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 bulk and single-cell transcriptome analyses reveal mitochondria and immune-related biomarkers in atherosclerosis with experimental validation

Zhecheng Xing
Kewei Zhang
Dongbin Zhang
Junhui Zhang
Weilin Wang
Yafei Qin
qinyafei92@tmu.edu.cn

Article

Keywords: Atherosclerosis, Mitochondria, Immunity, Single-cell sequencing analysis, Biomarkers

Posted Date: March 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9037999/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 The development of atherosclerosis (AS) is associated with mitochondrial function and immunity, yet the intrinsic mechanisms involved remain insufficiently elucidated. Therefore, the objective of the study was to explore mitochondria- and immune-related biomarkers in AS via single-cell RNA sequencing (scRNA-seq) and bulk RNA analysis.

Methods In this study, transcriptomic and scRNA-seq data of AS, mitochondria-related genes (MRGs), and immune-related genes (IRGs) were obtained from public databases. The research team pinpointed candidate genes by conducting both differential expression analysis and weighted gene co-expression network analysis (WGCNA). To identify potential biomarkers, they employed protein-protein interaction (PPI) analysis, machine learning techniques, and verified expression levels. Following these initial discoveries, the investigators carried out enrichment and immune infiltration analyses to assess how these biomarkers might contribute to the regulatory mechanisms of AS. Moreover, scRNA-seq analysis was performed to identify key cells. Ultimately, the reverse transcription quantitative PCR (RT-qPCR) was employed to validate the expression levels of biomarkers in clinical specimens.

Results The research successfully identified five biomarkers for AS, namely PLCβ4, RAC2, TYROBP, VAV1, and VCL. The dilated cardiomyopathy pathway showed positive enrichment in PLCβ4 and VCL, but negative enrichment in RAC2, TYROBP, and VAV1. Furthermore, 10 differential immune cells (DICs) were discerned between the AS and control groups. Seven DICs demonstrated notable correlations with five biomarkers, such as naive B cells. The scRNA-seq analysis identified seven distinct cell types, with macrophages, endothelial cells, and vascular smooth muscle cells (VSMCs) defined as the key cells. Finally, the RT-qPCR analysis revealed that the mRNA expression levels of RAC2, TYROBP, and VAV1 were significantly increased in the AS group, while PLCβ4 expression was markedly reduced.

Conclusion This study integrated bulk and scRNA-seq analyses to identify five candidate biomarkers, of which four were experimentally validated, along with three key cell types, providing novel insights into the molecular mechanisms and potential therapeutic targets for AS.

1. Introduction

Atherosclerosis (AS) is a chronic condition characterized by the deposition of lipids in the arterial walls, which eventually form plaques. The narrowing of the arterial lumen caused by these atherosclerotic plaques can lead to various clinical manifestations, such as angina pectoris and myocardial ischemia{Doran, #6}{Doran, #9}{Doran, #9}{Doran, #9}(Doran){Doran, #6}{Doran, #6}{Doran, #6}{Doran, #6}{Doran, #6}. Moreover, AS is the primary driver behind the high morbidity and mortality of most cardiovascular diseases(Lumeng, et al.). Plaques can progress into vulnerable plaques, vulnerable lesions marked by a substantial necrotic lipid center, a thin fibrous covering, and the infiltration of macrophages(Libby, et al.). The rupture of such plaques may precipitate various acute cardiovascular events, including myocardial infarction, arrhythmia, and ischemic stroke(Nayor, et al.).AS significantly impacts human health. The development of AS is multifactorial, involving multiple cellular components (vascular smooth muscle cells, endothelial cells, platelets) and pathological processes (oxidative stress, inflammation, lipid metabolism dysregulation)(Lee, and Olefsky). Numerous hypotheses (lipid infiltration, response to injury, inflammation, and immunity) have been proposed; however, none alone provides a complete and systematic explanation for its pathogenesis. At the molecular level, research on AS primarily focuses on lipid-related indicators and biomarkers(Linton, et al.; Tyrrell, and Goldstein). However, these traditional biomarkers have limitations in early detection and risk stratification of AS. Consequently, identifying novel biomarkers that can better capture the complexity of AS pathogenesis remains a critical research priority.

Mitochondrial biogenesis and functional homeostasis in smooth muscle cells play a crucial role in AS development, and enhancing these processes can slow atherosclerotic plaque formation and vascular calcification(Xiaojiao Wang, et al.). It stands as a critical contributor to AS development and advancement(Orekhov, et al.; Salnikova, et al.). Research confirms immunometabolism is a key regulator in conditions such as obesity and type 2 diabetes, both of which are risk factors for AS development(Vuong, et al.). Mitochondria play a critical role in the signaling of innate immunity, oxidative stress, and cell death, and their dysfunction can lead to arterial wall cell injury, triggering inflammatory and immune responses in the body(Stocker, and Keaney). Mitochondrial and immune-related genes influence the initiation and progression of AS through various pathways. While their significance is undeniable, the precise mechanisms remain elusive. Thus, integrated analysis of mitochondrial and immune-related genes may identify novel biomarkers for AS diagnosis and prognosis, as well as potential therapeutic targets for intervention.

To elucidate these complex mechanisms at cellular resolution, advanced technologies like single-cell RNA sequencing (scRNA-seq) now serve as formidable instruments.. In AS research, single-cell has markedly enhanced our grasp of cellular diversity, particularly macrophage diversity, within atherosclerotic lesions(Cochain, et al.; Jian-Da Lin, et al.){Lin, #22}. The emergence of it marks a fundamental paradigm shift in life science research, transitioning from studying "population averages" to resolving "individual heterogeneity." This approach not only enables precise identification of all cellular subpopulations within a tissue, thereby refining our understanding of biological system diversity, but also reconstructs cellular trajectories across pseudotemporal dynamics to elucidate key regulatory mechanisms. Furthermore, it provides accurate identification of specific cell subsets required for experimental investigations, creating critical foundations for subsequent research. The integration of single-cell analysis with transcriptomic technologies significantly enhances the resolution of immune cell subpopulations and increases the potential for discovering novel cellular subtypes. In complex pathological contexts, single-cell approaches may struggle to detect disease-specific cellular signatures; combining them with transcriptomic profiling substantially improves the precise spatial localization of diseased cells.

This study leveraged publicly available transcriptomic and scRNA-seq data of AS. Through comprehensive bioinformatic analyses, including differential expression analysis, weighted gene co-expression network analysis, and machine learning approaches, we aimed to identify mitochondrial and immune-related biomarkers associated with AS. Subsequently, single-cell analysis was employed to pinpoint key cell populations, which were then used to delineate the specific expression patterns and distribution characteristics of these biomarkers within them. This work provides novel scientific insights and potential therapeutic targets for the early detection and treatment of AS.

2. Materials and Methods

2.1 Data acquisition

The GSE43292, GSE100927, and GSE159677 were sourced from the Gene Expression Omnibus (GEO) database. The GSE43292 The GSE43292 served as the training dataset, which included 32 tissue samples sourced from atheroma plaques of individuals suffering from hypertension (the AS group), matched by an equal number of carotid tissue samples obtained from hypertensive patients in the control group. 32 tissue samples from atheroma plaques of hypertensive patients (AS group) and an equivalent number of carotid tissue samples from hypertensive patients (control group). The GSE100927 functioned as a verification dataset, including carotid artery tissue samples from 29 AS patients and 12 normal controls. The scRNA-seq data set from GSE159677 comprised three calcified atherosclerotic core plaques (AS group) and three patient-matched proximal adjacent portions of carotid artery (control group), all obtained from patients undergoing carotid endarterectomy. Additionally, 1,136 mitochondria-related genes (MRGs) were originated from MitoCarta 3.0 (Table S1). The immune-related genes (IRGs) were originated from ImmPort Shared Data, with 1,793 IRGs remaining after the removal of duplicates (Table S2).

2.2 Differential expression analysis

To pinpoint genes showing significant variations in expression between the AS and control cohorts within the GSE43292 dataset, differentially expressed genes (DEGs) in both groups were identified by the limma package, with a threshold of $|\log_2$fold change (FC)$| > 0.5$ and adj.P < 0.05. Moreover, we visualized these DEGs through a volcano plot generated with the ggplot2 package and a heatmap produced using the ComplexHeatmap package.

2.3 Weighted gene co-expression network analysis (WGCNA)

Further intersections between MRGs and IRGs were identified to obtain genes related to both mitochondria and immunity, which were defined as MRG_IRG. Based on all samples in the GSE43292, MRG_IRG were scored utilizing the ssGSEA algorithm in the GSVA (v 1.46.0) package. To obtain the modular genes with the highest degree of association with MRG_IRG, WGCNA was conducted utilizing the WGCNA (v 1.71) package on all samples in the GSE43292. The expression matrix was subjected to filtration to remove genes with zero expression values. Afterwards, the samples underwent clustering analysis utilizing the Hclust function to construct a sample clustering tree, and the necessity of excluding outlier samples was evaluated. To optimize the scale-free topological compatibility in gene interaction relationships, A soft threshold of power was employed to build the co-expression network, chosen specifically because it yielded a scale-free fit index (signed R^2) that surpassed 0.80 while keeping the mean connectivity hovering near zero. The filtered expression matrix formed the basis for the WGCNA network, with a minimum of 50 genes per module. Subsequently, co-expression modules were identified, resulting in the generation of a hierarchical clustering tree. The MRG_IRG scores were utilized as observable characteristics, and the Pearson correlation analysis was performed to calculate the correlation matrix between MRG_IRG scores and co-expression modules ($|cor| > 0.30$, $P < 0.05$). The modules exhibiting the most robust positive and negative correlations with the MRG_IRG score were identified and designated as key modules. Subsequently, the key module genes were identified using a threshold of |module gene significance (GS)| > 0.8 and |module membership (MM)| > 0.8.

2.4 Identification and enrichment analysis of candidate genes

The intersection of up-regulated DEGs, down-regulated DEGs, and key module genes was implemented utilizing the ggvenn package to identify MRG_IRG-related genes in AS, which were designated as candidate genes. Then, the candidate genes' biological roles were clarified using the clusterProfiler package, which facilitated the performance of GO and KEGG enrichment analysis on the candidate genes. Besides, to investigate protein-level interactions among candidate genes, the STRING database was utilized to create a PPI network (interaction score = 0.40).

2.5 Discernment of biomarkers through PPI network, machine learning, and expression validation

Further filtering of the candidate genes was conducted. Initially, the candidate genes were filtered utilizing two distinct algorithms (degree and closeness) via the Cytohubba plug-in within the Cytoscape software. Subsequently, the respective top 10-ranked genes identified by the aforementioned two different algorithms were intersected utilizing the ggvenn package to obtain the hub genes. SVM-RFE analysis of hub genes was employed the Mlbench package, and the highest accuracy rate was employed as the criterion for identifying the feature genes through SVM-RFE. The optimal number of variables was decided by the root mean square error (RMSE), where a lower RMSE value indicated a higher level of predictive accuracy for the SVM-RFE model. Concurrently, a Boruta analysis was executed with the default parameters utilizing the Boruta package, with genes exceeding shadowMax being selected as Boruta feature genes. The ggvenn package was applied to identify the intersection of the feature genes obtained from the two machine learning methods to identify candidate biomarkers. Additionally, The Wilcoxon test was executed utilizing the rstatix software package, comparing the expression patterns of candidate biomarkers between AS and control groups in the GSE43292 and GSE100927. Following this, candidate biomarkers that exhibited intergroup significant differences ($P < 0.05$) and displayed consistent expression trends in both GSE43292 and GSE100927 were identified as biomarkers.

2.6 Establishment and assessment of an artificial neural network (ANN) model

To assess biomarker prognostic potential for AS, an ANN model was developed. The ANN model was constructed based on biomarkers utilizing the neuralnet package in the GSE43292. The data were initially subjected to a process of normalisation. Subsequently, the min-max method within the range [0, 1] was employed for the purpose of separating the zoom data before the commencement of neural network training. Before calculation, the data values were standardized by determining their maximum and minimum values, three hidden layers were designated for the architecture. Following this, a confusion matrix was plotted utilizing the GSE43292, and the weights of the biomarkers in the ANN model were predicted. Furthermore, the ROC curves were generated utilizing the pROC package in the GSE43292 and GSE100927. The predicted performance of the ANN model for AS was evaluated by calculating the AUC value, with an AUC greater than 0.7 indicating favorable predictive ability.

2.7 Gene-gene interaction (GGI) network construction and gene set enrichment analysis (GSEA)

To gain a deeper understanding of the interrelationships between biomarkers and functionally similar genes, as well as the common functions of these genes, the GeneMANIA database was used to build the GGI network. The purpose of the GSEA was the clarification of the roles of biomarkers throughout AS development. Initially, Spearman correlation coefficients were calculated utilizing the corrplot package for each biomarker and every other genes across all samples in the GSE43292, with the results then arranged in descending sequence. Afterwards, the GSEA was performed utilizing the clusterProfiler package (|normalized enrichment score (NES)| >1, adj.P < 0.05).

2.8 Correlation between biomarkers and inflammatory and immune factors

The development of AS was associated with the production of inflammatory factors, which highlighted the necessity for further investigation into the correlation between biomarkers and inflammatory factors. A total of 200 inflammation-related genes (IRGs) were identified from the molecular signatures database (MSigDB) by querying the "Hallmark_inflammatory_response" keyword. Subsequently, in the GSE43292, ssGSEA scores for IRGs were performed on AS and control samples, utilizing the GSVA package. The Wilcoxon test was employed to analyze the differences in IRGs scores between the AS and control groups. Afterwards, the Spearman correlation analysis was conducted to examine the relationship between biomarkers and IRGs scores (|cor| > 0.30, P < 0.05). The interaction between biomarkers and immune factors was subjected to further investigation. The data on 30 chemokines, 15 chemokine receptors, and 20 major histocompatibility complex (MHC) in the AS samples of the GSE43292 were downloaded utilizing the Tumor Immune System Interaction Database. Thereafter, the Spearman correlation assessed the relationship between biomarker expression and immunity, with a particular focus on chemokines, chemokine receptors, and MHC.

