic tree cut algorithm, with minimum module size and merge cut height set at 30 and 0.25, respectively. Each module's expression pattern was summarized by its ME, defined as the first principal component. Module–trait associations were quantified via Pearson correlation between MEs and the binary IA trait, and modules reaching $P < 0.05$ were deemed disease-relevant. Within these modules, genes satisfying $|MM| > 0.5$ and $|GS| > 0.5$ were designated as hub genes.

A candidate gene set was then assembled by intersecting the DEGs, VSMC-derived marker genes from scRNA-seq, and WGCNA hub genes. This integrative filtering strategy was designed to prioritize genes concurrently associated with VSMC biology and IA pathology.

# 2.5 Functional enrichment analysis of candidate genes

The DAVID platform (https://david.ncifcrf.gov) was utilized for GO and KEGG pathway analyses of the candidate gene set. GO annotation encompassed BP, CC, and MF categories. Enrichment significance was assessed by Fisher's exact test, with $P < 0.05$ as the cutoff. Complementary GSEA was conducted to evaluate the directional expression patterns of the top enriched pathways between IA and control groups.

## 2.6 PPI network analysis

Candidate genes were submitted to the STRING database (https://cn.string-db.org/) with a minimum interaction confidence of 0.4. The resultant network was imported into Cytoscape for topological characterization, wherein degree, betweenness centrality, and closeness centrality were computed to rank node importance.

## 2.7 Key feature gene screening using machine learning algorithms

Three complementary algorithms were deployed to distill the most discriminative diagnostic features from the candidate gene pool. The RF algorithm, executed via the "randomForest" package, ranked gene importance by mean decrease in Gini impurity. SVM-RFE, implemented through the "e1071" package, recursively eliminated the least informative features based on classifier-derived weight coefficients. LASSO logistic regression, performed with the "glmnet" package, imposed an L1 penalty to shrink non-contributory coefficients to zero. Genes consistently selected across all three methods were retained as key diagnostic markers.

## 2.8 Construction of a risk prediction model based on key genes

A logistic regression–based nomogram was developed using the identified key genes as predictor variables. The dataset was partitioned into training (70%) and testing (30%) subsets. Regression coefficients were estimated by maximum likelihood, and their statistical relevance was evaluated via the Wald test. Each gene was assigned a weighted score in the nomogram proportional to its coefficient, and individual scores were aggregated to yield a composite risk probability. Model calibration was appraised through calibration curves comparing predicted against observed event rates, and clinical applicability was assessed via DCA to quantify net benefit across a spectrum of threshold probabilities.

## 2.9 Immune infiltration analysis

The CIBERSORTx algorithm (https://cibersortx.stanford.edu/) was employed to deconvolve the relative fractions of 22 immune cell subsets in each sample. Differences in cell-type abundance between IA and control groups were evaluated by Wilcoxon rank-sum test. Pearson correlation analysis was further applied to examine inter-cell-type associations and the relationship between key gene expression and immune cell infiltration levels.

## 2.10 ceRNA network construction and transcription factor (TF) prediction

A ceRNA network was assembled by first querying the miRWalk database (http://mirwalk.umm.uni-heidelberg.de/) for miRNAs targeting the key genes and then retrieving their upstream lncRNA regulators from the starBase database (http://starbase.sysu.edu.cn/). The resulting lncRNA–miRNA–mRNA interaction map was visualized in Cytoscape.

Transcription factor enrichment analysis was carried out via NetworkAnalyst (https://www.networkanalyst.ca/) using JASPAR-curated binding site information. Hypergeometric testing was performed to evaluate over-representation of transcription factor binding motifs in the promoter regions of key genes ($P < 0.05$).

## 2.11 Drug prediction and molecular docking analysis

Upregulated and downregulated gene lists were submitted to the CMap platform (https://clue.io) to identify small molecules with potential disease-reversing signatures, ranked by negative connectivity score. Candidate compounds were then subjected to molecular docking using AutoDock Vina. Three-dimensional ligand structures were retrieved from PubChem, and both ligands and receptor proteins were preprocessed with AutoDockTools (hydrogen addition, Gasteiger charge assignment). Docking grids were defined around the predicted binding regions, and binding free energies and interaction modes (hydrogen bonds, hydrophobic contacts) were computed. The top-scoring ligand–receptor complexes for each key target were visualized to illustrate binding geometry.

## 2.12 Expression validation and ROC analysis of key genes

Expression levels of the key genes were compared between IA and control groups in both the training (GSE75436) and validation (GSE122897) cohorts using the Wilcoxon rank-sum test. Diagnostic discriminative capacity was evaluated by ROC curve analysis with AUC computation. At the single-cell level, expression patterns of key genes were projected onto the UMAP embedding of the GSE193533 dataset, and inter-group differences among the sham, formed, and ruptured conditions were statistically assessed.

## 2.13 qRT-PCR

RNA was isolated with TRIzol reagent (Invitrogen, USA) and reverse-transcribed using the PrimeScript RT reagent kit (Takara, Japan) per the manufacturer's instructions. Amplification was performed with TB Green Premix Ex Taq under the following thermal profile: 95°C for 30 s, then 40 cycles of 95°C for 5 s and 60°C for 30 s. Primer sequences are provided in Supplementary Table 1. Expression fold-changes were derived by the $2^{-\Delta\Delta Ct}$ method with GAPDH as the endogenous reference.

## 2.14 Western blot

Whole-cell lysates were prepared with RIPA buffer supplemented with protease and phosphatase inhibitor cocktails. Protein concentration was measured by BCA assay (Beyotime, Shanghai, China). Equal protein loads (30 μg) were resolved on 10% SDS-PAGE gels, electrotransferred to PVDF membranes (Millipore, Bedford, MA, USA), and blocked in 5% non-fat milk/TBST for 1 h at ambient temperature. Membranes were probed overnight at 4°C with primary antibodies against COL5A1, IGFBP2, RASL12, PLCB4, and GAPDH (Proteintech, Shanghai, China), followed by HRP-conjugated secondary antibody incubation for 1 h. Chemiluminescent signals were captured with the Tanon 5200 system (Tanon Science & Technology, Shanghai, China), and densitometric quantification was performed using ImageJ with GAPDH as the loading reference.

## 2.15 Statistical Analysis

All bioinformatic workflows were executed in R (version 4.2.2). Relative transcript abundance from qRT-PCR was calculated via the $2^{-\Delta\Delta Ct}$ approach. Between-group comparisons were performed with the Wilcoxon rank-sum test, and statistical significance was defined as $P < 0.05$.

# 3 Results

## 3.1 Single-cell data analysis of scRNA-seq

Following quality control with the Seurat pipeline, high-quality cells from the GSE193533 dataset (nFeature_RNA > 200, nFeature_RNA < 7,000, percent.mt < 20%) were retained and normalized (Supplementary Fig. 1A). Variance-based feature selection identified highly variable genes across the merged dataset (Supplementary Fig. 1B, 1C). PCA was performed on the integrated expression matrix, and the leading 18 statistically significant PCs ($P < 0.05$) were selected for downstream analysis (Supplementary Fig. 2A–2C). UMAP-based non-linear embedding resolved the cells into 19 distinct clusters (Supplementary Fig. 2D, 2E).

Cluster identities were assigned through literature-guided manual annotation, and canonical marker gene expression across cell types was depicted in a bubble chart (Fig. 1A). Ten major populations were delineated: VSMC, monocyte/macrophage (Mo/M), neutrophil, fibroblast, endothelium, T lymphocyte, pericyte, B lymphocyte, dendritic cell, and mast cell (Fig. 1B). Compositional analysis across disease stages revealed a striking and progressive reduction in VSMC representation—from 66% in the sham group to 46% in the formed group and 17% in the ruptured group (Fig. 1C–E)—indicating that VSMC depletion accompanies IA development and rupture. Differential expression testing via FindAllMarkers yielded 2,936 VSMC-enriched marker genes ($P < 0.05$) for subsequent integrative analyses.

## 3.2 Differential analysis of transcriptomic data

Comparison of the IA and control groups in the GSE75436 cohort identified 4,307 DEGs, of which 2,325 were upregulated and 1,982 were downregulated (Fig. 2A). A heatmap depicting the leading 20 upregulated and 20 downregulated transcripts is presented in Fig. 2B.

## 3.3 WGCNA

Soft-thresholding power selection and adjacency-to-TOM transformation were executed as described (Supplementary Fig. 3A, 3B). Hierarchical clustering coupled with dynamic tree cutting yielded 11 co-expression modules (Supplementary Fig. 3C), each exhibiting inter-module distances exceeding 0.25, confirming module independence (Supplementary Fig. 3D). Correlation analysis between MEs and disease status pinpointed five modules—Black, darkgrey, magenta, darkturquoise, and midnightblue—as significantly associated with IA ($P < 0.05$), from which 547 hub genes were extracted (Supplementary Fig. 3E). Overlapping these hub genes with the DEGs and VSMC marker genes produced 113 candidate genes at the triple intersection (Supplementary Fig. 3F), representing a prioritized set of genes implicated in both VSMC biology and IA pathogenesis.

