

Identifying and Diagnosing Lytic Cell Death Genes in Atherosclerosis Using Machine Learning and Bioinformatics

Guolin Zhang^{1,*}, Ruicong Ma^{2,*}, Hongjin Jin^{1,*}, Qian Zhang², Wenhui Li², Yanchun Ding¹

¹Department of Cardiology, The Second Hospital of Dalian Medical University, Dalian, Liaoning, People's Republic of China; ²The Second Hospital of Dalian Medical University, Dalian, Liaoning, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yanchun Ding, Email yanchunding@dmu.edu.cn

Background: Lytic cell death (LCD) is gaining research attention in chronic inflammatory diseases such as atherosclerosis (AS). Our study investigates the role and mechanism of LCD in AS using machine learning and bioinformatics.

Methods: We sourced gene expression data and single-cell sequencing from the GEO database. Differential analysis identified differentially expressed genes (DEGs), which were then intersected with LCD-related genes to determine LCD-associated DEGs (LCDEGs). Machine learning was used to screen characteristic LCDEGs, and an artificial neural network (ANN) model was developed. The diagnostic accuracy of the model was assessed using ROC curves.

Results: The results demonstrated that the ANN model possesses a robust diagnostic ability in distinguishing between normal and AS cases, as well as identifying early and advanced stages. Unique AS subtypes were identified using a consensus clustering method. Two subtypes, C1 (non-immune subtype) and C2 (immune subtype), were delineated based on immune landscape analysis and gene set variation analysis functional enrichment. The chi-square test revealed that C1 was linked to early-stage (low-risk) atherosclerotic plaques, whereas C2 was associated with advanced-stage (high-risk) atherosclerotic plaques. At the single-cell level, LCDEG activity was calculated using AUCell and AddModuleScore. LCDEGs exhibited increased activity levels in macrophages within the initially classified cell subtypes. Moreover, they displayed higher activity in the “inflammation” subtype of specific macrophage subtype analysis.

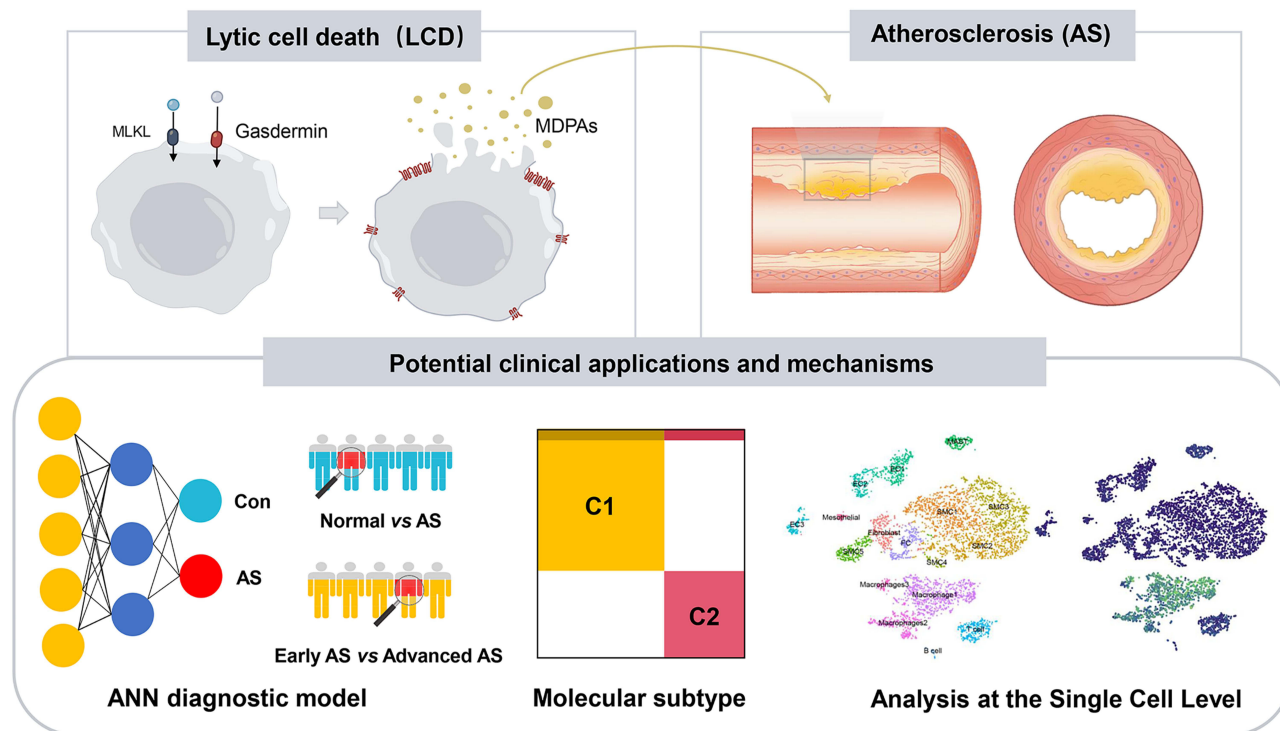
Conclusion: This study highlights the clinical potential of LCD in AS and suggests it involves a macrophage-mediated mechanism. We also experimentally identified and validated cytochrome B-245 β chain (CYBB) as a potential biomarker for AS.