2.9 Immune microenvironment analysis

To gain further insight into immune infiltration levels, the abundance of 22 distinct immune cell types was calculated for each sample in the AS and control groups from the GSE43292, utilizing the CIBERSORT algorithm. The analysis included samples with a p-value of less than 0.05 and non-zero cell content. Next, the differential immune cells (DICs) between the two groups were compared via the Wilcoxon test (P < 0.05). Additionally, to evaluate the correlation between biomarkers and DICs, as well as the interconnections among DICs themselves, a Spearman correlation analysis examined data from the entire GSE43292 cohort.

2.10 Construction of regulatory networks

Molecular regulatory networks provided insights into the fundamental understanding of the core mechanisms behind gene regulation in disease processes. The prediction of miRNA regulation of biomarkers could be achieved through the miRDB database (target score > 80) and the TargetScan database (context++ score percentile ≥ 90). Next, overlapping miRNAs from these databases were analyzed to ascertain crucial regulatory elements. Through the miRNet database, key miRNAs' upstream lncRNAs were forecasted. Furthermore, through the Cytoscape iRegulon plugin (NES > 5) forecasted transcription factors engaging with biomarkers.Ultimately, the resultant lncRNAs-key miRNAs-mRNA and TF-mRNA networks were visualized in Cytoscape software.

2.11 Drug prediction and molecular docking

Biomarker-targeting medications were pinpointed via the DGIdb drug-gene interaction repository. Then, using Cytoscape, the drug-biomarker network map was then rendered. The drugs with the highest interaction scores, as predicted by the DGIdb, were identified as key drugs. Afterwards, the key drugs were added to the the public chemistry database to acquire their three-dimensional structures. Moreover, the biomarker protein targets' crystal and ligand molecular structures were sourced from pharmaceutical database. Subsequently, the biomarkers underwent molecular docking with the pivotal medication through the AutoDock Tool software. In the end, the outcomes of the molecular docking were brought to life using PyMOL software.

2.12 The scRNA-seq data processing

The 10x scRNA-seq data of GSE159677 were subjected to quality control (QC) utilizing the Seurat package. Initially, cells with fewer than 200 genes and genes covered in less than three cells were excluded. Concurrently, cells expressing fewer than 100 or more than 6000 genes, as well as those with over 5% mitochondrial genes, were filtered out. Afterwards, the LogNormalize function of the Seurat package was utilized for standardized data. After that, 2,000 hypervariable genes were extracted utilizing the FindVariableFeatures function. Subsequently, principal component analysis (PCA) was employed to downscale the scRNA-seq data. Next, the optimal principal components (PCs) count for dimensionality reduction was ascertained by creating an elbow plot via the ElbowPlot function. To identify highly significant PCs, the JackStraw and ScoreJackStraw methods were utilized to identify PCs enhanced by a higher number of genes with low P-values (P < 0.05). The FindNeighbors and FindClusters functions were employed for unsupervised clustering analysis at a resolution of 0.4 to obtain cell clusters. To visualise the cell clusters, the uniform manifold approximation and projection (UMAP) method was used. Cellular annotation of various cell clusters was conducted in accordance with the marker genes documented in the scholarship.

2.13 Identification of key cells

The distribution and expression levels of biomarkers across annotated cells were subjected to further analysis through the use of UMAP and violin plots, respectively. Moreover, the Wilcoxon test evaluated biomarker expression in annotated cells between the AS and control groups within GSE159677. Subsequently, the aforementioned cells that significant inter-group variances (P < 0.05) in all biomarkers were identified as key cells.

2.14 Cellular communication and pseudo-temporal trajectory analyses

The CellChat package was employed to infer interaction patterns among the annotated cellular entities, consequently elucidating the number of interactions and the interaction weight/strength between key cells and other cells in the GSE159677. Moreover, the analysis of ligand-receptor interactions within the context of the single-cell expression profile employed the CellChat software. Then, the number of ligand-receptor interactions per annotated cell was counted using the CellChat package. Furthermore, the differentiation trajectories of all key cells were analyzed utilizing the monocle package. The segmentation of key cells' trajectories was conducted based on the trajectory nodes, thus enabling an investigation of biomarker expression at each stage. Subsequently, the monocle package was employed to delineate the dynamic trend of expression of biomarkers throughout cell differentiation.

2.15 Reverse transcription-quantitative PCR (RT-qPCR)

RT-qPCR analysis was performed to further confirm and compare the expression levels of biomarkers between the AS group and the control group. In total, 11 clinical samples were collected from patients with a confirmed diagnosis of AS, along with 9 control samples. Total RNA was extracted from frozen tissue samples of AS cases and control subjects using the TRizol kit (Ambion, Cat. 15596-018CN, USA), with all extraction procedures meticulously following the manufacturer's guidelines. A 1 μL volume of the extracted RNA underwent concentration analysis using a NanoPhotometer N50. Following this, the RNA's purity and concentration levels were documented, enabling the determination of the necessary RNA quantity for later reverse transcription procedures. Then, following the manufacturer's protocol, we employed Hifair® ▯ 1st Strand cDNA Synthesis SuperMix for qPCR Kit (Yeasen Biotechnology, Shanghai, China) to convert the RNA into cDNA through reverse transcription. Next, The cDNA underwent a dilution of 5- to 20-fold using RNase/ARase-free ddH2O, followed by the addition of 3uL of the diluted cDNA, 5uL of 2x Universal Blue SYBR Green qPCR Master Mix, 1uL of forward primer (10 μM), and 1uL of reverse primer (10 μM). Moreover, 40 cycles (exclusive of pre-denaturation) of reactions were executed with the CFX Connect real-time quantitative PCR machine (BIO-RAD, XLFZ006). The specifics of the program are detailed in Table S3. The data on the primer sequences for the biomarkers can be found in Table S4. GAPDH served as the reference gene, and relative gene expression levels were determined by employing the $2^{-\triangle\triangle CT}$ method.

2.16 Statistical analysis

The statistical analyses were conducted utilizing the R language, and the significant group differences were evaluated via the Wilcoxon test. In RT-qPCR, the Ct values were compared utilizing paired and two-tailed t-tests, which were calculated with GraphPad Prism 5. The $P < 0.05$ was deemed to be statistically significant.

3. Results

3.1 Identification of 1,239 DEGs and 239 key module genes

The analysis of differential gene expression identified 1,239 DEGs between the AS and the control cohorts. Among these, 682 were up-regulated genes and 557 were down-regulated genes in the AS group (Fig. 1a). The volcano plot labeled the top 15 most significantly up-regulated and down-regulated genes (ranked by $|\log_2 FC|$ in descending order). Additionally, a heatmap depicted the expression profiles of the aforementioned genes (Fig. 1b). Furthermore, 13 MRG_IRGs were identified through the intersection of 1,136 MRGs with 1,793 IRGs (Fig. 1c). The AS group demonstrated markedly elevated MRG_IRG scores in comparison to the control group (Fig. 1d). Subsequently, the expression matrix of 18,851 genes was subjected to WGCNA. One outlier sample was excluded from the subsequent network construction process, with the remaining 63 samples included instead (Fig. 1e). A soft threshold of 8 was identified, along with a signed R^2 value exceeding 0.80 and a mean connectivity of approximately zero (Fig. 1f). Afterwards, the hierarchical clustering tree identified a total of six co-expression modules (Fig. 1g). Moreover, the analysis linking gene modules to MRG_IRG scores showed that the MEbrown module exhibited the strongest positive correlation, while the MEturquoise module demonstrated the highest negative correlation (Fig. 1h). Following the application of thresholds of $|GS| > 0.8$ and $|MM| > 0.8$, 49 genes were retained in the MEbrown module and 190 genes in the MEturquoise module, resulting in a total of 239 key module genes (Fig. i).

3.2 Identification of 165 candidate genes and exploration of their functions

Then, 682 up-regulated DEGs, 557 down-regulated DEGs, and 239 key module genes were subjected to intersection analysis, resulting in 165 candidate genes (Fig. 2a). Subsequently, an enrichment analysis was conducted to provide preliminary insights into the signaling pathways implicated by the candidate genes. The candidate genes were significantly enriched in 63 GO entries with 22 biological processes (BPs), 40 cellular components (CCs), and one molecular function (MF) (adj.$P < 0.05$) (Table S5). In particular, the top 10 candidate genes' BPs, such as “release of sequestered calcium ion into cytosol” (Fig. 2b, Table S5), suggest that candidate genes played crucial roles in the regulation of intracellular calcium ions. Furthermore, a KEGG enrichment analysis revealed the enrichment of 12 pathways, like “mTOR signaling pathway” and “Sphingolipid signaling pathway” (Fig. 2c, Table S6). The enrichment of these pathways suggests that the candidate genes could exert a significant impact on both physiological and pathological role in cell signaling, metabolic regulation and immune response. Besides, a PPI network was constructed, including 98 candidate genes and 139 interactive relationships (Fig. 2d). Most of these candidate genes (such as TYROBP and RAC2) exhibited interactions with multiple other candidate genes at the protein level.

3.3 Acquisition of five biomarkers: PLCβ4, RAC2, TYROBP, VAV1, and VCL

The construction of the PPI network incorporated 98 candidate genes, which were subsequently integrated into Degree and Closeness algorithms. The overlap of the top 10 genes from both aforementioned algorithms resulted in eight hub genes: DMD, FERMT3, PATJ, PLCβ4, RAC2, TYROBP, VAV1, and VCL (Fig. 3a-c). Next, these 8 hub genes were incorporated into the SVM-RFE algorithm. The SVM-RFE model exhibited optimal prediction accuracy when the number of corresponding hub genes was set to five (RMSE = 0.4248) (Fig. 3d). A total of five SVM-RFE-feature genes were identified: PLCβ4, RAC2, TYROBP, VAV1, and VCL. Moreover, the Boruta analysis identified seven genes that exhibited significantly higher scores than ShadowMax, thus establishing them as Boruta featured genes (Fig. 3e). The aforementioned genes were: DMD, FERMT3, PLCβ4, RAC2, TYROBP, VAV1, and VCL. Afterwards, five candidate

biomarkers (PLCβ4, RAC2, TYROBP, VAV1, and VCL) were determined by cross-referencing the characteristic genes obtained from the two aforementioned machine learning algorithms (Fig. 3f). Furthermore, the results revealed substantially increased levels of RAC2, TYROBP, and VAV1 expression ($P < 0.0001$). along with notably decreased levels of PLCβ4 and VCL ($P < 0.0001$) in the AS group within the GSE43292 (Fig. 3g). Interestingly, the differential expression of these candidate biomarkers between the inter-group comparisons in the GSE100927 was consistent with that observed in the GSE43292 ($P < 0.0001$) (Fig. 3h). The results demonstrated that PLCβ4, RAC2, TYROBP, VAV1, and VCL exhibited stability and reliability, suggesting their potential value in diagnosing AS and qualifying them as biomarkers.

3.4 ANN model demonstrated favourable performance in assessing the diagnosis of AS

The ANN model was developed to further evaluate the diagnostic value of biomarkers for AS. The results of the ANN model indicated that the entire training process was performed in 38,231 steps (Fig. 4a). The termination condition was satisfied when the error function reached a value of less than 0.01, with an error value of 8.216055. The findings demonstrated ANN model weights spanned from -13.80 to 3.70.In addition, the weight predictions were -13.84514 (PLCβ4), 3.63852 (RAC2), -0.01164 (TYROBP), -4.27724 (VAV1), and 2.29652 (VCL). Furthermore, the AUC values of the ANN model in the GSE43292 and GSE100927 were 0.8291 and 0.9943, respectively (Fig. 4b-c), demonstrating that the model's high accuracy and stability.

3.5 GGI network construction as well as GSEA for PLCβ4, RAC2, TYROBP, VAV1, and VCL

The GGI network, featuring 25 genes, was formed, with five biomarkers surrounded by the remaining 20 genes (Fig. 5a). Among them, VAV1 showed the highest number of predicted functions, such as platelet activation and phagocytosis. Among the top 10 significantly enriched pathways for five biomarkers ($|NES| > 1$, $adj.P < 0.05$), the “arrhythmogenic right ventricular cardiomyopathy arvc”, “dilated cardiomyopathy”, and “hypertrophic cardiomyopathy hcm” were collectively activated in PLCβ4 and VCL, while they were inhibited in RAC2, TYROBP, and VAV1 (Fig. 5b). Conversely, the pathways involving “hematopoietic cell lineage” and “lysosome” were observed to be activated in RAC2, TYROBP, and VAV1, while exhibiting inhibition in PLCβ4 and VCL. The findings demonstrated that the five biomarkers were instrumental in regulating the underlying mechanisms of cardiovascular disease.

3.6 Correlation of PLCβ4, RAC2, TYROBP, VAV1, and VCL with inflammatory and immune factors

In consideration of the inflammatory characteristics inherent to AS, ssGSEA scores were assigned to IRGs across all samples within the GSE43292. Notably, the IRG scores were significantly higher in the AS group ($P < 0.0001$) (Fig. 6a). Furthermore, the IRG scores exhibited a notable negative correlation with PLCβ4 and VCL, while showing a pronounced positive correlation with RAC2, TYROBP, and VAV1 (Fig. 6b). Furthermore, 29 chemokines, 15 chemokine receptors, and 11 MHC were identified in the GSE43292. The majority of chemokines, chemokine receptors, and MHC exhibited a notable inverse correlation with PLCβ4 and VCL, along with substantial direct links to TYROBP, VAV1, and RAC2 (Fig. 6c-e). The findings indicated that PLCβ4 and VCL might exert a suppressive influence on inflammatory and immune responses, whereas TYROBP, VAV1, and RAC2 might facilitate such processes.