## 3.4 Functional characterization and interaction network of candidate genes

GO enrichment of the 113 candidates revealed predominant involvement in neutrophil degranulation, immune-associated neutrophil activation, calcium ion homeostasis, and actin filament regulation at the BP level (Fig. 3A). Enriched CC terms encompassed vacuolar lumen, lysosomal lumen, azurophil granule, myofibril, and secretory granule membrane (Fig. 3B), while prominent MF terms included immune receptor activity, phospholipase activity, and structural constituent of muscle (Fig. 3C). KEGG mapping highlighted lysosome, complement and coagulation cascades, platelet activation, lipid and atherosclerosis, hematopoietic cell lineage, sphingolipid signaling pathway, chemokine signaling pathway, neutrophil extracellular trap formation, vascular smooth muscle contraction, and apoptosis as the most significantly enriched pathways (Fig. 3D, 3E). GSEA corroborated that the majority of these pathways were markedly activated in the IA group ($NES > 1$, $P < 0.05$), with the notable exception of vascular smooth muscle contraction, which was significantly suppressed ($NES < 1$, $P < 0.05$) (Fig. 3F). The PPI network comprised 109 protein nodes interconnected by 1,426 edges; topological analysis highlighted CD68, CSF1R, TYROBP, HCLS1, APOE, CTSB, IGFBP2, IFI30, NCF4, PLCB4, RASL12, C1QA, CORO1A, CTSD, and HCK as high-degree nodes with potential regulatory significance (Fig. 3G).

## 3.5 Machine learning screening of key genes

Three machine learning strategies were applied to the 113 candidates. RF importance ranking retained 28 features (Fig. 4A); SVM-RFE selected 20 features through iterative backward elimination (Fig. 4B, 4C); and LASSO regression identified 6 features at the optimal penalization parameter (Fig. 4D, 4E). The consensus of these three outputs converged on four genes: COL5A1, IGFBP2, RASL12, and PLCB4 (Fig. 4F). A nomogram incorporating these four predictors was subsequently built. Elevated COL5A1 and IGFBP2 expression, together with diminished RASL12 and PLCB4 expression, conferred increased IA risk, with RASL12 downregulation exerting the greatest predictive weight (Fig. 5A). The calibration curve demonstrated close agreement between estimated and observed probabilities (Fig. 5B), and DCA confirmed that the nomogram delivered a superior net benefit over treat-all and treat-none strategies across a broad range of decision thresholds (Fig. 5C).

## 3.7 Immune infiltration analysis

CIBERSORTx deconvolution quantified 22 immune cell fractions per sample (Fig. 6A, 6B). Among inter-cell-type correlations, gamma delta T cells and activated CD4 memory T cells exhibited the strongest positive association (r = 0.75), whereas M2 macrophages and gamma delta T cells displayed the strongest inverse relationship (r = − 0.54) (Fig. 6C). Between-group comparisons demonstrated significant enrichment of M0 macrophages and gamma delta T cells in IA specimens ($P < 0.05$), alongside significant depletion of memory B cells, M2 macrophages, resting mast cells, plasma cells, resting CD4 memory T cells, and naive CD4 T cells ($P < 0.05$) (Fig. 6D). Correlation of key gene expression with immune fractions revealed that COL5A1 and IGFBP2 were positively associated with gamma delta T cells and activated mast cells yet negatively associated with M2 macrophages and resting mast cells. Reciprocal patterns were observed for RASL12 and PLCB4 ($P < 0.05$) (Fig. 6E–6H), implicating these genes in the modulation of immune cell composition during IA progression.

## 3.8 Upstream regulatory network delineation

The integrated ceRNA network comprised 72 nodes (4 mRNAs, 19 miRNAs, 80 lncRNAs) connected by 439 edges (Fig. 7A). Topological assessment identified KCNQ1OT1 as the highest-degree lncRNA hub, the let-7 family members (hsa-let-7b-5p, hsa-let-7a-5p, hsa-let-7i-5p) as the most connected miRNAs, and COL5A1 and IGFBP2 as the most interconnected mRNA targets.

Transcription factor enrichment via NetworkAnalyst/JASPAR uncovered 25 significantly over-represented binding motifs in the promoter regions of the key genes (Fig. 7B). Leading candidates included SREBF1, TFAP2A, HNF4A, GATA2, NFKB1, PAX2, HINFP, USF2, PPARG, and FOXC1, each potentially contributing to transcriptional dysregulation in IA.

## 3.9 Candidate drug identification and docking validation

CMap analysis prioritized five small molecules exhibiting the most negative connectivity scores: carteolol, siguazodan, vorinostat, dacomitinib, and BRD-K36796217. Molecular docking confirmed favorable binding for all compound–target pairs, with binding energies uniformly below − 5 kcal/mol (Supplementary Table 2). The strongest interactions per target were BRD-K36796217–COL5A1 (-7.822 kcal/mol), carteolol–IGFBP2 (-7.927 kcal/mol), vorinostat–RASL12 (-8.884 kcal/mol), and dacomitinib–PLCB4 (-8.726 kcal/mol). Representative docking poses illustrating key hydrogen-bond and hydrophobic contacts are presented in Fig. 8A–D.

## 3.10 Multi-level expression validation and diagnostic assessment

In both the training (GSE75436) and validation (GSE122897) cohorts, COL5A1 and IGFBP2 were significantly elevated in IA relative to controls, whereas RASL12 and PLCB4 were significantly reduced ($P < 0.05$) (Fig. 9A, 9B). ROC analysis yielded AUC values exceeding 0.7 for all four genes in both cohorts (Fig. 9C, 9D), supporting robust diagnostic performance.

At single-cell resolution (GSE193533), projection of key gene expression onto the UMAP embedding (Supplementary Fig. 4A) confirmed that COL5A1 and IGFBP2 were upregulated, while RASL12 and PLCB4 were downregulated, in the formed and ruptured groups relative to the sham group ($P < 0.05$) (Supplementary Fig. 4B), corroborating the bulk-level findings at cellular granularity..

## 3.11 Experimental confirmation by RT-qPCR and Western blot

RT-qPCR demonstrated that COL5A1 and IGFBP2 transcript levels were markedly elevated in IA specimens relative to healthy controls ($P < 0.001$), while RASL12 and PLCB4 mRNA abundance was substantially decreased ($P < 0.001$) (Fig. 10A). Immunoblotting yielded concordant protein-level changes: COL5A1 and IGFBP2 bands were conspicuously intensified, whereas RASL12 and PLCB4 signals were attenuated in the IA group (Fig. 10B). These results establish that the four key genes undergo consistent transcriptional and translational alterations in IA tissue, reinforcing their candidacy as disease-associated biomarkers.

# 4 Discussion

In this study, the molecular mechanisms of VSMC involvement in IA development were investigated through the integration of single-cell sequencing and bulk transcriptomic data. A progressive decrease in VSMCs was observed during IA formation and rupture (66% in sham group, 46% in formation group, 17% in rupture group), which is consistent with the findings of Chalouhi *et al.*[11] that VSMC phenotypic transformation is a hallmark event in early IA development[12]. It was further confirmed by Wang *et al.* [13] that VSMC dysfunction and depletion lead to decreased vascular wall stability, which are identified as key factors in aneurysm occurrence and rupture[14, 15]. The downregulation of the vascular smooth muscle contraction pathway in the IA group was further validated through GSEA analysis, supporting the central role of VSMC impairment in IAs.

Four key genes were selected through cross-validation using multiple machine learning algorithms: upregulated COL5A1 (encoding type V collagen α1 chain) and IGFBP2 (insulin-like growth factor binding protein 2), as well as downregulated RASL12 (small G protein) and PLCB4 (phospholipase Cβ4). COL5A1 has been shown to be involved in vascular remodeling, and its increased expression may represent a compensatory response to vascular wall injury[16]. IGFBP2 was found by Shen *et al.*[17] to promote VSMC transformation toward an inflammatory phenotype, accelerating aneurysm progression. RASL12 expression changes were identified in unstable atherosclerotic plaques and have been recognized as a potential diagnostic marker for atherosclerotic plaque instability[18], suggesting its important role in vascular disease development. Although limited research is currently available on the detailed mechanisms by which RASL12 directly regulates VSMCs, its molecular characteristics and expression patterns suggest it may be an important participant in VSMC functional regulation. It was observed by Paulhe *et al.*[19] that PLCB4 downregulation is closely associated with impaired VSMC contractile function, affecting vascular wall stability. In general, all four genes demonstrated good diagnostic value

in ROC analysis (AUC > 0.7) and were validated in independent datasets, indicating their potential as biomarkers for IAs. Following this, qRT-PCR and Western blot experiments were consistent with those observed in the bulk transcriptomic and single-cell transcriptomic analyses, further supporting the role of these genes in the regulation of VSMC function and aneurysm progression. This experimental validation strengthens the potential of COL5A1, IGFBP2, RASL12, and PLCB4 as potential biomarkers for IA, confirming their relevance in vascular disease development and highlighting their utility for future diagnostic and therapeutic strategies.