Keywords: atherosclerosis, lytic cell death, lytic cell death-related genes, machine learning, bioinformatics, cytochrome B-245 β chain, CYBB

Introduction

The prevalence of cardiovascular diseases caused by atherosclerosis (AS) is a growing concern that poses a serious threat to human health and places a substantial burden on global society.¹ Generally, men are more prone to developing AS compared to women, particularly in middle age, with a male-to-female ratio of 2:1 in many populations.² However, the risk for women increases post-menopause, eventually equaling or even surpassing that of men.³ By 2030, it is estimated that the number of deaths due to cardiovascular diseases will reach a staggering 20 million.⁴ A key contributor to the rise in cardiovascular diseases is an increase in AS, a chronic arterial inflammatory disease characterized by the buildup of fatty deposits in arterial walls.² AS begins with the activation of endothelial cells in the blood vessels, leading to changes such as lipid deposition, fibrosis, calcification, and narrowing of the arteries.³ This condition significantly contributes to the incidence of high-risk cardiovascular events, such as stroke and myocardial infarction.⁵ In the early stages of AS, symptoms are generally mild, and the plaques formed are often stable with intimal thickening and macula formation.⁶

Graphical Abstract



However, as AS progresses, the risk of acute cardiovascular events significantly increases. This may be attributed to advanced plaques becoming thinner or thicker fibrous cap atheromatous plaques, which are unstable and prone to rupture.^{7,8} Consequently, acquiring a comprehensive understanding of the initiation and progression of AS is essential for efforts aimed at mitigating the incidence of acute cardiovascular events. By identifying and addressing the factors contributing to the development and instability of atherosclerotic plaques, the risk of cardiovascular events may be mitigated and health outcomes enhanced.

Programmed cell death includes apoptosis and lytic cell death (LCD). Pyroptosis, necroptosis, and ferroptosis—three forms of regulated cell death—are classified under the LCD category.⁹ Apoptosis maintains the integrity of the plasma membrane, with apoptotic effectors contributing to immune silencing.¹⁰ In contrast, LCD can result in the loss of plasma membrane integrity, leading to the release of cellular components, including nuclear contents, into the extracellular space.¹¹ Additionally, LCD allows the release of immunogenic cellular components, such as damage-associated molecular patterns (DAMPs) and multiple inflammatory cytokines, which elicit a robust inflammatory response.¹² Previous studies have shown that LCD dysregulation may contribute to various inflammatory diseases, including the development and progression of AS.¹³ RIPK3/MLKL-mediated necroptosis promotes plaque necrotic core formation and inflammatory response by destroying the integrity of the plasma membrane, releasing DAMPs and inflammatory factors; while ferroptosis aggravates oxidative stress through lipid peroxidation and enhances plaque vulnerability.^{14,15} These processes together can lead to the instability of atherosclerotic plaques.¹⁶ In the context of atherosclerosis, various types of cell death (including macrophage) are involved, and multiple forms of cell death are implicated. In this study, we primarily explore the potential pathophysiological mechanisms of lytic cell death in AS and identify key diagnostic and biological markers associated with lytic cell death in AS.

Recent breakthroughs in microarray technology, single-cell sequencing, and bioinformatics have significantly advanced the field of cardiovascular biomedical research.^{17,18} The reanalysis and integration of a large amount of data

stored in public databases can uncover new information about diseases.¹⁹ Recently, various machine learning models have been widely applied in the analysis of gene expression data.²⁰ Machine learning models combined with bioinformatics have become powerful tools for analyzing gene expression data and developing diagnostic models for cardiovascular diseases.²¹ For example, the integration of bioinformatics and machine learning approaches has identified biomarkers linking coronary AS to fatty acid metabolism, highlighting the potential of these methods in AS research.²¹ However, the application of this method in AS remains underexplored. This study used machine learning approaches with bioinformatics to screen for a cohort of disease-specific genes linked to LCD, develop a diagnostic model demonstrating substantial diagnostic efficacy, and identify a novel clinically significant subtype of AS. Analysis at the single-cell level partially elucidated the underlying mechanisms of LCD in AS. Additionally, cytochrome B-245 β chain (CYBB), identified as a disease signature biological macromolecule associated with LCD, emerged as a potential biomarker closely linked to the prognosis of AS. To the best of our knowledge, this study is an inaugural investigation into the potential impact of LCD on the development of AS, using an integrated approach that combines machine learning and bioinformatics. It establishes a foundation for subsequent research and presents alternative insights into the treatment strategies for AS.