3.7 Immune cell distribution and correlation analysis of PLCβ4, RAC2, TYROBP, VAV1, and VCL

The immune microenvironment between the AS and control groups was subjected to further investigation. The analysis included 62 samples from the GSE43292, which exhibited a p-value of less than 0.05 and a complete cell content. A stacked bar chart depicted immune cell infiltration levels across 22 cell types in AS and control groups (Fig. 7a). Notably, eosinophils and activated mast cells were not identified in all samples. Moreover, 10 DICs were identified between the AS and control groups, such as memory B cells (Fig. 7b). Naive B cells demonstrated the highest number of correlations with other differential immune cells, including notable positive associations with CD8 T cells, activated Natural Killer (NK) cells, and monocytes, as well as significant negative associations with memory B cells, M0 macrophages, and neutrophils (Fig. 7c). Especially, naive B cells, CD8 T cells, activated NK cells, and monocytes demonstrated significant positive correlations with PLCβ4 and VCL, and notable negative correlations with RAC2, TYROBP, and VAV1. In contrast, memory B cells, M0 macrophages, and neutrophils demonstrated notable positive correlations with RAC2, TYROBP, and VAV1, in addition to PLCβ4 and VCL (Fig. 7d).

3.8 PLCβ4, RAC2, TYROBP, VAV1, and VCL were regulated by multiple factors

Further analysis examined biomarker regulatory factors. The miRDB database predicted 111 miRNAs targeting five biomarkers, whereas the TargetScan database forecasted 466 miRNAs. In total, 47 key miRNAs were identified, such as hsa-miR-374c-3p. Subsequently, 226 lncRNAs were predicted to target the key miRNAs, like LINC02381. Accordingly, an lncRNA-miRNA-mRNA network containing four biomarkers (PLCβ4, RAC2, TYROBP, and VCL), 47 key miRNAs, and 226 lncRNAs was constructed (Fig. 8a), showing that PLCβ4, RAC2, TYROBP, and VCL were influenced by various factors. Moreover, the five biomarkers predicted eight transcription factors. (Fig. 8b). In the TF-mRNA network, KLF7 was observed to regulate PLCβ4, VAV1, and VCL; FOS was also found to regulate PLCβ4, VCL, and RAC2; and SPI1 regulated RAC2, TYROBP, and VAV1.

3.9 Drug prediction and molecular docking of RAC2 and VAV1

During the drug prediction process, three potential drugs were discovered for VAV1: phorbol 12-myristate 13-acetate, epoetin alfa, and tetradecanoylphorbol acetate. Additionally, two drugs were identified for RAC2: IDARUBICIN and doxorubicin hydrochloride. However, no drugs were predicted for PLCβ4, TYROBP, and VCL (Fig. 9a). The pair with the highest interaction score was that of RAC2 and IDARUBICIN, which indicated that IDARUBICIN was the key drug in this context (Table 1). The molecular docking of RAC2 and IDARUBICIN was detailed in Fig. 9b. The free binding energies for the interactions between RAC2 and IDARUBICIN were found to be -7.6 kcal/mol. A free binding energy less than -5 kcal/mol indicated that the key drug had a high binding affinity for the biomarkers. Consequently, the results demonstrated a robust binding affinity between RAC2 and IDARUBICIN. Notably, the hydrogen bonding facilitated the binding of RAC2 and IDARUBICIN.

3.10 Macrophages, endothelial cells, and vascular smooth muscle cells (VSMCs) identified as key cells

The scRNA-seq data were subjected to further analysis. After QC, the cell tally was refined to 32,045, whereas the gene tally stayed unchanged at 22,615.. The results of the process after QC were presented in Fig. 10a. Following standard data processing, a subset of 2,000 highly variable genes was identified (Fig. 10b). The PCA was performed, the single-cell samples from the AS and control groups were found to be distributed reasonably (Fig. 10c). Afterwards, the top 30 PCs were chosen to subsequent analysis (Fig. 10d). The P-values associated with each gene across the top 30 PCs were predominantly below 0.05 (Fig. 10e), indicating a heightened statistical significance for these PCs. Moreover, the utilization of UMAP cluster analysis led to the identification of 23 distinct cell clusters (Fig. 10f). The expression of important marker genes for each cell type was visualized through bubble plots (Fig. 10g-h). A total of six cell clusters were annotated by finding marker genes, including macrophages, endothelial cells, VSMCs, NK T cells, T lymphocytes, and B lymphocytes (Fig. 10i). The biomarkers were further analyzed for their expression in the annotated cell types. Detailed representations of the biomarkers' distribution and expression within the annotated cells are illustrated in Fig. 10j-k. Notably, the expression of all biomarkers was notable in macrophages, endothelial cells, and VSMCs between the AS and control groups (Fig. 10l). Therefore, macrophages, endothelial cells, and VSMCs were identified as key cells.

3.11 Cellular communication and pseudo-temporal trajectory analyses of key cells

Cellular communication analysis uncovered complex interactions among the six annotated cell types. Remarkably, three key cells (macrophages, endothelial cells, and VSMCs) exhibited a high frequency and intensity of interaction with one another (Fig. 11a-b). The heatmap illustrating ligand-receptor pair interactions revealed a higher level of interaction between macrophages, endothelial cells, and vascular smooth muscle cells (Fig. 11c). In particular, the probability of CCL5-ACKR1 (ligand-receptor) interaction was notably higher between NK T cells and endothelial cells (Fig. 11d). Endothelial cell differentiation along the cell trajectory was depicted in Figure 11e, which illustrated the progressive differentiation of cells from left to right over time, with the deepest blue shade corresponding to the cells that differentiated earliest in the process. Further analysis revealed that endothelial cells underwent differentiation in three distinct states (state 1, 2, and 3), with the state 1 representing their early phase of differentiation (Fig. 11f-g). Furthermore, PLCβ4 expression was initially unaltered and subsequently increased over time, eventually reaching a stable state (Fig. 11h), suggesting that PLCβ4 potential significance primarily in the middle stages (state 1-2) of endothelial cell development. Following an initial decline, the expression of RAC2 and TYROBP was demonstrated to be stable, suggesting their important roles in the early stages (state 1) of endothelial cell development. In the state 1, VCL expression declined gradually, stayed constant in the state 2, and rose slowly in the state 3, suggesting that VCL expression might play a significant role in both the early and late phases of endothelial cell development. The pseudo-temporal trajectory analysis of macrophages and VSMCs was detailed in Fig. 11i and Fig. 11j. Subsequently, a comprehensive account of the pseudo-temporal trajectory analysis of endothelial cells was provided.

3.12 Validation of PLCβ4, RAC2, TYROBP, VAV1, and VCL

Following the extraction of total RNA, the concentration assay revealed that all the samples fell comfortably within the standard range of RNA concentration (Table S7). The RT-qPCR outcomes demonstrated a considerable elevation in mRNA expression levels for RAC2, TYROBP, and VAV1 in the AS group ($P < 0.05$) (Fig. 12a-c), while PLCβ4 mRNA expression exhibited a notable decrease ($P < 0.01$) (Fig. 12d). The outcomes aligned with the forecasted expression of the aforementioned 4 biomarkers, as derived from the GSE43292 and GSE100927. Furthermore, VCL mRNA expression exhibited a downward tendency in the AS group; however, this variation did not attain statistical significance (Fig. 12e).

4. Discussion

AS is a chronic disease pathologically characterized by lipid deposition in the arterial walls, resulting in plaque formation. Mitochondria contribute substantially to AS development; when dysfunctional, they lead to alterations in various mitochondrial components, such as enzymes and proteins, triggering subsequent pathological changes. Enhancing mitochondrial biogenesis and functional homeostasis in smooth muscle cells during the initiation and development of AS helps slow plaque progression and vascular calcification. Furthermore, mitochondria are integral to signaling in innate immunity, oxidative stress, and cell death. Mitochondrial dysfunction can induce arterial wall cell injury and promote inflammatory and immune responses. Using bioinformatic approaches, we explored the potential relationship among AS, mitochondria, and immunity. By applying two machine learning algorithms and validating expression levels, we identified five biomarkers (PLCβ4, RAC2, TYROBP, VAV1, and VCL). Drug prediction analysis suggested five potential therapeutics targeting two of these biomarkers (VAV1 and RAC2). Single-cell analysis revealed six annotated cell types, and cell communication analysis highlighted three key cell types: macrophages, endothelial cells, and vascular smooth muscle cells (VSMCs). Pseudotime trajectory analysis of these key cells indicated that one or more subclusters within each type were specifically associated with AS diseased tissue.

This study integrated mitochondrial and immune-related pathways, identifying five context-dependent biomarkers—PLCβ4,RAC2, TYROBP, VAV1, and VCL—that play significant roles in the pathogenesis of AS. PLCβ4 encodes phospholipase C-β, which hydrolyzes phosphatidylinositol 4,5-bisphosphate (PIP_2) to generate the second messengers inositol trisphosphate (IP_3) and diacylglycerol (DAG), thereby regulating intracellular calcium mobilization and PKC activation (Berridge, and Irvine). Although the direct link between PLCβ4 and AS remains incompletely characterized, its involvement in lipid catabolism (Li, et al.) and its association with lipoprotein(a), cholesterol, and high-density lipoprotein (HDL) levels (Ying-Ju Lin, et al.) suggest that the PLCβ4-governed IP_3/DAG signaling axis may contribute to vascular smooth muscle cell proliferation, endothelial inflammation, and dysregulated lipid metabolism in atherosclerotic lesions. Both RAC2 and VAV1 operate within the Rho GTPase signaling network and are critically involved in immune cell function and oxidative stress. RAC2, predominantly expressed in hematopoietic cells, modulates neutrophil and macrophage activities, thereby participating in efferocytosis—a protective process in AS (Barrett; Kojima, et al.). Moreover, RAC2 directly regulates the assembly of NADPH oxidase, a primary source of reactive oxygen species that promotes plaque instability (Jian Zhang, et al.). Similarly, the guanine nucleotide exchange factor VAV1 contributes to foam cell formation via CD36-dependent uptake of oxidized low-density lipoprotein (oxLDL) (Rahaman, et al.) and enhances oxidative stress through Rac1 activation (Miletic, et al.). These two proteins represent potential therapeutic targets for modulating inflammatory and oxidative pathways in AS. TYROBP (also known as DAP12) is a transmembrane signaling adaptor protein primarily expressed in myeloid and NK cells and has been repeatedly implicated in cardiovascular pathologies. Evidence indicates that TYROBP expression is upregulated in advanced atherosclerotic plaques, where it regulates immune cell function during

late-stage disease (Liu, et al.). Its recurrent association with multiple cardiovascular conditions—including atrial fibrillation (Yan-Fei Zhang, et al.) and aortic valve calcification (Qiao, et al.)—highlights its potential therapeutic target and warrants further mechanistic investigation. VCL, a cytoskeletal protein essential for cell adhesion and migration, exhibits a dual role in AS. Phosphorylation at Serine-721 increases endothelial permeability, thereby accelerating disease progression (Shih, et al.), whereas its regulatory effect on smooth muscle cell migration may contribute to plaque stabilization (Wu, et al.). Our findings suggest that VCL may indirectly influence lipid metabolism through cytoskeleton-mediated signaling, although this hypothesis requires experimental validation. In summary, these five biomarkers collectively intersect with three core mechanisms underlying AS: mitochondrial dysfunction, immune dysregulation, and vascular remodeling. Their identification provides novel insights into disease mechanisms and opens avenues for potential therapeutic strategies.

Our immunoinfiltration analysis revealed ten types of immune cells with notable disparities in abundance between the AS and control cohorts. Among these, naive B cells, $CD8^+$ T cells, activated NK cells, and monocytes showed positive correlations with PLCβ4 and VCL but negative correlations with RAC2, TYROBP, and VAV1, while memory B cells, M0 macrophages, and neutrophils exhibited the opposite pattern. These correlations suggest distinct roles of these biomarkers in regulating different immune cell subsets during AS progression. B cells are primarily responsible for mediating humoral immunity. As a major population in lymphoid and hematopoietic organs, they hold critical roles in the development, activation, and regulation of cellular immune responses. Accumulating evidence indicates that B cell-mediated immunity, as a element of the immune reaction associated with AS, can directly or indirectly promote disease progression. For instance, a study by Hamze et al. showed the existence of B cells within the vascular walls of human atherosclerotic plaques (Hamze, et al.). Furthermore, immunodeficient mice exhibited increased susceptibility to AS upon immune reconstitution (George, et al.). In line with this, B cell depletion therapy using Rituximab was associated with a significant reduction in total carotid intima-media thickness in patients(Edwards, et al.; Novikova, et al.) {Edwards, #38}{Edwards, #38}. Based on these findings, we propose that B cell immunity plays a particularly important role in the pathogenesis of AS(Tsiantoulas, et al.; Srikakulapu, and McNamara). $CD8^+$ T cells play pivotal parts in both innate and adaptive immune defense mechanisms (Zhang, and Bevan). However, they can also contribute to excessive immune responses, leading to pathological immune-mediated damage. In the context of AS, $CD8^+$ T cells accumulate within the vascular wall(Jian-Da Lin, et al.) and are predominantly localized in the plaque shoulder regions and fibrous caps .The majority of studies indicate that the activation and cytokine secretion by $CD8^+$ T cells exert pro-atherogenic effects. In contrast, other research suggests a protective role, proposing that $CD8^+$ T cells may help reduce lesion formation by eliminating immunogenic dendritic cells through their cytotoxic activity, thereby attenuating the atherogenic process (Paul, et al.). NK cells function as promoters of AS. Their prevalence is significantly higher in vulnerable human plaques compared to stable plaques. Moreover, glycosphingolipids, which act as potent agonists for NK cells, have been detected within human atherosclerotic lesions but are absent in normal arteries(Bobryshev, and Lord). These findings collectively indicate an important and active role for NK cells in the pathogenesis of AS. Monocyte-derived macrophages are essential in lipid accumulation and atherosclerosis advancement(Li, and Glass). A study involving patients with hypercholesterolemia demonstrated that statin treatment suppresses the expression of pro-inflammatory cytokines—including TNF and IL-1β—in monocytes. This anti-inflammatory impact on monocytes plays a pivotal role in lessening adverse cardiovascular incidents among those suffering from stable coronary artery disease(Ferro, et al.) .