In the immune infiltration analysis, significant upregulation of M0 macrophages[20] and gamma delta T cells[21] was detected in IA tissues, while M2 macrophages were downregulated, indicating a shift of the immune microenvironment toward a pro-inflammatory state[22]. This is similar to Shao's proposal regarding the critical role of macrophage polarization imbalance in IAs[23]. According to recent studies, macrophage infiltration is closely associated with IA rupture[24], and pro-inflammatory M1 macrophages accelerate vascular wall degradation by secreting inflammatory factors and matrix metalloproteinases[25]. T cells were found by Munjal *et al.*[26] to secrete pro-inflammatory factors such as IL-17, triggering VSMC apoptosis and phenotypic transformation. Our correlation analysis showed that upregulated COL5A1 and IGFBP2 were positively correlated with pro-inflammatory T cells and negatively correlated with anti-inflammatory M2 macrophages, further supporting the importance of immune inflammation in IAs. GO and KEGG enrichment analyses revealed that candidate genes were primarily enriched in pathways such as neutrophil degranulation and lysosome function, while GSEA analysis confirmed significant upregulation of multiple inflammation-related pathways in the IA group, indicating the central role of inflammatory responses in IA development[27]. It is worth noting that the important role of inflammatory response in the progression of IA is also verified in multiple regulatory networks. Members of the miR-7 family have been identified as key regulators in intracranial aneurysm pathogenesis[28]. Previous studies have demonstrated that miR-7 can inhibit homocysteine-induced VSMC proliferation, migration, and inflammatory factor expression by targeting matrix metalloproteinase-14 (MMP-14) and inhibiting the TLR4/NF-κB signaling pathway miR-7[29].. Additionally, matrix metalloproteinases, including MMP-14 targeted by miR-7, have been shown to play crucial roles in extracellular matrix remodeling processes and vascular smooth muscle cell migration and proliferation, which are critical factors in aneurysm wall weakening and potential rupture[30]. It was also found by Zhang *et al.*[31] that when various pathogenic factors act, nuclear pore complex protein 98 (Nup98) promotes the migration of histidine triad nucleotide-binding protein 1 (HINT1) into the nucleus. The nuclear accumulation of HINT1 subsequently enhances binding with transcription activator protein-2α (TFAP2A). This intermolecular interaction initiates TFAP2A-mediated transcriptional regulation of the integrin α6 (ITGA6) gene. Following ITGA6 upregulation, downstream focal adhesion kinase (FAK) and signal transducer and activator of transcription 3 (STAT3) signaling cascades are activated. Ultimately, this molecular event chain induces phenotypic transformation of VSMCs, thereby exacerbating the formation and development of aortic aneurysms. These findings suggest that IA development involves complex transcriptional and post-transcriptional regulatory networks.

In the drug prediction analysis, five potential therapeutic agents (carteolol, siguazodan, vorinostat, dacomitinib, and BRD-K36796217) were identified, with molecular docking analysis indicating good binding affinity to key targets (all <-5 kcal/mol). Among these, carteolol (a β-receptor blocker) may inhibit the inflammatory activation of VSMCs through the cAMP/PKA/NF-κB pathway, as shown by Jacob *et al.* [32]. Vorinostat (SAHA), as a histone deacetylase (HDAC) inhibitor, plays an important role in regulating VSMC function through epigenetic mechanisms. Studies have shown that vorinostat can inhibit HDAC1/2/3 activity, increase histone H3 and H4 acetylation levels, thereby altering the chromatin structure and accessibility of specific genes[33]. In vascular smooth muscle cells, vorinostat can upregulate the expression of anti-proliferation genes such as KLF4 and p21 by inhibiting HDAC activity, while downregulating contractile markers such as smooth muscle α-actin (α-SMA) and smooth muscle myosin heavy chain (SM-MHC), promoting the transformation of VSMCs from a contractile to a synthetic phenotype[34]. Additionally, vorinostat can reduce the responsiveness of vascular smooth muscle cells to growth factors through epigenetic mechanisms, inhibiting cell proliferation and migration, and alleviating neointima formation and vascular remodeling[35]. However, the specific therapeutic effects of these drugs on IA require more research and data to elucidate and support their potential in the treatment of IA.

Through multi-omics analysis, VSMC dysfunction has been revealed to play a key role in IA pathogenesis, and related ceRNA and transcriptional regulatory networks have been constructed for the first time. However, several limitations should be acknowledged in this study. The relatively small sample size limits the generalizability of the findings, and there are inherent differences between in vitro models and in vivo conditions that may affect the relevance of the results. Additionally, species-specific differences may influence gene expression and functional outcomes, which needs to be considered when translating these findings to human studies. Furthermore, the absence of prospective clinical validation of these findings remains a major limitation.

Despite these challenges, this research provides new insights into early diagnosis and treatment of IAs. The integration of genomics and transcriptomics offers a valuable approach to understanding the molecular pathophysiology of IA, and through this strategy, potential diagnostic markers and therapeutic targets have been successfully identified. Future research should focus on validating the functional mechanisms of the key genes, exploring immune-VSMC interactions, and evaluating the efficacy of candidate drugs in preclinical and clinical settings. To improve the robustness and generalizability of these findings, the establishment of a clear research timeline, detailed experimental designs, and multi-center collaboration will be essential in further advancing the understanding and management of IA.

# Declarations

### Author Contributions

YH conceived and designed the study, performed the key analyses, interpreted the findings, and drafted the manuscript. SZ supervised the study, provided critical feedback, and revised the manuscript. SL and

QL collected and curated the data. CW contributed to study design and data interpretation. All authors discussed the results, contributed to manuscript revision, and approved the final version..

**Acknowledgments**

We would like to express our sincere gratitude to all the participants for their contributions to this study.

**Availability of data**

The findings presented in this study are supported by data incorporated within the manuscript. Source datasets utilized for analysis were accessed from publicly available repositories including Gene Expression Omnibus (GEO) databases and the Connectivity Map (CMap) platform. Any processed or derived datasets generated during our analytical procedures can be made available by contacting the corresponding author with a justified request. The transparency and accessibility of our research data adhere to standard scientific practices to facilitate result verification and potential further investigations by the scientific community.

**Conflicts of Interest**

The authors declare no conflicts of interest related to this study.

**Consent to Publish declaration**

Not applicable.

**Funding**

This study was supported by the Institutional Research Fund of The Third People's Hospital of Chengdu (Grant No. CSY-YW-03-2024-059), the Technology Support Program of Chengdu (Grant No. 2024-YF05-00748-SN), and The Third People's Hospital of Chengdu Clinical Research Program (Grant No. CSY-YN-01-2023-005).

**Ethics Statement**

**This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (IRB) of [**The Third People's Hospital of Chengdu**]. The approval number is [2025-S-367].**

**Clinical trial number: not applicable.**

# References

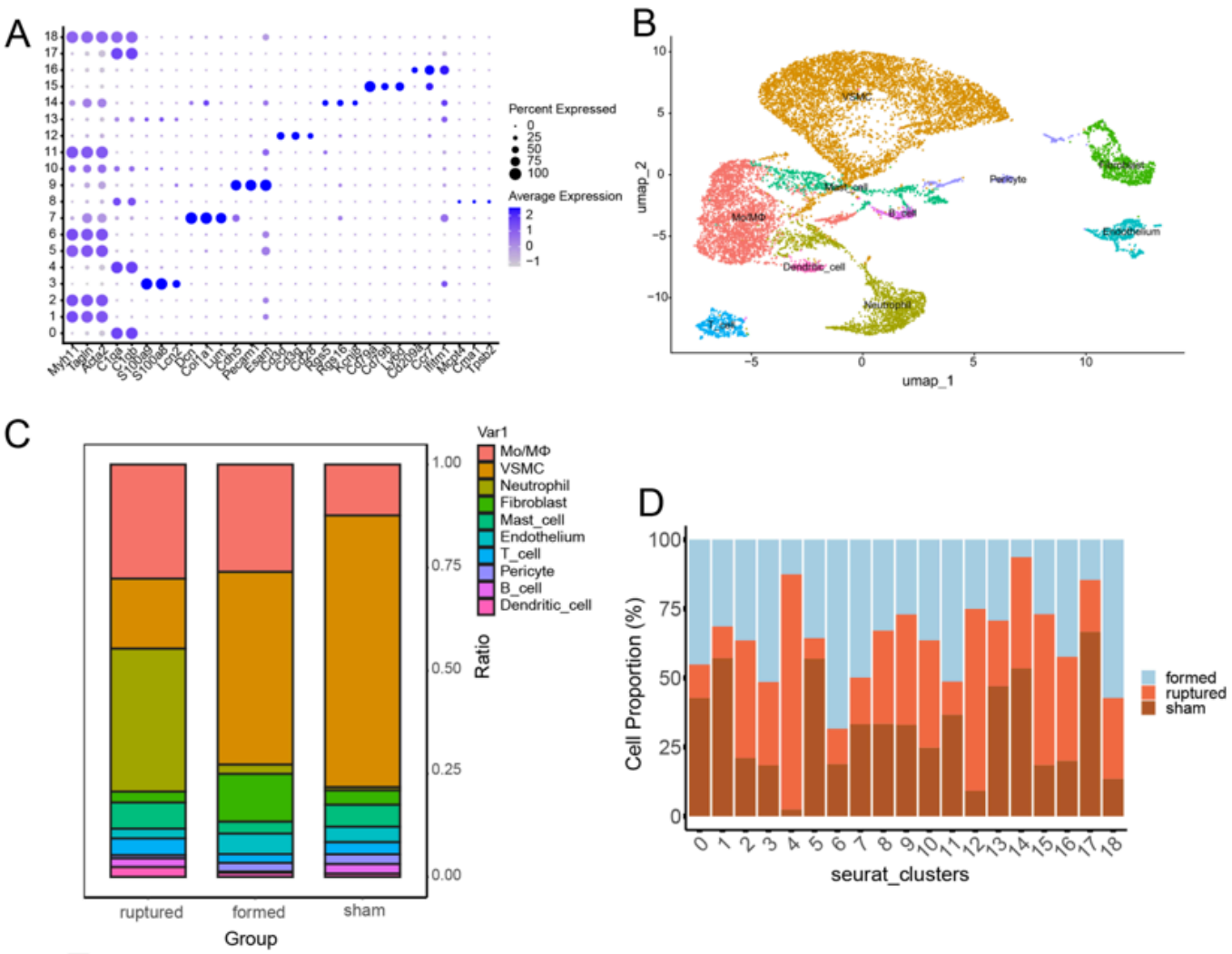
1. Texakalidis P, Sweid A, Mouchtouris N, Peterson E, Sioka C, Rangel-Castilla L, Reavey-Cantwell J, Jabbour P: **Aneurysm Formation, Growth, and Rupture: The Biology and Physics of Cerebral Aneurysms**. *World neurosurgery* 2019, **130**:277-284.

2. Etminan N, Dörfler A, Steinmetz H: **Unruptured Intracranial Aneurysms- Pathogenesis and Individualized Management**. *Deutsches Arzteblatt international* 2020, **117**(14):235-242.
3. Wang Z, Ma J, Yue H, Zhang Z, Fang F, Wang G, Liu X, Shen Y: **Vascular smooth muscle cells in intracranial aneurysms**. *Microvascular research* 2023, **149**:104554.
4. Khan D, Cornelius J, Muhammad S: **The Role of NF-κB in Intracranial Aneurysm Pathogenesis: A Systematic Review**. *International journal of molecular sciences* 2023, **24**(18).
5. Zhu L, Tang H, Wu C, Wei Y, Li Q, Dai D, Yang P, Huang Q, Xu Y, Liu J *et al*: **Activation of BMP4-pSmad1/5 pathway impairs the function of VSMCs in intracranial aneurysms**. *Vascular pharmacology* 2023, **153**:107236.
6. Ali M, Starke R, Jabbour P, Tjoumakaris S, Gonzalez L, Rosenwasser R, Owens G, Koch W, Greig N, Dumont A: **TNF-α induces phenotypic modulation in cerebral vascular smooth muscle cells: implications for cerebral aneurysm pathology**. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism* 2013, **33**(10):1564-1573.
7. Hong M, Tao S, Zhang L, Diao L, Huang X, Huang S, Xie S, Xiao Z, Zhang H: **RNA sequencing: new technologies and applications in cancer research**. *Journal of hematology & oncology* 2020, **13**(1):166.
8. Li L, Yang X, Jiang F, Dusting G, Wu Z: **Transcriptome-wide characterization of gene expression associated with unruptured intracranial aneurysms**. *European neurology* 2009, **62**(6):330-337.
9. Kurki M, Häkkinen S, Frösen J, Tulamo R, von und zu Fraunberg M, Wong G, Tromp G, Niemelä M, Hernesniemi J, Jääskeläinen J *et al*: **Upregulated signaling pathways in ruptured human saccular intracranial aneurysm wall: an emerging regulative role of Toll-like receptor signaling and nuclear factor-κB, hypoxia-inducible factor-1A, and ETS transcription factors**. *Neurosurgery* 2011, **68**(6):1667-1675; discussion 1675-1666.
10. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y: **Single-cell RNA sequencing technologies and applications: A brief overview**. *Clinical and translational medicine* 2022, **12**(3):e694.
11. Chalouhi N, Ali MS, Jabbour PM, Tjoumakaris SI, Gonzalez LF, Rosenwasser RH, Koch WJ, Dumont AS: **Biology of intracranial aneurysms: Role of inflammation**. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism* 2012(9):32.
12. Zhou D, Li T, Tan X, Yun C, Jiang P, Zhang T, Kuang H, Wang Y: **Secreted protein acidic and rich in cysteine (SPARC) induces apoptosis of human brain vascular smooth muscle cells through regulating HK2 in intracranial aneurysm**. *Front Mol Neurosci* 2023, **16**:1290556.
13. Wang Z, Ma J, Yue H, Zhang Z, Fang F, Wang G, Liu X, Shen Y: **Vascular smooth muscle cells in intracranial aneurysms**. *Microvasc Res* 2023, **149**:104554.
14. Martinez AN, Tortelote GG, Pascale CL, Ekanem UI, Leite APO, McCormack IG, Dumont AS: **Dimethyl Fumarate Mediates Sustained Vascular Smooth Muscle Cell Remodeling in a Mouse Model of Cerebral Aneurysm**. *Antioxidants (Basel)* 2024, **13**(7).

15. Lansdell TA, Fisher C, Simmonds K, Reeves MJ, Woo D, Dorrance AM, Demel SL: **Rs10230207 genotype confers changes in HDAC9 and TWIST1, but not FERD3L in lymphoblasts from patients with intracranial aneurysm**. *Neurogenetics* 2019, **20**(2):83-89.
16. Sheak JR, Jones DT, Lantz BJ, Maston LD, Vigil D, Resta TC, Resta MM, Howard TA, Kanagy NL, Guo Y *et al*: **NFATc3 regulation of collagen V expression contributes to cellular immunity to collagen type V and hypoxic pulmonary hypertension**. *Am J Physiol Lung Cell Mol Physiol* 2020, **319**(6):L968-l980.
17. Shen X, Xi G, Maile LA, Wai C, Rosen CJ, Clemmons DR: **Insulin-like growth factor (IGF) binding protein 2 functions coordinately with receptor protein tyrosine phosphatase β and the IGF-I receptor to regulate IGF-I-stimulated signaling**. *Mol Cell Biol* 2012, **32**(20):4116-4130.
18. Wang J, Kang Z, Liu Y, Li Z, Liu Y, Liu J: **Identification of immune cell infiltration and diagnostic biomarkers in unstable atherosclerotic plaques by integrated bioinformatics analysis and machine learning**. *Front Immunol* 2022, **13**:956078.
19. Paulhe F, Bogyo A, Chap H, Perret B, Racaud-Sultan C: **Vascular smooth muscle cell spreading onto fibrinogen is regulated by calpains and phospholipase C**. *Biochem Biophys Res Commun* 2001, **288**(4):875-881.
20. Okada A, Koseki H, Ono I, Kayahara T, Kurita H, Miyamoto S, Kataoka H, Aoki T: **Identification of The Unique Subtype of Macrophages in Aneurysm Lesions at the Growth Phase**. *J Stroke Cerebrovasc Dis* 2022, **31**(12):106848.
21. Zhou D, Zhu Y, Jiang P, Zhang T, Zhuang J, Li T, Qi L, Wang Y: **Identifying pyroptosis- and inflammation-related genes in intracranial aneurysms based on bioinformatics analysis**. *Biol Res* 2023, **56**(1):50.
22. Xu Y, Guo P, Wang G, Sun X, Wang C, Li H, Cui Z, Zhang P, Feng Y: **Integrated analysis of single-cell sequencing and machine learning identifies a signature based on monocyte/macrophage hub genes to analyze the intracranial aneurysm associated immune microenvironment**. *Front Immunol* 2024, **15**:1397475.
23. Shao L, Qin X, Liu J, Jian Z, Xiong X, Liu R: **Macrophage Polarization in Cerebral Aneurysm: Perspectives and Potential Targets**. *J Immunol Res* 2017, **2017**:8160589.
24. Han Y, Li G, Zhang Z, Zhang X, Zhao B, Yang H: **Axl promotes intracranial aneurysm rupture by regulating macrophage polarization toward M1 via STAT1/HIF-1α**. *Front Immunol* 2023, **14**:1158758.
25. Huang R, Sun Y, Liu R, Zhu B, Zhang H, Wu H: **ZeXieYin formula alleviates atherosclerosis by inhibiting the MAPK/NF-κB signaling pathway in APOE-/- mice to attenuate vascular inflammation and increase plaque stability**. *J Ethnopharmacol* 2024, **327**:117969.
26. Munjal A, Khandia R: **Atherosclerosis: orchestrating cells and biomolecules involved in its activation and inhibition**. *Adv Protein Chem Struct Biol* 2020, **120**:85-122.
27. Li C, Su Z, Su W, Wang Q, Wang S, Li Z: **Profiling of immune infiltration landscape of ruptured intracranial aneurysm**. *Medicine* 2024, **103**(12):9.