Based on single-cell findings, we identified that three key cell types—macrophages, endothelial cells, and VSMCs—exhibit frequent and high-intensity interactions with each other. A study on the pathogenesis of atherosclerotic cardiovascular disease indicated that macrophages help sustain local inflammatory responses, promote plaque development and thrombus formation. More specifically, classically activated M1 macrophages initiate and maintain inflammation in atherosclerotic cardiovascular disease, while alternatively activated M2 macrophages contribute to inflammation resolution(Barrett). Therefore, we consider that macrophages significantly influence the development of AS. Recent studies have further demonstrated that macrophages engage throughout the full course of AS, including the initiation, growth, eventual rupture, and wound-healing stages of arterial plaques. These actions further alter the plaque microenvironment, which not only impacts ongoing and impending inflammatory responses within the lesion but also directly determines whether AS progresses or regresses (Hou, et al.). Regarding endothelial cell dysfunction and AS, one study proposed—through analysis of molecular mechanisms and typical clinical risk factors—a hypothesis that a drug could be developed to act on endothelial cells in AS-prone regions of the arterial vasculature, reprogramming them to express a vasoprotective phenotype. This might mitigate systemic risk impacts and slow the progression of atherosclerotic lesions (Gimbrone, and García-Cardeña). This hypothesis may provide valuable direction for future AS research, particularly regarding the development of targeted endothelial therapies. Additionally, insulin-mediated activation of endothelial cells enhances the differentiation of perivascular progenitor cells into brown adipocytes and their associated adipokines, thereby ameliorating AS (Park, et al.). We therefore propose that endothelial dysfunction is a major cause of energy imbalance in patients with obesity, and since obesity is a high-risk factor for AS, the relationship between endothelial cells and AS is undoubtedly close. Abnormal proliferation of VSMCs promotes plaque formation; however, in advanced plaques, VSMCs can play important beneficial roles, such as preventing fibrous cap rupture. Phenotypic switching of VSMCs leads to a dedifferentiated form lacking typical VSMC markers, and this transition directly promotes AS(Bennett, et al.) . Traditionally, foam cells were considered to originate from bone marrow-derived macrophages within arterial wall lipid deposits, which take up low-density lipoprotein, exhibiting a foamy appearance. Although macrophages indeed contribute to foam cells in atherosclerotic plaques, studies using both mouse and human medial and intimal cultures under atherosclerotic conditions have shown that VSMCs may convert into foam cells upon encountering aggregated or oxidized LDL. In fact, most foam cells in human atherosclerotic plaques appear to originate from VSMCs (Allahverdian, et al.; Ying Wang, et al.). In summary, these three key cell types influence the initiation and progression of AS through various pathways, and we believe these insights provide a new perspective for AS research.

The RT-qPCR results showed that the expression levels of PLCβ4, TYROBP, VAV1, and RAC2 were consistent with database predictions. This concordance between RT-qPCR results and bioinformatic predictions validates the reliability of our computational analysis and strengthens the clinical relevance of these biomarkers. In contrast, the results for VCL did not align with expectations and were not statistically significant. We speculate that this discrepancy may be attributed to various factors, such as sample source and processing methods. It is also possible that the relatively small sample size led to insufficient statistical power, or that sample heterogeneity masked any genuine expression differences of VCL. While the successful validation of four biomarkers provides strong evidence for their roles in AS pathogenesis, the inconsistent results for VCL highlight the need for further investigation with larger cohorts to

fully elucidate its contribution to disease progression. To improve the accuracy and general applicability of the results, it is recommended that future studies include larger sample sizes and adhere to standardized protocols for replication.

This study identified potential mitochondrial and immune-related biomarkers (PLCβ4, RAC2, TYROBP, VAV1, and VCL) in patients with AS through bioinformatics analysis, suggesting their promising applications in early AS identification and therapeutic strategy development. RT-qPCR validation successfully confirmed the differential expression of four biomarkers (PLCβ4, TYROBP, VAV1, and RAC2) in clinical samples, demonstrating strong concordance with bioinformatic predictions. Although VCL showed expression patterns consistent with predictions but without statistical significance, all five biomarkers warrant continued investigation, given their central roles in atherosclerotic pathways identified through multi-omics analysis. In addition, single-cell analysis identified three key cell types—Macrophages, Endothelial cells, and VSMCs—closely associated with AS progression. These findings provide new candidate genes for AS diagnosis, provide innovative understanding of pathogenesis and progression of the disease, and deliver valuable information for the development of future treatment strategies. Building upon these promising findings, future validation studies should employ larger clinical cohorts and complementary experimental approaches, including animal models, to further validate causality of these biomarkers in AS development and assess their therapeutic potential. We will continue to closely monitor research progress related to these five biomarkers and three key cell types in the field of AS to gain deeper mechanistic insights.

5. Conclusion

This study integrated bulk and scRNA-seq analyses to identify five candidate biomarkers (PLCβ4, TYROBP, VAV1, RAC2, VCL), of which four were experimentally validated (PLCβ4, TYROBP, VAV1, RAC2), along with three key cell types (Macrophages, endothelial cells, and VSMCs), providing novel insights into the molecular mechanisms and potential therapeutic targets for AS. Drug prediction identified five potential compounds targeting AS-related biomarkers. Although this study has achieved significant progress, further clinical trials are required to validate these findings and translate them into practical clinical applications.

Declarations

Author Contributions

ZCX: conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing; KWZ and DBZ: collection and assembly of data, data analysis and interpretation, manuscript writing; JHZ and WLW: collection of data, data analysis and interpretation; YFQ: conception and design, financial support, administrative support, manuscript writing, final approval of manuscript. All the authors have read and approved the final content of this manuscript.

Conflict of Interest Statement

All methods were performed in accordance with the relevant guidelines and regulations.Written informed consent was obtained from all subjects and/or their legal guardian(s) prior to their participation in this study.

Ethics Approval and Consent to Participate

The study was approved by Ethics Committee of Henan Provincial People's Hospital, Henan Provincial People's Hospital.

Funding

This work was supported by grants to Ya-fei Qin from Natural Science Foundation of Henan Province (No. 242300420412)

Acknowledgements

Not applicable

Availability of Data and Materials

The datasets analysed during the current study are publicly available from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The accession numbers are GSE43292, GSE100927, and GSE159677. Mitochondria-related genes were obtained from MitoCarta 3.0 (https://www.broadinstitute.org/mitocarta), and immune-related genes were obtained from ImmPort (https://www.immport.org). All data generated or analysed during this study are included in this published article and its supplementary information files.

References

1. Doran, Amanda C. "Inflammation Resolution: Implications for Atherosclerosis." *Circulation research*, vol. 130, no. 1, pp. 130–48, PubMed, doi:10.1161/CIRCRESAHA.121.319822.

2. Lumeng, Carey N. et al. "Obesity Induces a Phenotypic Switch in Adipose Tissue Macrophage Polar Ization." *The Journal of clinical investigation*, vol. 117, no. 1, pp. 175–84, PubMed, doi:10.1172/JCI29881.
3. Libby, Peter et al. "Atherosclerosis." *Nature reviews. Disease primers*, vol. 5, no. 1, p. 56, PubMed, doi:10.1038/s41572-019-0106-z.
4. Nayor, Matthew et al. "The Molecular Basis of Predicting Atherosclerotic Cardiovascular Disea Se Risk." *Circulation research*, vol. 128, no. 2, pp. 287–303, PubMed, doi:10.1161/CIRCRESAHA.120.315890.
5. Lee, Yun Sok and Jerrold Olefsky. "Chronic Tissue Inflammation and Metabolic Disease." *Genes & development*, vol. 35, no. 5-6, pp. 307–28, PubMed, doi:10.1101/gad.346312.120.
6. Linton, MacRae F. et al. "Hdl Function and Atherosclerosis: Reactive Dicarbonyls as Promising Ta Rgets of Therapy." *Circulation research*, vol. 132, no. 11, pp. 1521–45, PubMed, doi:10.1161/CIRCRESAHA.123.321563.
7. Tyrrell, Daniel J. and Daniel R. Goldstein. "Ageing and Atherosclerosis: Vascular Intrinsic and Extrinsic Factors a Nd Potential Role of Il-6." *Nature reviews. Cardiology*, vol. 18, no. 1, pp. 58–68, PubMed, doi:10.1038/s41569-020-0431-7.
8. Wang, Xiaojiao et al. "Molecular Dysfunctions of Mitochondria-Associated Endoplasmic Reticulu M Contacts in Atherosclerosis." *Oxidative medicine and cellular longevity*, vol. 2021, p. 2424509, PubMed, doi:10.1155/2021/2424509.
9. Orekhov, Alexander N. et al. "Mitochondrion as a Selective Target for the Treatment of Atheroscleros Is: Role of Mitochondrial DNA Mutations and Defective Mitophagy in the Pathogenesis of Atherosclerosis and Chronic Inflammation." *Current neuropharmacology*, vol. 18, no. 11, pp. 1064–75, PubMed, doi:10.2174/1570159X17666191118125018.
10. Salnikova, Diana et al. "Mitochondrial Dysfunction in Vascular Wall Cells and Its Role in Ather Osclerosis." *International journal of molecular sciences*, vol. 22, no. 16, p. 8990, PubMed, doi:10.3390/ijms22168990.
11. Vuong, Jacqueline T. et al. "Immune Checkpoint Therapies and Atherosclerosis: Mechanisms and Clinic Al implications: Jacc State-of-the-Art Review." *Journal of the American College of Cardiology*, vol. 79, no. 6, pp. 577–93, PubMed, doi:10.1016/j.jacc.2021.11.048.
12. Stocker, Roland and John F. Keaney, Jr. "Role of Oxidative Modifications in Atherosclerosis." *Physiological reviews*, vol. 84, no. 4, pp. 1381–478, PubMed, doi:10.1152/physrev.00047.2003.
13. Cochain, Clément et al. "Single-Cell Rna-Seq Reveals the Transcriptional Landscape and Heteroge Neity of Aortic Macrophages in Murine Atherosclerosis." *Circulation research*, vol. 122, no. 12, pp. 1661–74, PubMed, doi:10.1161/CIRCRESAHA.117.312509.
14. Lin, Jian-Da et al. "Single-Cell Analysis of Fate-Mapped Macrophages Reveals Heterogeneity, Including Stem-Like Properties, during Atherosclerosis Progression an D Regression." *JCI insight*, vol. 4, no. 4, p. e124574, PubMed, doi:10.1172/jci.insight.124574.
15. Berridge, M. J. and R. F. Irvine. "Inositol Trisphosphate, a Novel Second Messenger in Cellular Signal Tr Ansduction." *Nature*, vol. 312, no. 5992, pp. 315–21, PubMed, doi:10.1038/312315a0.
16. Li, Chien-Feng et al. "Plcb4 Copy Gain and Plcß4 Overexpression in Primary Gastrointestinal S Tromal Tumors: Integrative Characterization of a Lipid-Catabolizing En Zyme Associated with Worse Disease-Free Survival." *Oncotarget*, vol. 8, no. 12, pp. 19997–20010, PubMed, doi:10.18632/oncotarget.15306.
17. Lin, Ying-Ju et al. "Genetic Variants in Plcb4/Plcb1 as Susceptibility Loci for Coronary Ar Tery Aneurysm Formation in Kawasaki Disease in Han Chinese in Taiwan." *Scientific reports*, vol. 5, p. 14762, PubMed, doi:10.1038/srep14762.
18. Barrett, Tessa J. "Macrophages in Atherosclerosis Regression." *Arteriosclerosis, thrombosis, and vascular biology*, vol. 40, no. 1, pp. 20–33, PubMed, doi:10.1161/ATVBAHA.119.312802.
19. Kojima, Yoko et al. "Cd47-Blocking Antibodies Restore Phagocytosis and Prevent Atherosclero Sis." *Nature*, vol. 536, no. 7614, pp. 86–90, PubMed, doi:10.1038/nature18935.
20. Zhang, Jian et al. "Macrophage Aldh2 (Aldehyde Dehydrogenase 2) Stabilizing Rac2 Is Requir Ed for Efferocytosis Internalization and Reduction of Atherosclerosis Development." *Arteriosclerosis, thrombosis, and vascular biology*, vol. 42, no. 6, pp. 700–16, PubMed, doi:10.1161/ATVBAHA.121.317204.
21. Rahaman, Shaik O. et al. "Vav Guanine Nucleotide Exchange Factors Regulate Atherosclerotic Lesio N Development in Mice." *Arteriosclerosis, thrombosis, and vascular biology*, vol. 33, no. 9, pp. 2053–7, PubMed, doi:10.1161/ATVBAHA.113.301414.
22. Miletic, Ana V. et al. "Vav Proteins Control Myd88-Dependent Oxidative Burst." *Blood*, vol. 109, no. 8, pp. 3360–8, PubMed, doi:10.1182/blood-2006-07-033662.
23. Liu, Chiyu et al. "Identifying Rbm47, Hck, Cd53, Tyrobp, and Havcr2 as Hub Genes in Advan Ced Atherosclerotic Plaques by Network-Based Analysis and Validation." *Frontiers in genetics*, vol. 11, p. 602908, PubMed, doi:10.3389/fgene.2020.602908.
24. Zhang, Yan-Fei et al. "Cxcr4 and Tyrobp Mediate the Development of Atrial Fibrillation Via in Flammation." *Journal of cellular and molecular medicine*, vol. 26, no. 12, pp. 3557–67, PubMed, doi:10.1111/jcmm.17405.
25. Qiao, En et al. "Exploring Potential Genes and Pathways Related to Calcific Aortic Valv E Disease." *Gene*, vol. 808, p. 145987, PubMed, doi:10.1016/j.gene.2021.145987.
26. Shih, Yu-Tsung et al. "Vinculin Phosphorylation Impairs Vascular Endothelial Junctions Promot Ing Atherosclerosis." *European heart journal*, vol. 44, no. 4, pp. 304–18, PubMed, doi:10.1093/eurheartj/ehac647.
27. Wu, Chong et al. "Identification of a Hub Gene Vcl for Atherosclerotic Plaques and Disco Very of Potential Therapeutic Targets by Molecular Docking." *BMC medical genomics*, vol. 17, no. 1, p. 42, PubMed, doi:10.1186/s12920-024-01815-9.
28. Hamze, Moustafa et al. "Characterization of Resident B Cells of Vascular Walls in Human Athero Sclerotic Patients." *Journal of immunology (Baltimore, Md. : 1950)*, vol. 191, no. 6, pp. 3006–16, PubMed, doi:10.4049/jimmunol.1202870.