28. Wang K, Wang X, Lv H, Cui C, Leng J, Xu K, Yu G, Chen J, Cong P: **Identification of the miRNA-target gene regulatory network in intracranial aneurysm based on microarray expression data**. *Exp Ther Med* 2017, **13**(6):3239-3248.
29. Ma H, Wang L, Lv W, Lv Z: **Effects of miR-7 on Hcy-induced rat cerebral arterial vascular smooth muscle cell proliferation, migration and inflammatory factor expression by targeting MMP-14 to regulate TLR4/NF-κB signaling pathway**. *Cell Mol Biol (Noisy-le-grand)* 2020, **66**(7):12-17.
30. Atkinson G, Bianco R, Di Gregoli K, Johnson JL: **The contribution of matrix metalloproteinases and their inhibitors to the development, progression, and rupture of abdominal aortic aneurysms**. *Front Cardiovasc Med* 2023, **10**:1248561.
31. Zhang Y, Wu W, Yang X, Luo S, Wang X, Da Q, Yan K, Hu L, Sun S, Du X *et al*: **HINT1 aggravates aortic aneurysm by targeting ITGA6/FAK axis in vascular smooth muscle cells**. *J Clin Invest* 2025.
32. Jacob D, Shariff B, Bond B, Boyes N, Harper J: **β-Adrenergic Receptor Blockade Augments Sympathetically Mediated Vasoconstriction during Hypoxia in Healthy Young Adults**. *Physiology* 2024, **39**(S1):3.
33. Marks PA, Xu WS: **Histone deacetylase inhibitors: Potential in cancer therapy**. *J Cell Biochem* 2009, **107**(4):600-608.
34. Lacolley P, Regnault V, Segers P, Laurent S: **Vascular Smooth Muscle Cells and Arterial Stiffening: Relevance in Development, Aging, and Disease**. *Physiol Rev* 2017, **97**(4):1555-1617.
35. Findeisen HM, Gizard F, Zhao Y, Qing H, Heywood EB, Jones KL, Cohn D, Bruemmer D: **Epigenetic regulation of vascular smooth muscle cell proliferation and neointima formation by histone deacetylase inhibition**. *Arterioscler Thromb Vasc Biol* 2011, **31**(4):851-860.

# Figures




E

| | sham | formed | reptured |
|---|---|---|---|
| VSMC | 66% | 46% | 17% |
| Mo/MΦ | 12% | 26% | 28% |
| Fibroblast | 3% | 12% | 3% |
| Endothelium | 4% | 5% | 2% |
| Immune cells （myeloid and lymphoid) | 15% | 11% | 50% |
| Total cell counts | 4697 | 5530 | 4270 |

**Figure 1**

**Single-cell data clustering annotation and cell type distribution analysis.**

(A) Dot plot illustrating canonical marker gene expression across annotated cell clusters. Dot diameter reflects the fraction of expressing cells, and color gradient denotes mean expression intensity. (B) UMAP embedding of all cells, color-coded by assigned cell-type identity based on established lineage markers. (C) Bar chart quantifying the relative abundance of each cell type in the sham, formed, and ruptured

conditions. (D) Stacked bar representation showing the contribution of each disease stage within individual cell clusters. (E) Summary table listing the proportional composition of major cell populations (VSMCs, Mo/M, fibroblasts, endothelium, and immune cells) across the three experimental groups.

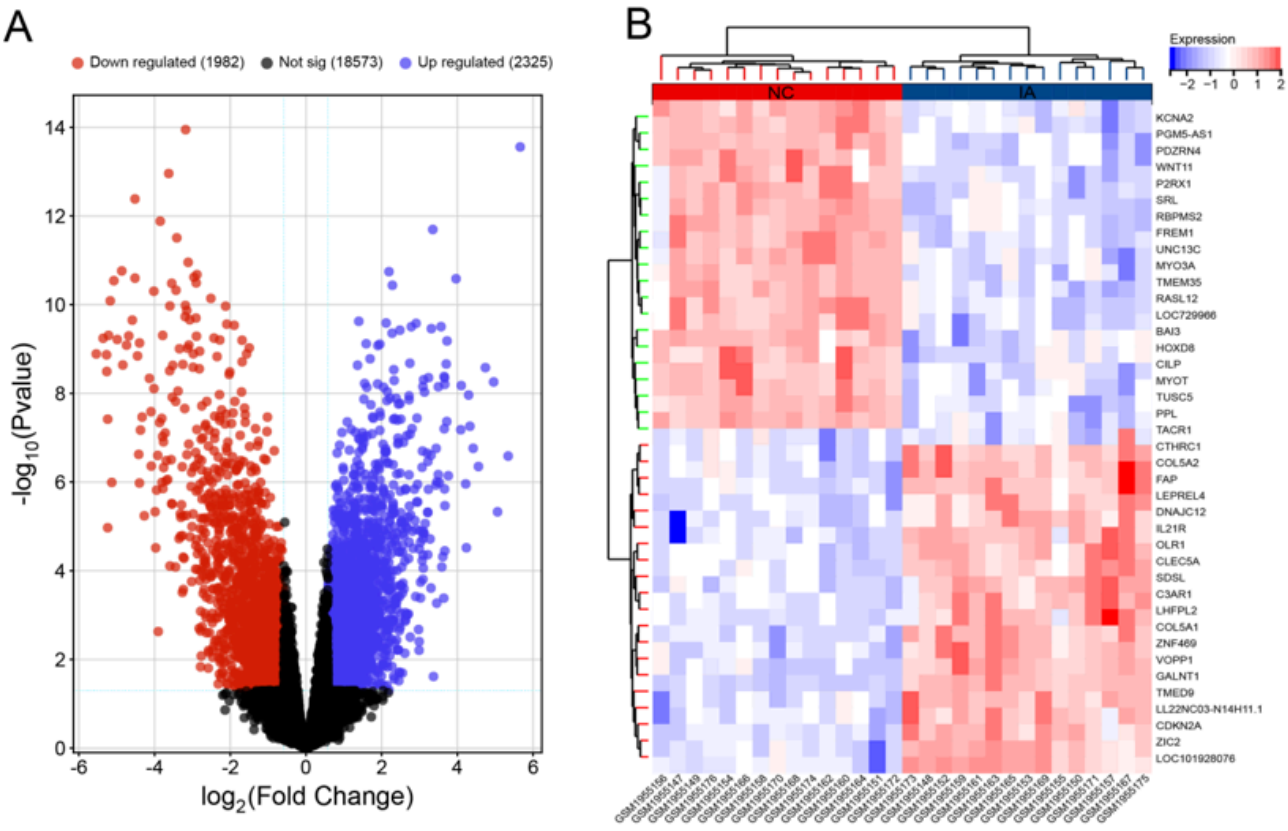
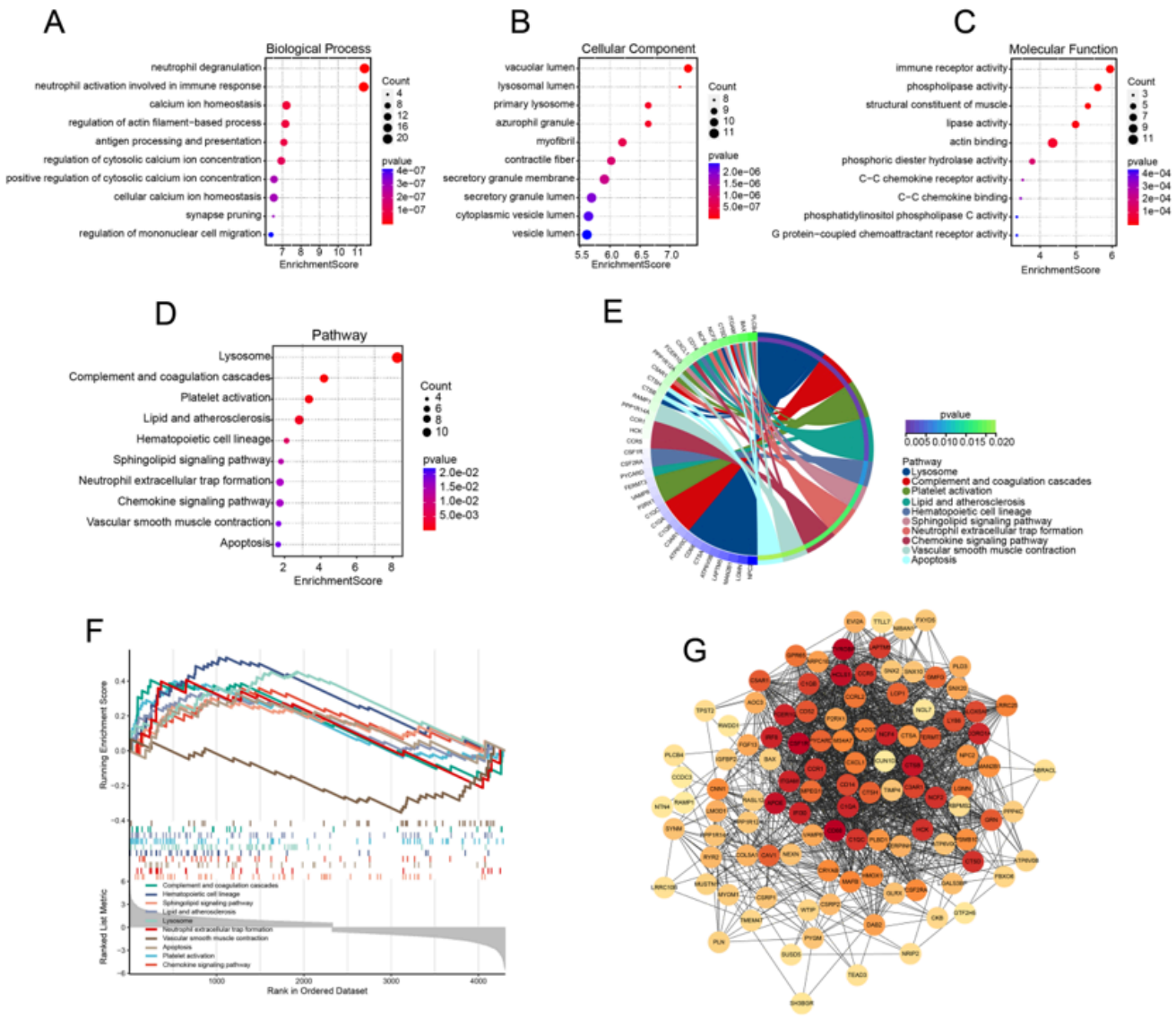



**Figure 2**

**Differential gene expression analysis between intracranial aneurysm (IA) and normal control (NC) tissues.**

(A) Volcano plot depicting genome-wide expression changes, with significantly upregulated and downregulated genes highlighted. (B) Heatmap visualizing the expression profiles of the top 20 upregulated and top 20 downregulated transcripts across all samples.




## Figure 3

**Functional enrichment analysis and protein-protein interaction (PPI) network.**

(A–C) Bubble charts presenting the leading 10 enriched GO terms for BP (A), CC (B), and MF (C). Bubble size corresponds to gene count, and color intensity reflects statistical significance. (D) Bubble chart of the top 10 significantly enriched KEGG pathways. (E) Gene–pathway association network illustrating the linkage between candidate genes and their enriched KEGG terms. (F) GSEA profiles of the top 10 differentially activated KEGG pathways in IA versus control comparisons ($P < 0.05$). (G) PPI map constructed from the candidate gene set. Node shading scales with connectivity degree, where darker tones signify higher interaction frequency.

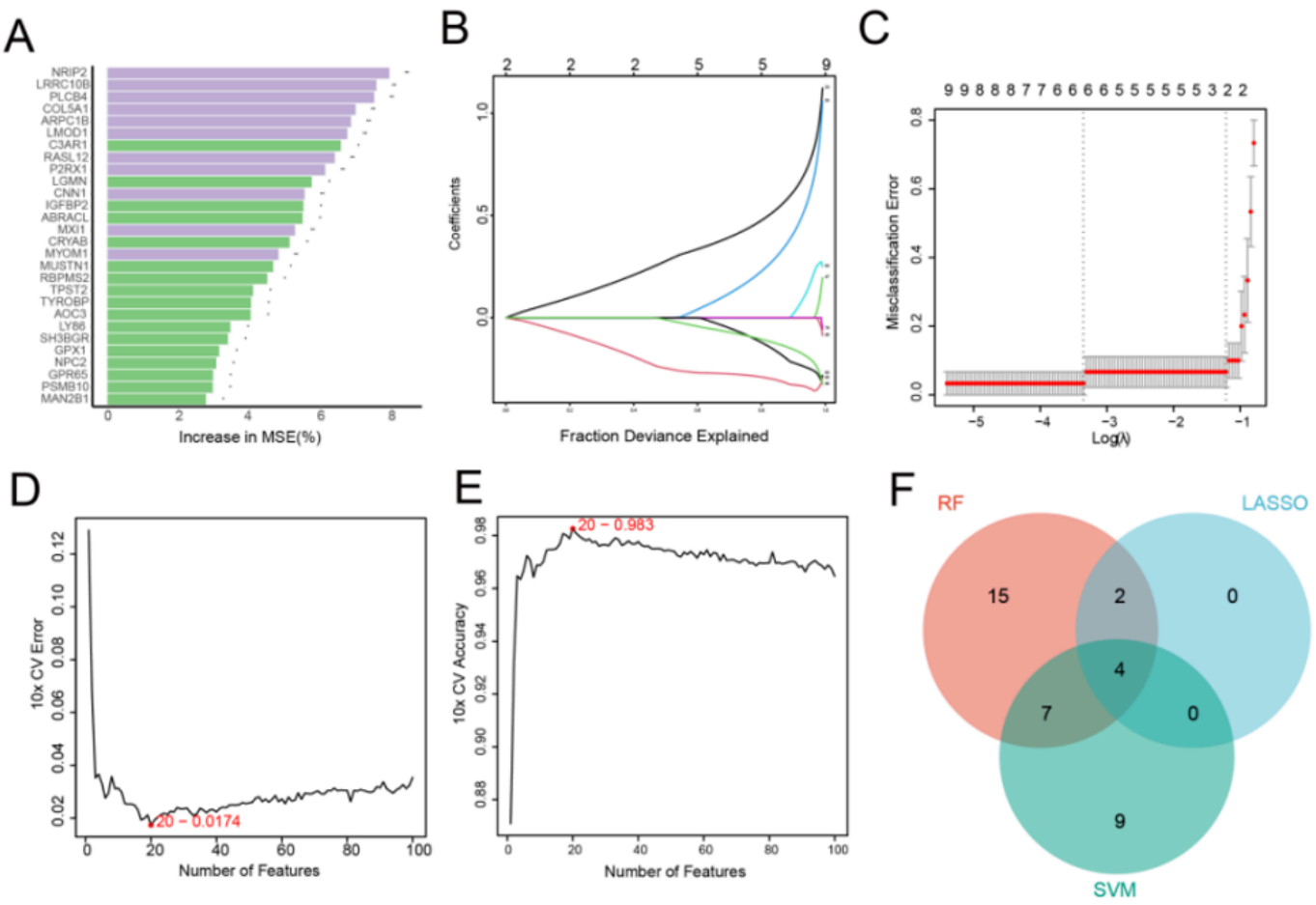



**Figure 4**

**Key gene identification using machine learning algorithms.**

(A) Gene importance ranking derived from the RF algorithm, displaying the top 28 contributors by mean decrease in Gini impurity. (B) Cross-validation accuracy curve from SVM-RFE as a function of retained feature number. (C) Feature priority list generated by the SVM-RFE procedure. (D) Regularization path of LASSO regression coefficients across varying penalty strengths. (E) Cross-validated error profile of the LASSO model plotted against log(λ); the vertical dashed line marks the optimal λ. (F) Venn diagram delineating the shared genes among RF, SVM-RFE, and LASSO outputs.