29. George, J. et al. "Cellular and Humoral Immune Responses to Heat Shock Protein 65 Are Bot H Involved in Promoting Fatty-Streak Formation in Ldl-Receptor Deficie Nt Mice." *Journal of the American College of Cardiology*, vol. 38, no. 3, pp. 900–5, PubMed, doi:10.1016/s0735-1097(01)01440-1.
30. Edwards, Jonathan C. W. et al. "Efficacy of B-Cell-Targeted Therapy with Rituximab in Patients with Rh Eumatoid Arthritis." *The New England journal of medicine*, vol. 350, no. 25, pp. 2572–81, PubMed, doi:10.1056/NEJMoa032534.
31. Novikova, Diana S. et al. "The Effects of Rituximab on Lipids, Arterial Stiffness and Carotid Int Ima-Media Thickness in Rheumatoid Arthritis." *Journal of Korean medical science*, vol. 31, no. 2, pp. 202–7, PubMed, doi:10.3346/jkms.2016.31.2.202.
32. Tsiantoulas, Dimitrios et al. "Targeting B Cells in Atherosclerosis: Closing the Gap from Bench to Be Dside." *Arteriosclerosis, thrombosis, and vascular biology*, vol. 35, no. 2, pp. 296–302, PubMed, doi:10.1161/ATVBAHA.114.303569.
33. Srikakulapu, Prasad and Coleen A. McNamara. "B Cells and Atherosclerosis." *American journal of physiology. Heart and circulatory physiology*, vol. 312, no. 5, pp. H1060–H67, PubMed, doi:10.1152/ajpheart.00859.2016.
34. Zhang, Nu and Michael J. Bevan. "Cd8(+) T Cells: Foot Soldiers of the Immune System." *Immunity*, vol. 35, no. 2, pp. 161–8, PubMed, doi:10.1016/j.immuni.2011.07.010.
35. Paul, Veronica Soundra Veena et al. "Quantification of Various Inflammatory Cells in Advanced Atherosclerot Ic Plaques." *Journal of clinical and diagnostic research : JCDR*, vol. 10, no. 5, pp. EC35–8, PubMed, doi:10.7860/JCDR/2016/19354.7879.
36. Bobryshev, Yuri V. and Reginald S. A. Lord. "Co-Accumulation of Dendritic Cells and Natural Killer T Cells within R Upture-Prone Regions in Human Atherosclerotic Plaques." *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*, vol. 53, no. 6, pp. 781–5, PubMed, doi:10.1369/jhc.4B6570.2005.
37. Li, Andrew C. and Christopher K. Glass. "The Macrophage Foam Cell as a Target for Therapeutic Intervention." *Nature medicine*, vol. 8, no. 11, pp. 1235–42, PubMed, doi:10.1038/nm1102-1235.
38. Ferro, D. et al. "Simvastatin Inhibits the Monocyte Expression of Proinflammatory Cytoki Nes in Patients with Hypercholesterolemia." *Journal of the American College of Cardiology*, vol. 36, no. 2, pp. 427–31, PubMed, doi:10.1016/s0735-1097(00)00771-3.
39. Hou, Pengbo et al. "Macrophage Polarization and Metabolism in Atherosclerosis." *Cell death & disease*, vol. 14, no. 10, p. 691, PubMed, doi:10.1038/s41419-023-06206-z.
40. Gimbrone, Michael A., Jr. and Guillermo García-Cardeña. "Endothelial Cell Dysfunction and the Pathobiology of Atherosclerosis." *Circulation research*, vol. 118, no. 4, pp. 620–36, PubMed, doi:10.1161/CIRCRESAHA.115.306301.
41. Park, Kyoungmin et al. "Endothelial Cells Induced Progenitors into Brown Fat to Reduce Atheros Clerosis." *Circulation research*, vol. 131, no. 2, pp. 168–83, PubMed, doi:10.1161/CIRCRESAHA.121.319582.
42. Bennett, Martin R. et al. "Vascular Smooth Muscle Cells in Atherosclerosis." *Circulation research*, vol. 118, no. 4, pp. 692–702, PubMed, doi:10.1161/CIRCRESAHA.115.306361.
43. Allahverdian, Sima et al. "Contribution of Intimal Smooth Muscle Cells to Cholesterol Accumulatio N and Macrophage-Like Cells in Human Atherosclerosis." *Circulation*, vol. 129, no. 15, pp. 1551–9, PubMed, doi:10.1161/CIRCULATIONAHA.113.005015.
44. Wang, Ying et al. "Smooth Muscle Cells Contribute the Majority of Foam Cells in Apoe (Apo Lipoprotein E)-Deficient Mouse Atherosclerosis." *Arteriosclerosis, thrombosis, and vascular biology*, vol. 39, no. 5, pp. 876–87, PubMed, doi:10.1161/ATVBAHA.119.312434.

Tables

Table 1 Biomarkers for Predicting Drug Response

Gene	Drug	Regulatory approval	Indication
RAC2	IDARUBICIN	Approved	Antineoplastic agents
RAC2	DOXORUBICIN HYDROCHLORIDE	Approved	Antineoplastic agents
VAV1	PHORBOL 12-MYRISTATE 13-ACETATE	Not Approved	
VAV1	EPOETIN ALFA	Approved	Antianemic agents, for treatment of anemia, for treatment of stroke, erythropoieti
VAV1	TETRADECANOYLPHORBOL ACETATE	Not Approved	

Figures

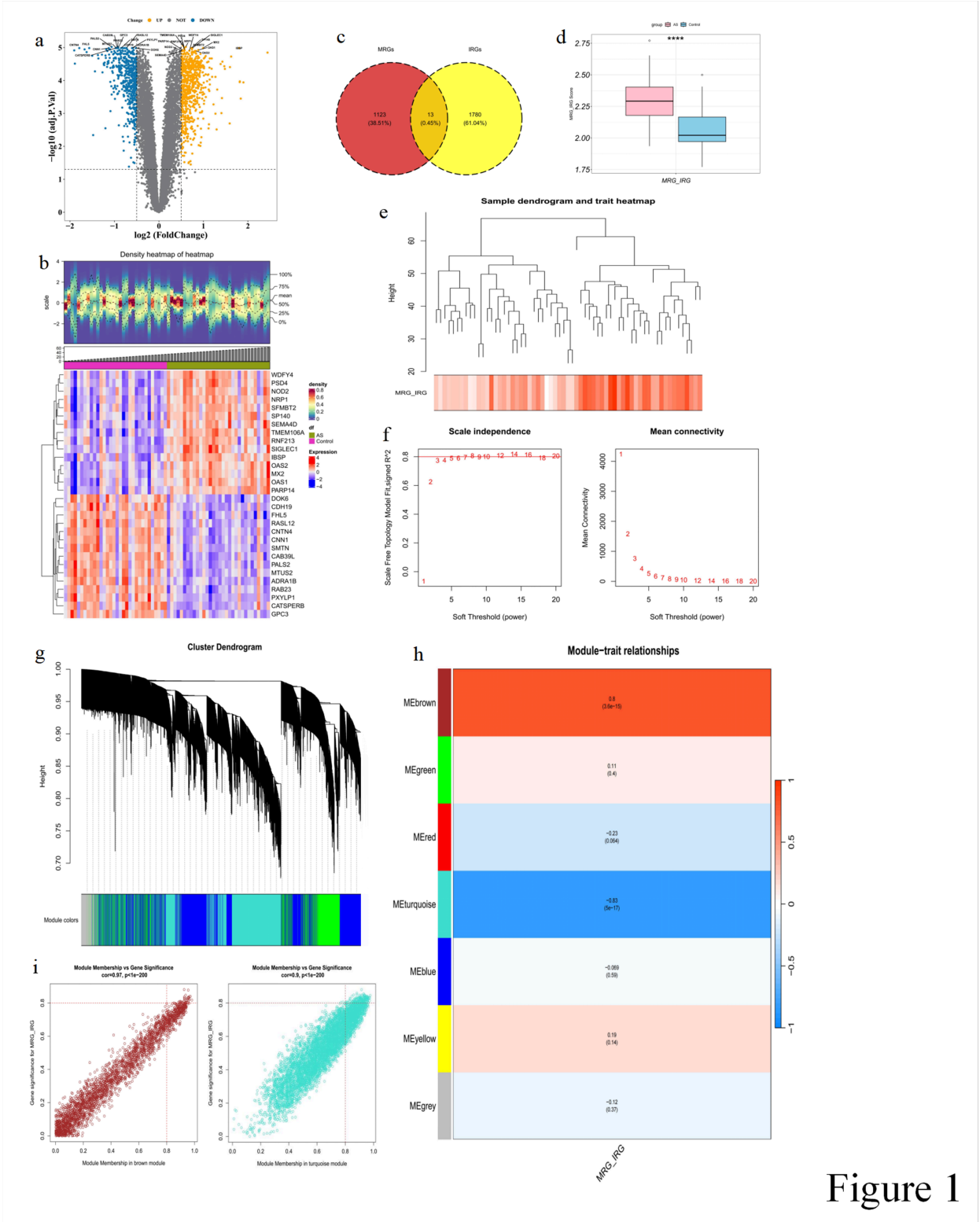
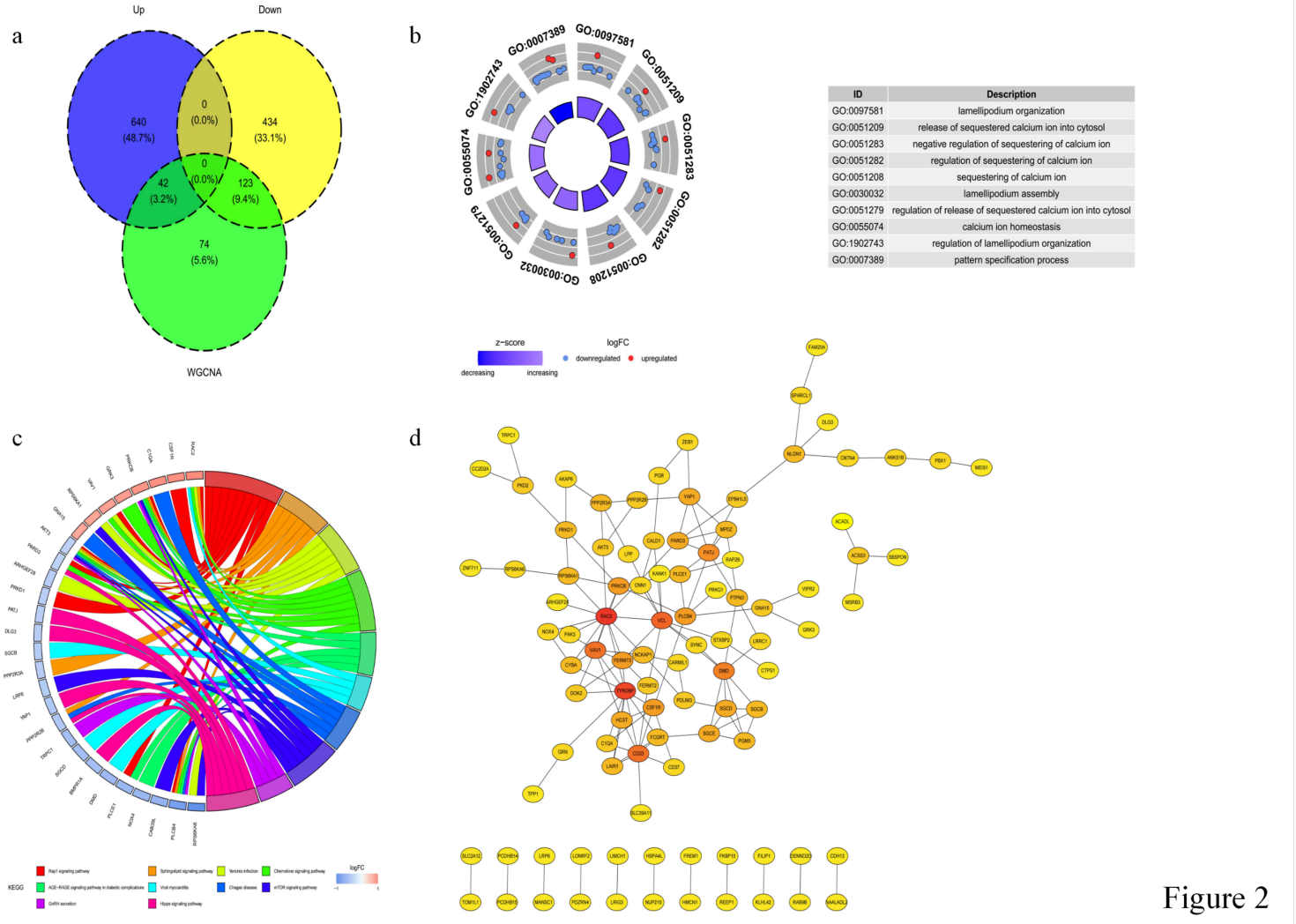


Figure 1

Figure 1

Identified DEGs and key module genes. (a) Volcano plot of DEGs: downregulated genes are shown in blue, upregulated genes in orange, and non-significant genes in gray. (b) Heatmap of DEGs: red indicates high expression, and blue indicates low expression. (c) Overlap of DEGs and key genes to obtain MRG_IRGs. (d) Between-group differences in MRG_PCD scores. (e) Sample clustering tree from WGCNA: each vertical line represents one sample, and the vertical axis indicates sample distance, with smaller values reflecting higher similarity. (f) Selection of the soft-thresholding power. (g) Hierarchical clustering dendrogram of genes: different colors represent different modules. (h) Correlation analysis between gene modules and MRG_IRG scores: red indicates positive correlation, blue indicates negative correlation; the numerical value represents the correlation coefficient, with larger absolute values indicating stronger correlation. P-values are shown in parentheses, with smaller values indicating greater significance. (i) Scatter plot illustrating the relationship between GS and MM in the key module.