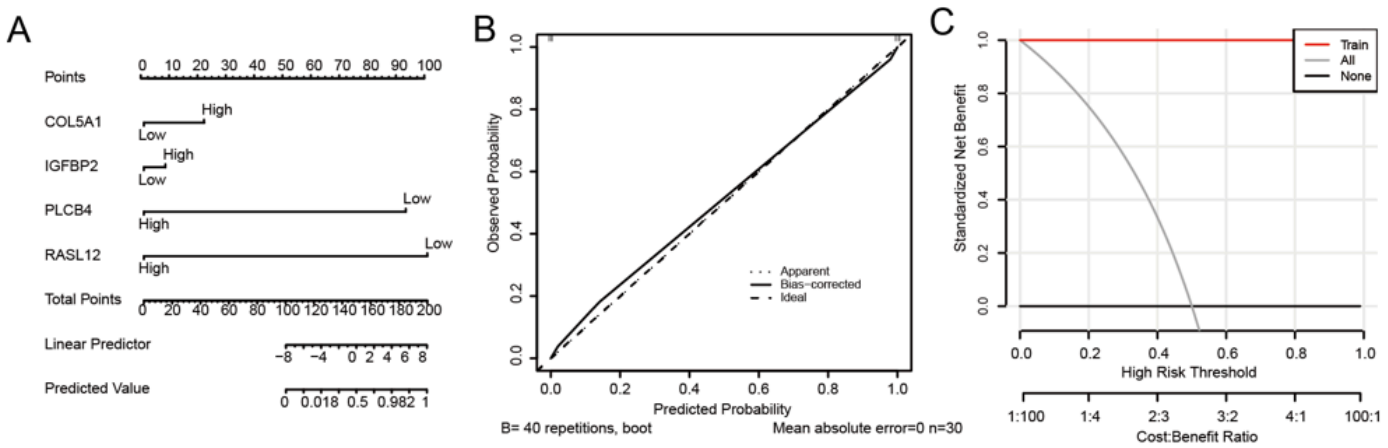
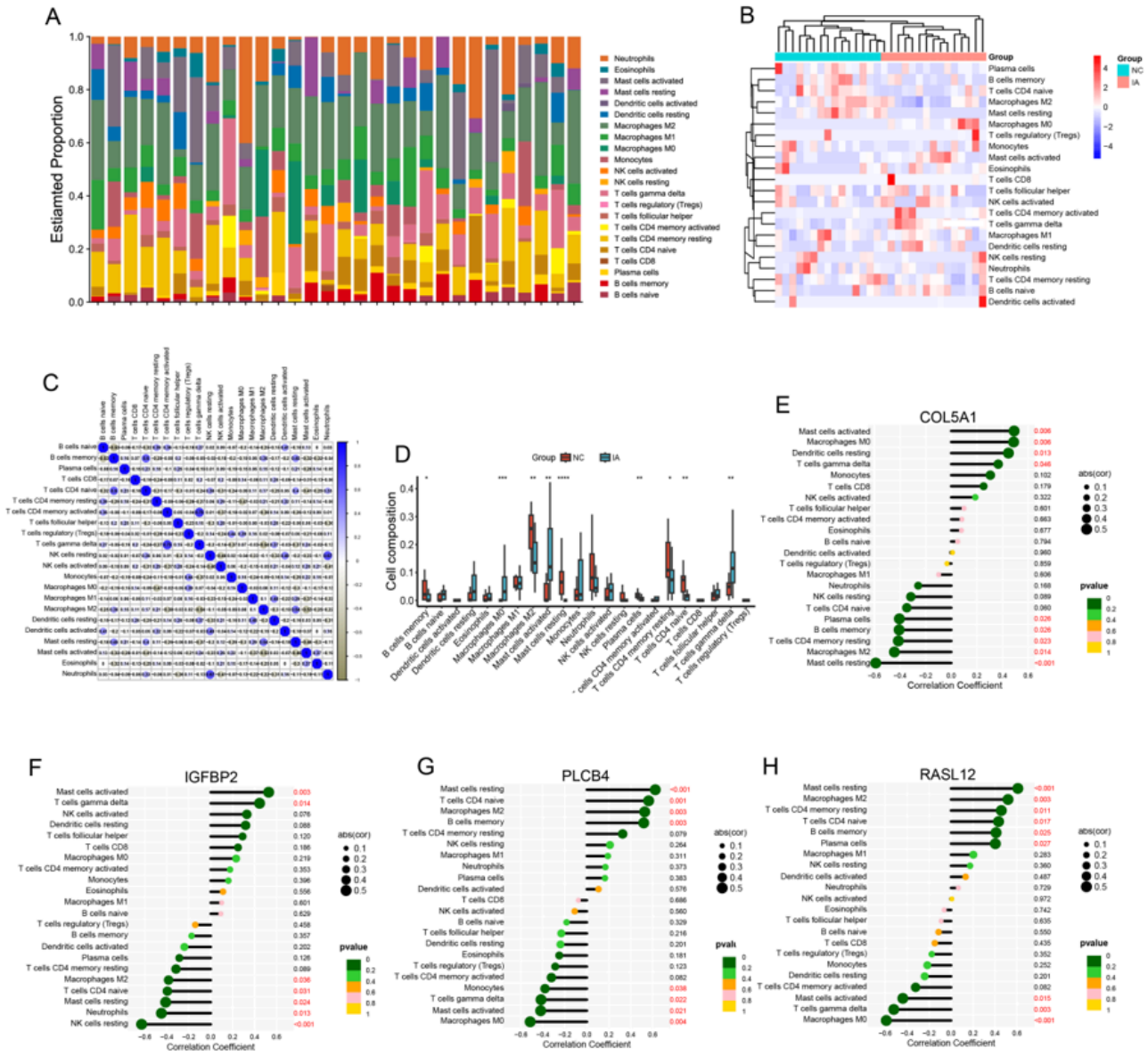



**Figure 5**

**Nomogram risk prediction model and evaluation.**

(A) Nomogram integrating four key predictors (COL5A1, IGFBP2, RASL12, and PLCB4) for estimating IA occurrence probability. (B) Calibration plot comparing model-predicted probabilities with observed event frequencies; the diagonal dashed line denotes ideal calibration. (C) DCA contrasting the net clinical benefit of the nomogram against treat-all and treat-none reference strategies across a range of threshold probabilities.




## Figure 6

**Immune cell infiltration analysis.**

(A) Stacked bar chart depicting the CIBERSORTx-estimated proportions of 22 immune cell subsets per sample. (B) Heatmap displaying the relative abundance of each immune population across all specimens. (C) Pairwise correlation matrix of immune cell fractions; circle size and color encode Pearson correlation coefficients. (D) Box plots contrasting the infiltration levels of significantly altered immune populations between IA and control groups. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. (E–H) Lollipop charts depicting the association between key gene expression (COL5A1, IGFBP2, RASL12, and PLCB4) and differentially abundant immune subsets. Horizontal axis indicates correlation coefficients; circle size and color represent statistical significance.

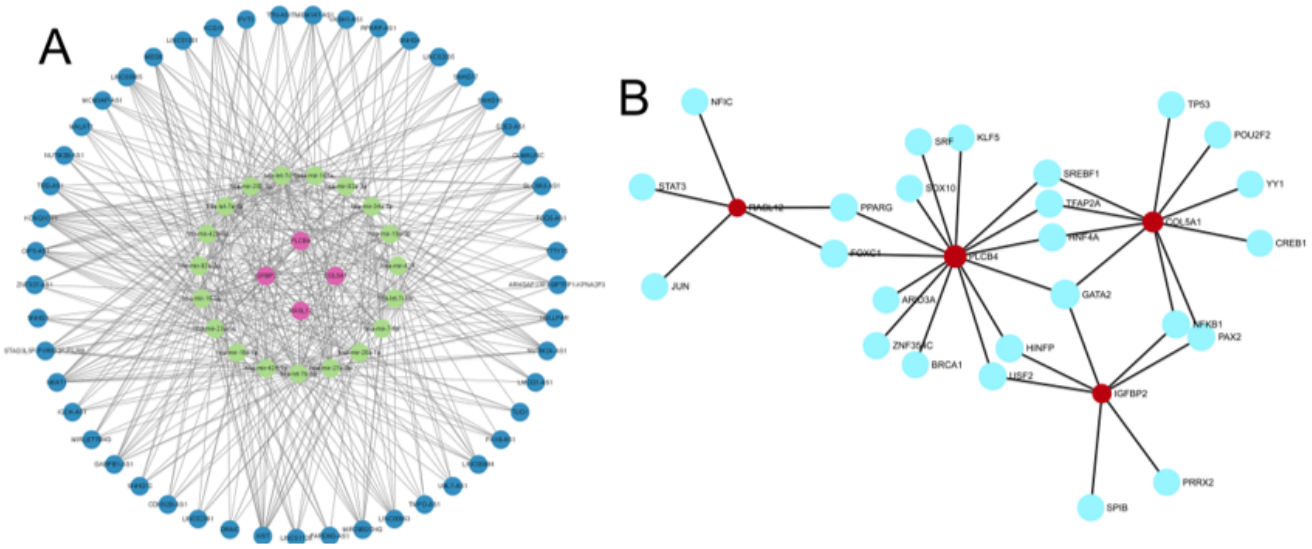



**Figure 7**

**Competing endogenous RNA (ceRNA) network and transcription factor (TF) prediction.**

(A) ceRNA interaction network. Blue, green, and red nodes denote lncRNAs, miRNAs, and mRNAs, respectively. Edge connections represent experimentally supported or computationally predicted regulatory pairs. (B) Transcription factor–target gene network. Cyan nodes indicate predicted transcription factors, and red nodes represent the four key genes.

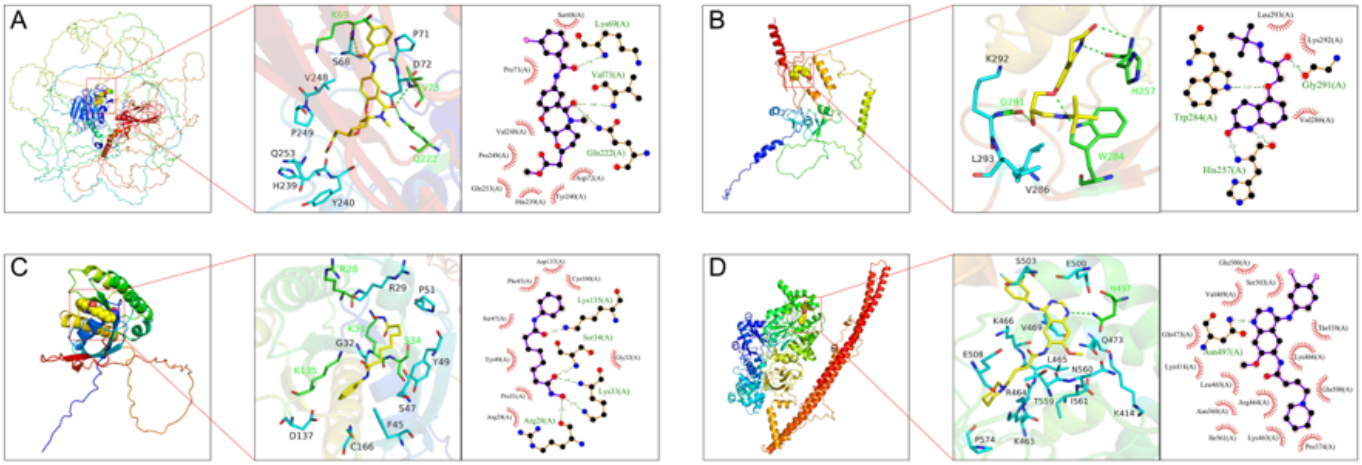



**Figure 8**

**Molecular docking analysis of small molecule drugs and key targets.**

(A–D) Three-dimensional depictions of the optimal docking conformations for COL5A1 (A), IGFBP2 (B), RASL12 (C), and PLCB4 (D), highlighting key intermolecular contacts between candidate small molecules and their respective protein targets.