ID	Description
GO:0097581	lamellipodium organization
GO:0051209	release of sequestered calcium ion into cytosol
GO:0051283	negative regulation of sequestering of calcium ion
GO:0051282	regulation of sequestering of calcium ion
GO:0051208	sequestering of calcium ion
GO:0030032	lamellipodium assembly
GO:0051279	regulation of release of sequestered calcium ion into cytosol
GO:0055074	calcium ion homeostasis
GO:1902743	regulation of lamellipodium organization
GO:0007389	pattern specification process

Figure 2

Figure 2

Candidate Genes and Their Functions. (a) Candidate genes were identified by intersecting differentially expressed MRGs, IRGs, and key module genes. (b) GO annotation results. Each dot represents a gene, with red indicating upregulation and blue indicating downregulation. The table provides specific descriptions of the GO term IDs. (c) KEGG enrichment analysis visualized using a Circos plot. Different colors represent distinct pathways, and the chords point to the genes enriched in each pathway. (d) Construction of a PPI (Protein-Protein Interaction) network. Different colors represent distinct functional pathways, and the chords point to the genes enriched in each pathway.

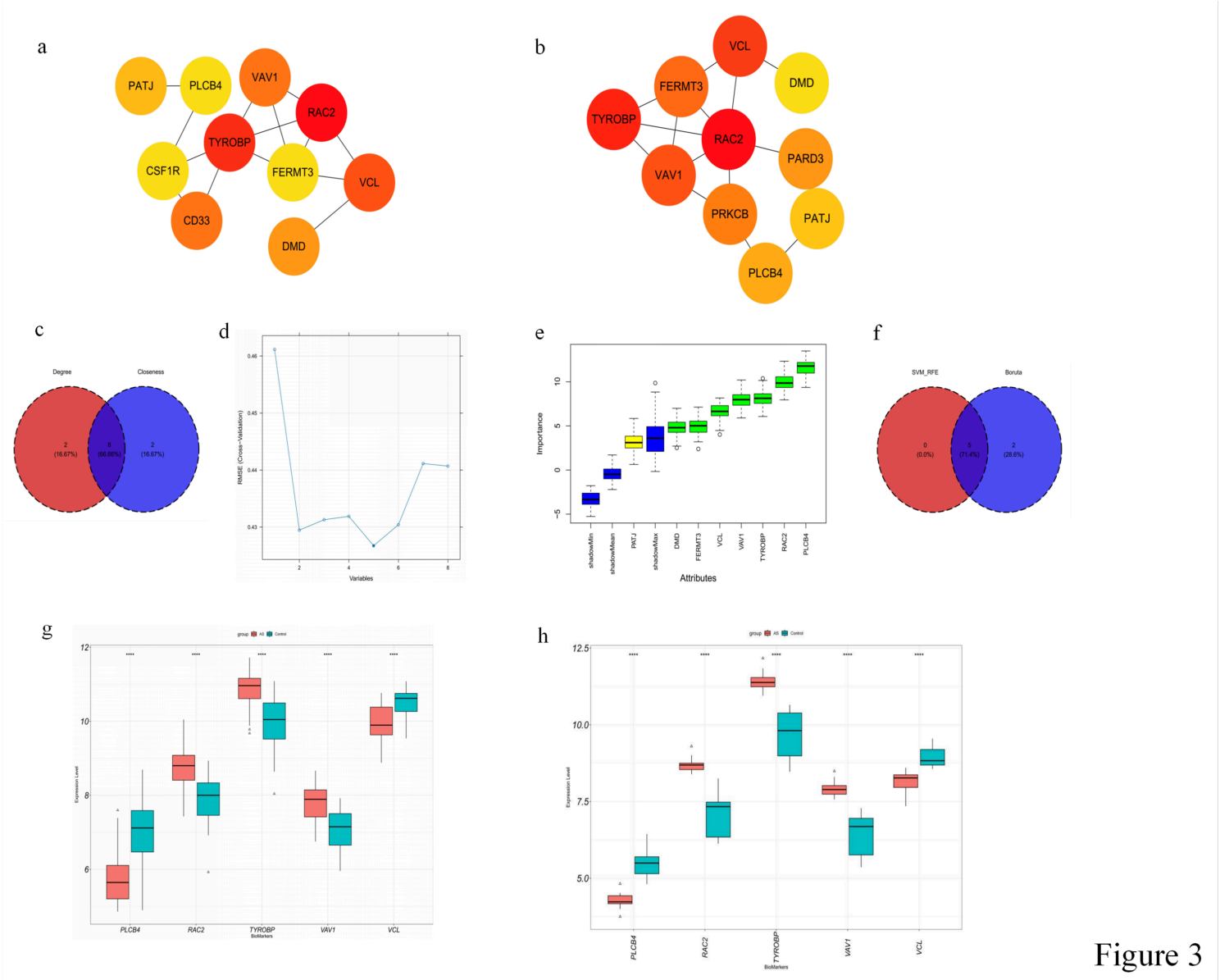


Figure 3

Figure 3

The PPI network was analyzed using two algorithms from Cytohubba in Cytoscape, yielding the TOP10 Hub genes identified by Closeness (a) and Degree (b), respectively. Genes highlighted in red indicate higher interaction scores within the algorithm, while those in yellow represent relatively lower scores. (c) The intersection of the Top 10 Hub genes from the two algorithms resulted in eight initially screened biomarkers. (d) SVM-RFE classification results: The eight initially screened biomarkers were further analyzed using the SVM-RFE method, leading to the identification of five candidate biomarkers. (e) Boruta results plot: Using the SVM-RFE method on the eight initially screened biomarkers, the plot shows the maximum Z-score among shadow attributes in yellow, revealing that seven genes scored significantly higher than the MZSA threshold. (f) Hub genes obtained from the two machine learning methods were intersected, resulting in five genes identified as candidate biomarkers. (g) Expression levels of the candidate biomarkers were validated in the training set. (h) Expression levels of the candidate biomarkers were validated in the validation set.

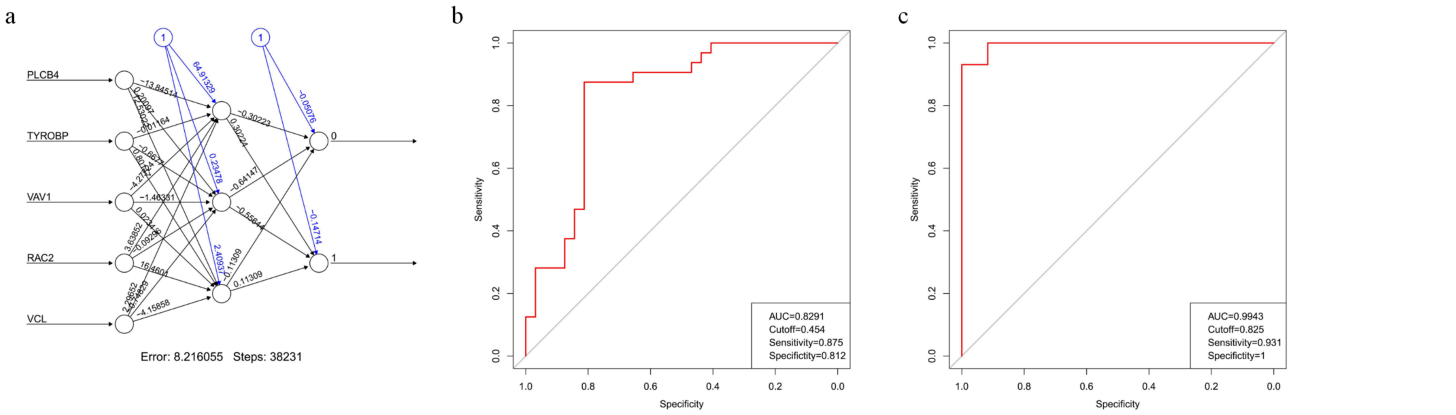


Figure 4

Figure 4

Verification of Biomarkers Using Expression Data. (a)Architecture of the artificial neural network composed of the five biomarkers. Numbers between nodes represent connection weights. The network consists of an input layer (five nodes, left), a hidden layer (three nodes, middle), and an output layer (two nodes, right). (b) ROC curve evaluating the performance of the ANN on the training set. (c) ROC curve evaluating the performance of the ANN on the validation set.

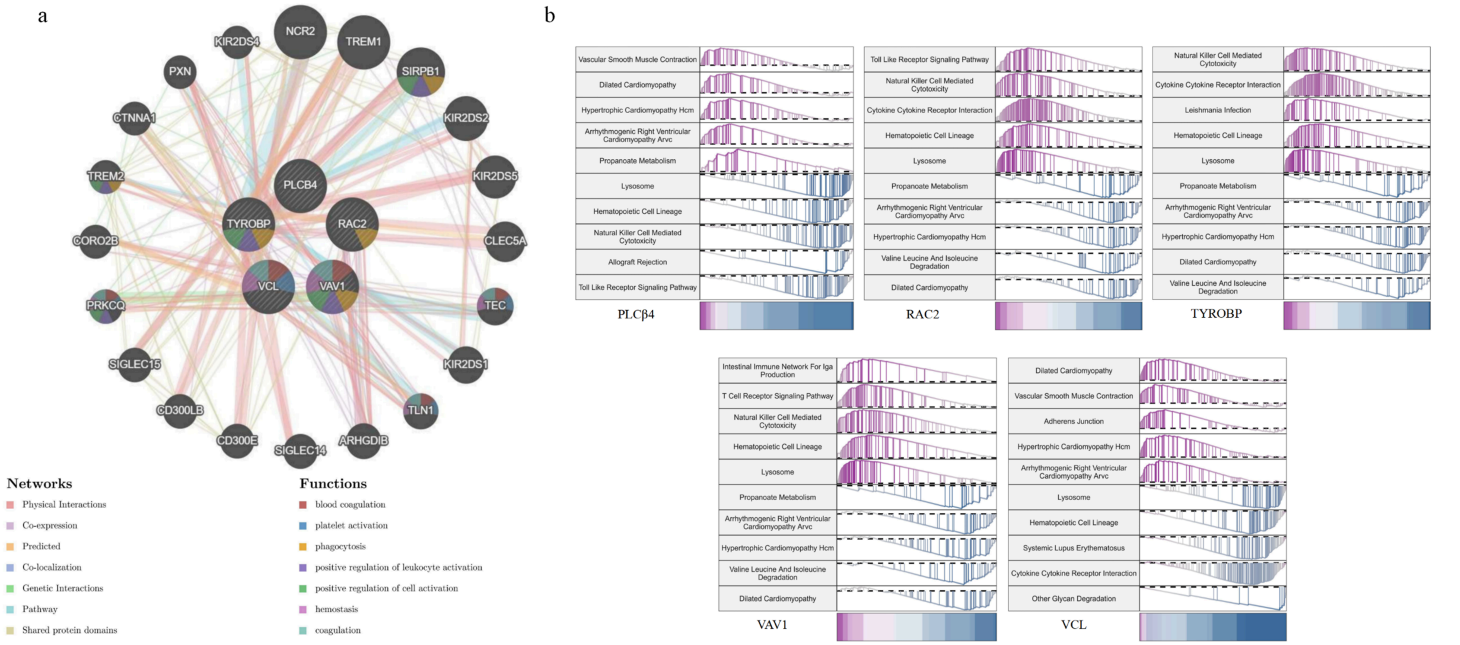


Figure 5

Figure 5

Functional Exploration of Biomarkers. (a) Interaction network of biomarkers and their predicted genes: Edge colors represent different interaction modes, while node colors denote the functional signaling pathways in which each gene is enriched. (b) Gene Set Enrichment Analysis (GSEA): The x-axis displays genes ranked by correlation coefficient, and the y-axis indicates the enrichment score. Curves above the dashed line suggest activation of the corresponding function/pathway under high gene expression, whereas curves below the dashed line indicate activation under low expression conditions.

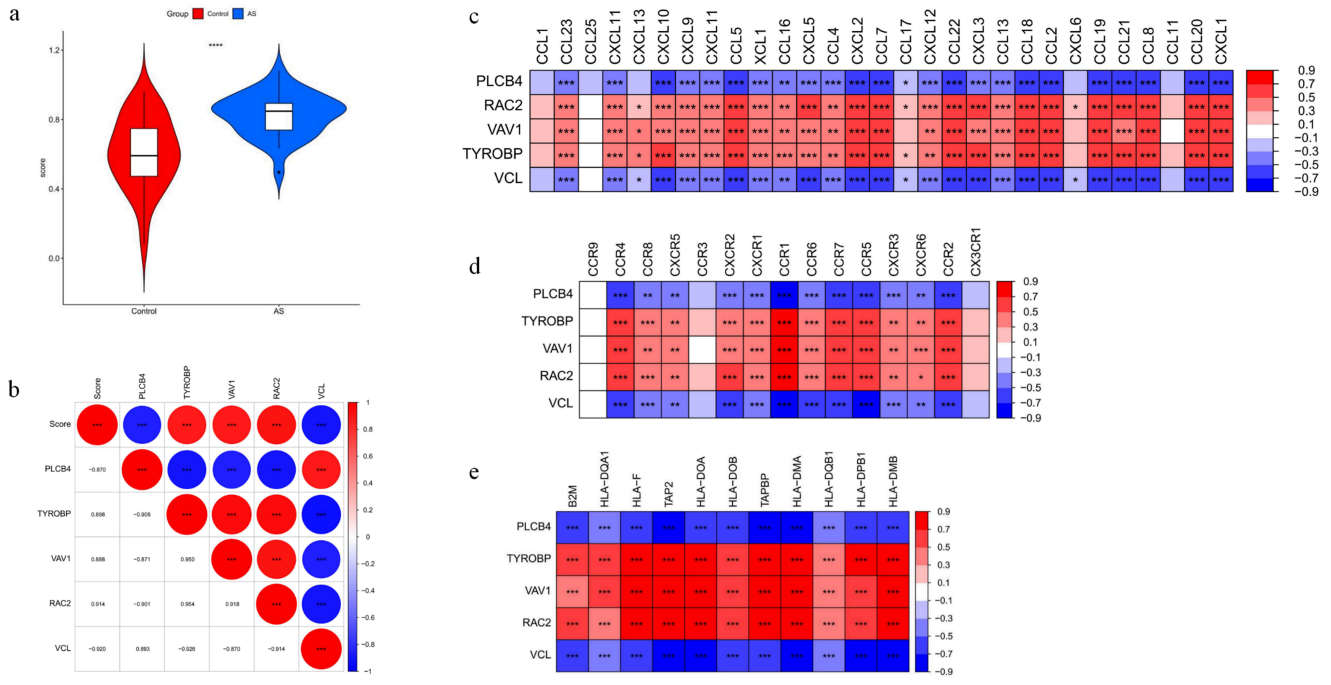


Figure 6

Figure 6

Correlation of Biomarkers with Inflammatory and Immune Factors. (a) Inter-group differences in inflammatory factor scores. (b) Correlation between inflammatory factor scores and biomarkers. (c) Correlation of 29 chemokines with biomarkers. (d) Correlation of 15 chemokine receptors with biomarkers. (e) Correlation of 11 MHC with biomarkers.

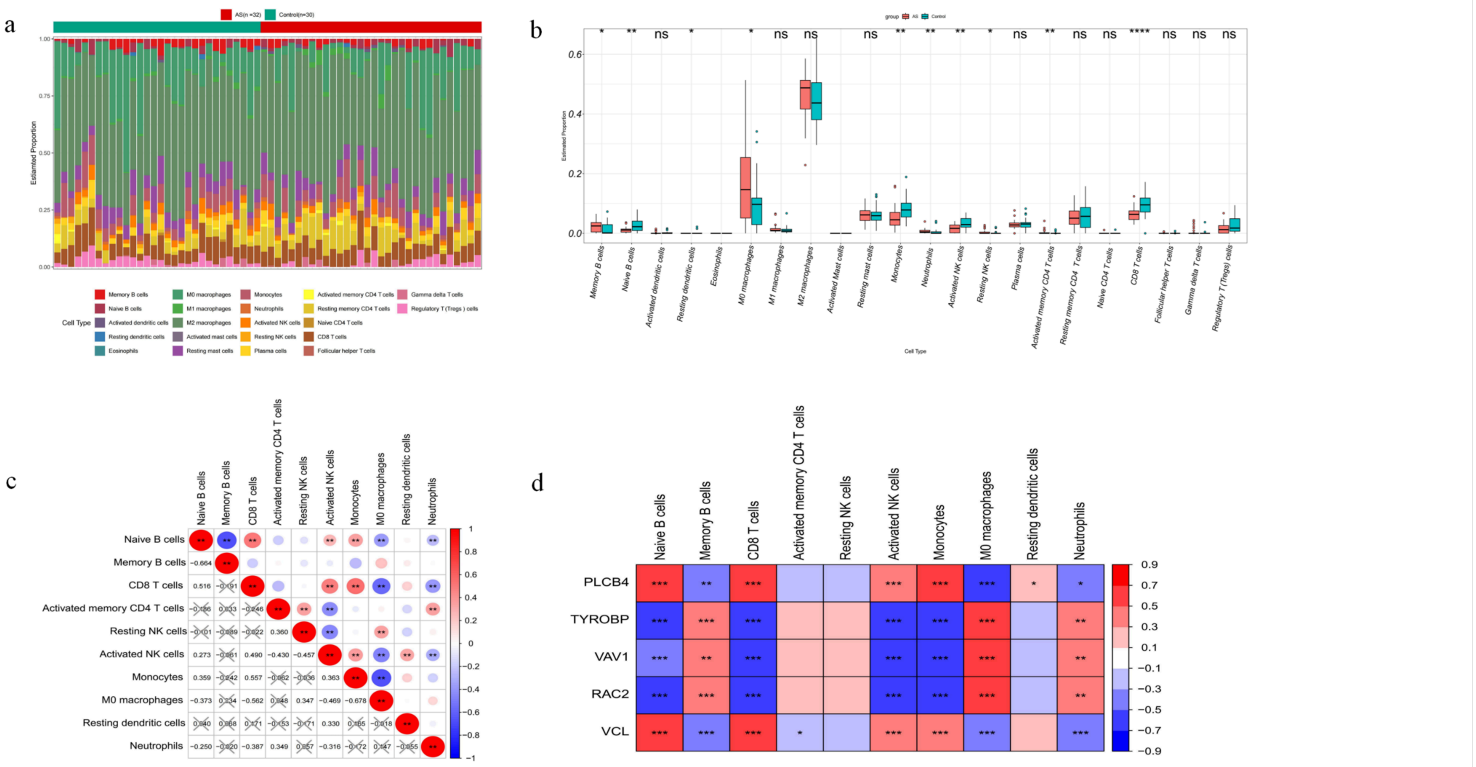


Figure 7

Figure 7

Immune Infiltration Analysis. (a) Distribution of 22 types of immune cells across samples in the training set. (b) Box plots showing expression differences of immune cells between the AS group and the control group in the training set. (c) Spearman correlation analysis among differentially expressed immune cells in the training set. (d) Spearman correlation analysis between the risk score and differentially expressed immune cells in the training set.

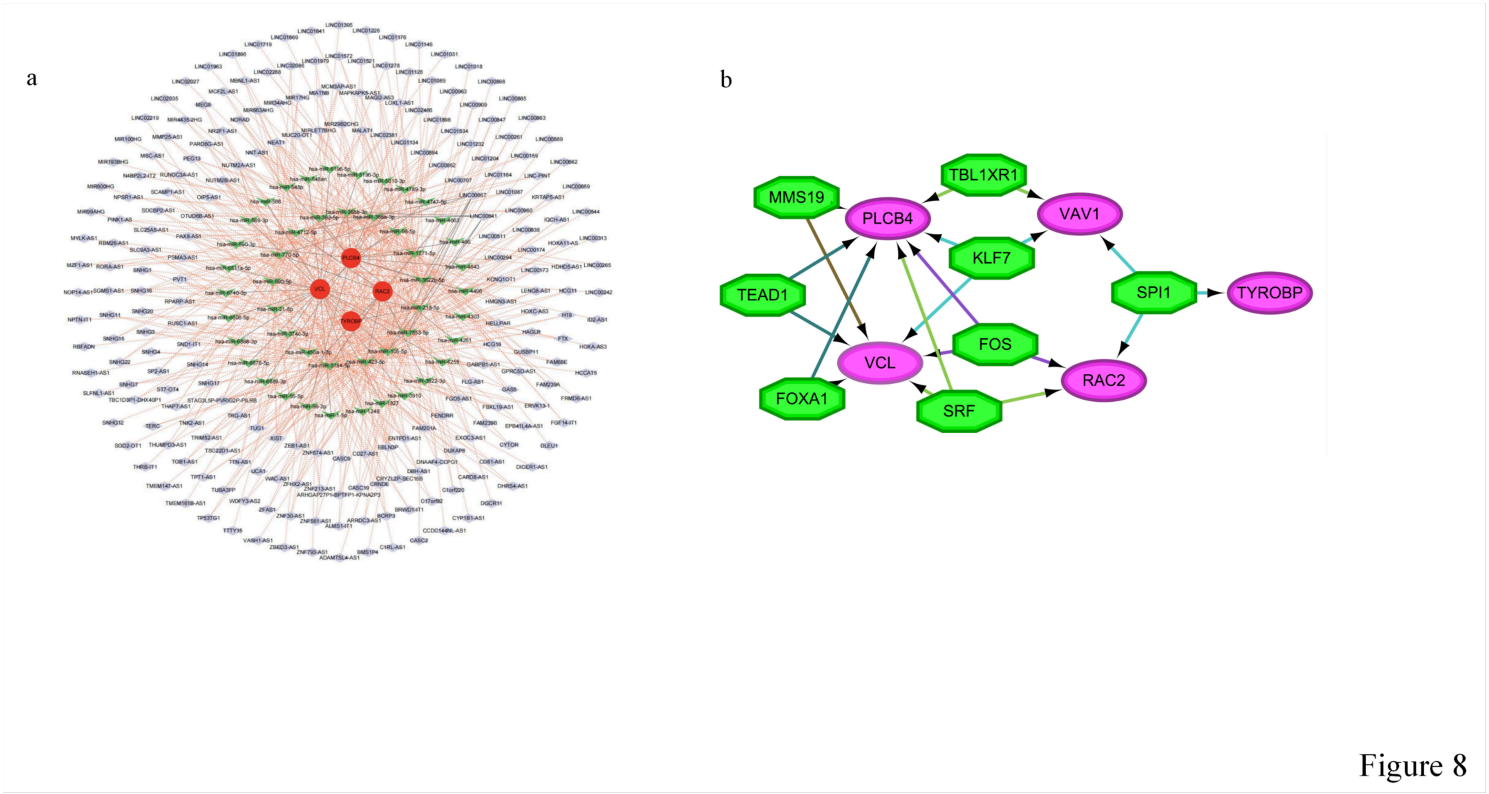


Figure 8

Regulatory Network. (a) LncRNA-miRNA-mRNA regulatory network: blue diamonds represent lncRNAs, green V-shapes represent miRNAs, and red circles represent biomarkers. (b) TF-mRNA regulatory network: green hexagons represent transcription factors, and purple ellipses represent biomarkers.

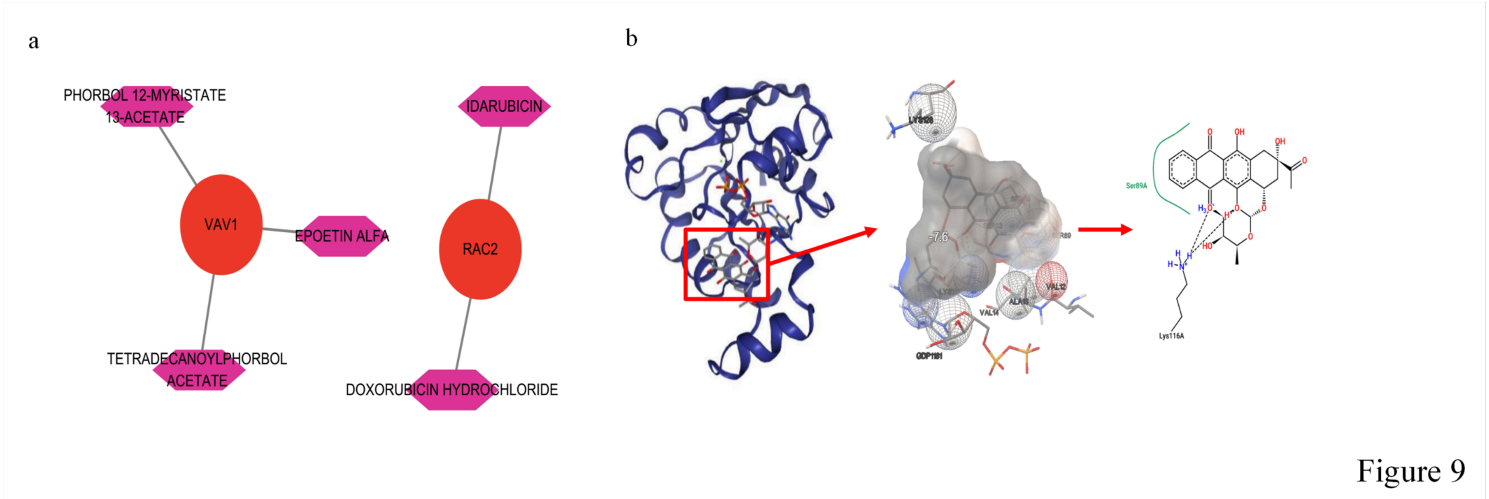


Figure 9

(a) Drug-Biomarker Network Diagram: Biomarkers are represented in red, and drugs are denoted by purple hexagons.(b) Molecular Docking Details of IDARUBICIN with RAC2.