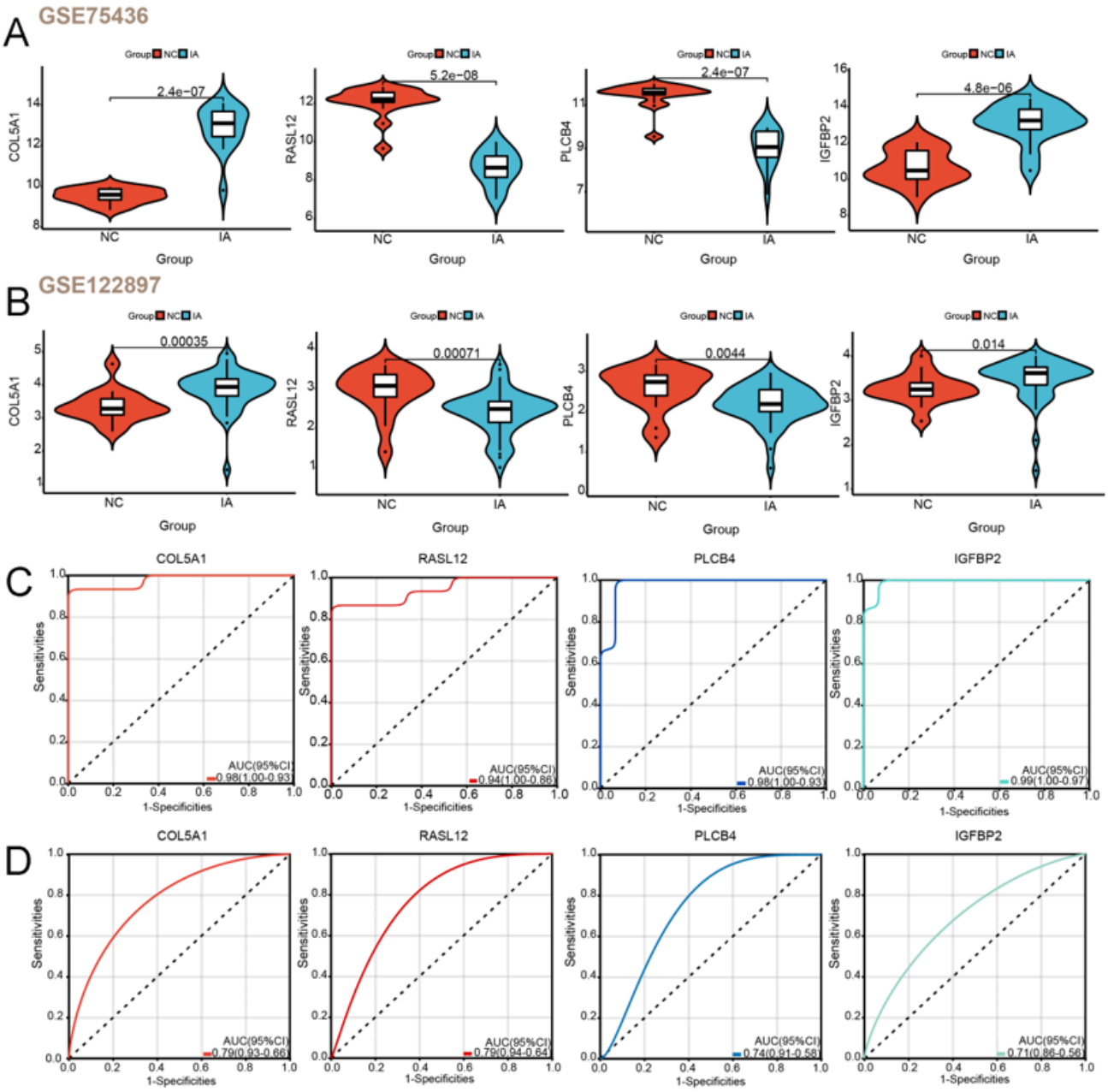



**Figure 9**

**Expression validation and receiver operating characteristic (ROC) curve analysis of key genes.**

(A, B) Violin plots contrasting key gene expression (COL5A1, IGFBP2, RASL12, and PLCB4) between IA and control specimens in the training cohort (A) and external validation cohort (B). (C, D) ROC curves assessing the discriminative capacity of each key gene in the training cohort (C) and validation cohort (D), with corresponding AUC values annotated in the legend.

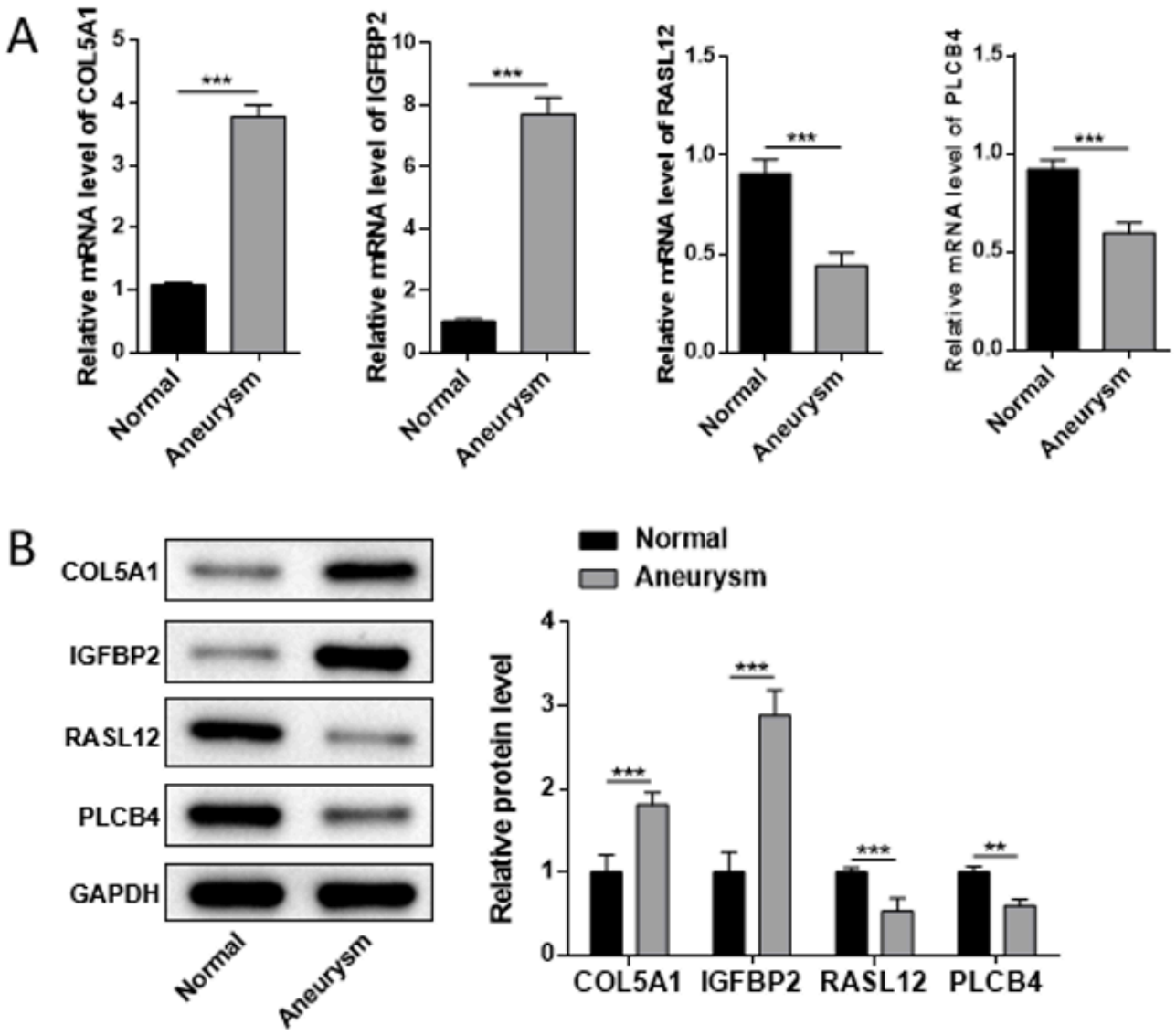



**Figure 10**

**Expression levels of COL5A1, IGFBP2, RASL12, and PLCB4 in normal and aneurysm groups.**

(A) Relative transcript abundance of COL5A1, IGFBP2, RASL12, and PLCB4 quantified by RT-qPCR in normal control and IA groups. (B) Representative immunoblot images and densitometric quantification of COL5A1, IGFBP2, RASL12, and PLCB4 protein levels. ***P < 0.001, **P < 0.01.

## Supplementary Files

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

- Supplementaryfigure.docx
- SupplementaryTable.docx
- SupplementaryFullLengthWBImages.pptx