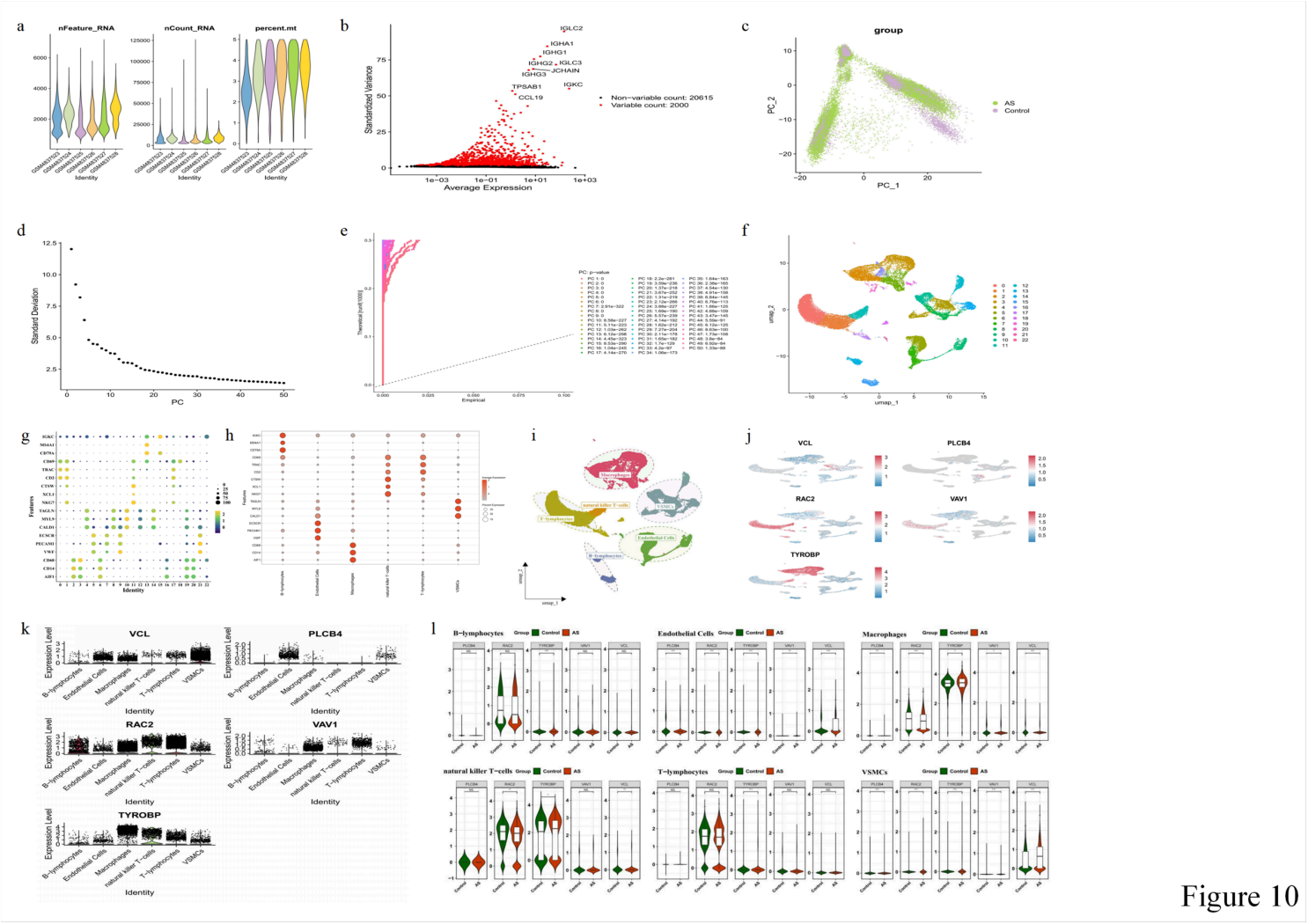


Figure 10

Figure 10

(a) Data Quality Control: Distribution plots showing gene counts, UMI counts per sample, and the percentage of mitochondrial gene expression. (b) Highly Variable Genes: Scatter plot of the top 2,000 highly variable genes. Genes highlighted in red represent the top 2,000 with the highest expression variance, while those in black indicate genes with low variance. (c) Linear dimensionality reduction was performed based on gene expression across all cells. (d) Visualization of the first 50 principal components, illustrating the distribution across selected dimensions. (e) Identification of principal components enriched with genes showing low P-values (P<0.05). (f) UMAP plot showing the distribution of cells across different clusters. (g) Cell Annotation–Cluster Level: Expression of marker genes within each identified cluster. (h)Expression of marker genes across annotated cell populations. (i) UMAP plot after cell type annotation, with colors representing distinct cell clusters. (j) Spatial expression visualization of biomarkers across cells in UMAP coordinates. (k) Distribution of biomarker expression levels across cells. (l) Comparative expression analysis of biomarkers across different sample groups within each cell type.

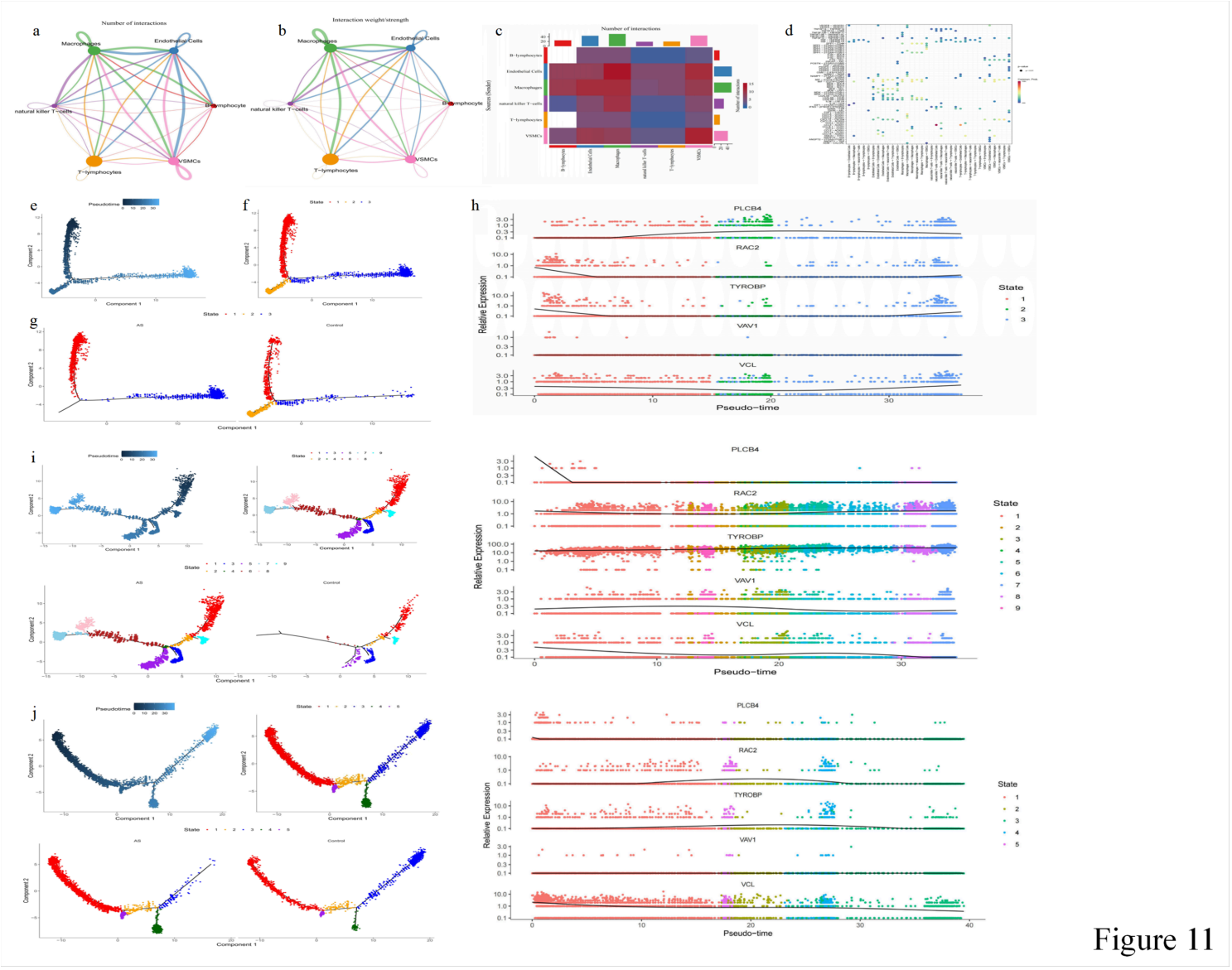


Figure 11

Figure 11

The analysis of cellular communication. (a) The size of each colored circle represents the number of cells in the corresponding population, with larger circles indicating higher cell counts. Arrows point from ligand-expressing cells to receptor-expressing cells, and the thickness of the lines corresponds to the number of ligand-receptor pairs identified between interacting cell populations. (b) Probability-weighted interaction strength (calculated as the sum of probability values). (c) Heatmap of ligand-receptor pair quantities. (d) Network of ligand-receptor interactions. (e,f,g) Pseudotemporal Analysis of Endothelial Cells Depicting the Differentiation Trajectory. (h) Temporal Expression Profiles of Biomarker Genes in Endothelial Cells. (i) The pseudo-temporal trajectory analysis of macrophages. (j) The pseudo-temporal trajectory analysis of macrophages and VSMCs.

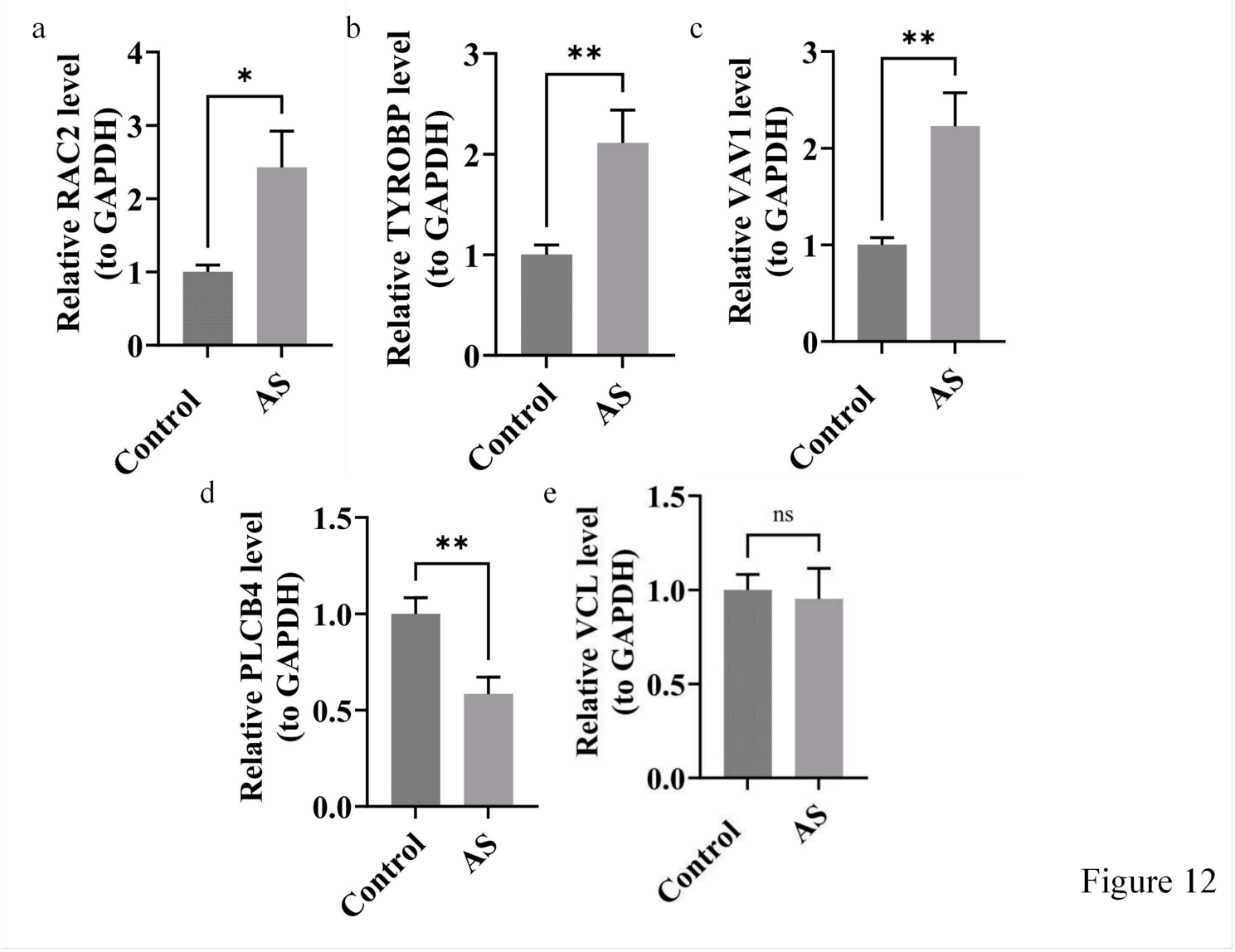


Figure 12

Figure 12

Analysis of differential gene expression by RT-qPCR.