Methods

The workflow of the research in this study is shown in Figure 1.

Data Collection

All bulk sequencing datasets and single-cell sequencing data were extracted from the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) database utilizing a comprehensive search strategy. The search was conducted using the following keywords: “Atherosclerosis” OR “Acute coronary syndrome” OR “Ischemic stroke” to ensure the inclusion of relevant datasets. Specific dataset platforms, such as Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray

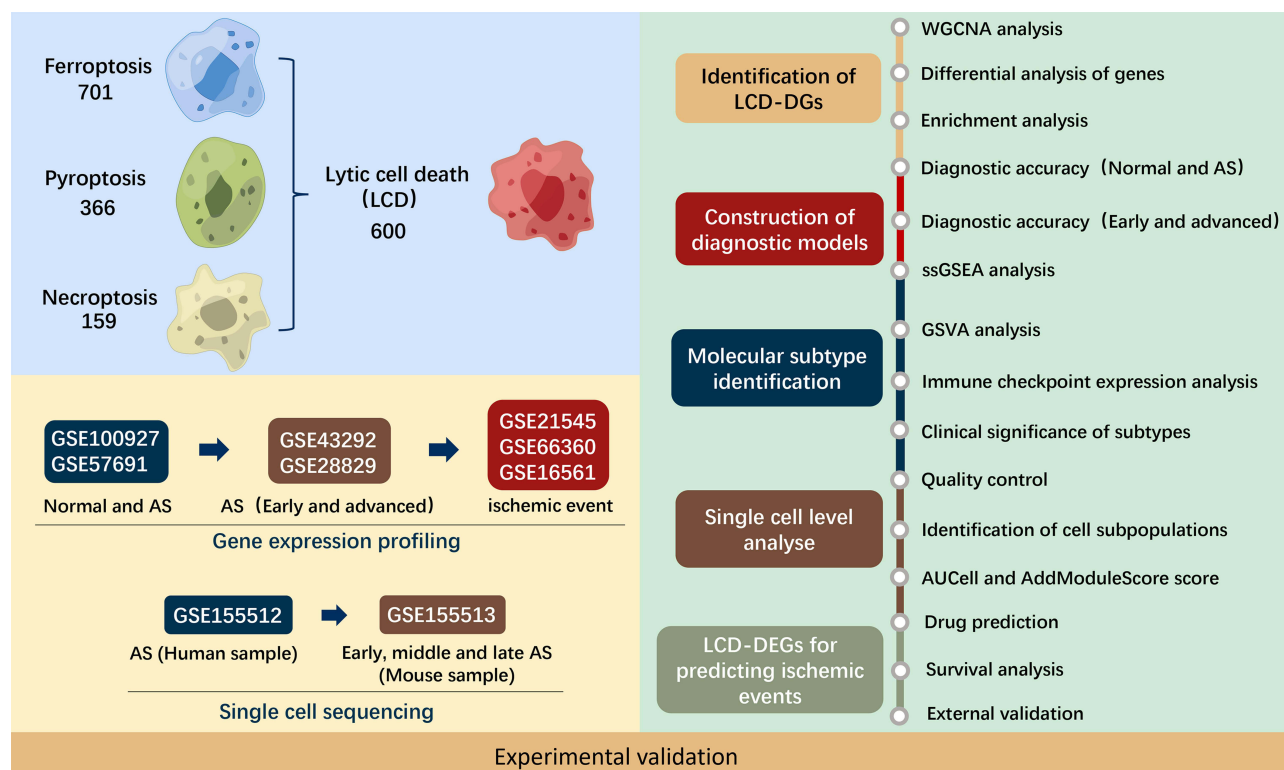


Figure 1 Flowchart of the workflow. Abbreviations are defined as follows: atherosclerosis (AS), differentially expressed genes (DEGs), gene set variation analysis (GSVA), Lytic cell death (LCD), Lytic cell death-associated genes (LCDGs), Lytic cell death-associated differentially expressed genes (LCD-DEGs), single-sample gene set enrichment analysis (ssGSEA), weighted gene co-expression network analysis (WGCNA).

(GPL17077) and Illumina HumanHT-12 V4.0 expression beadchip (GPL10558), were used for the data. Detailed information on the datasets is provided in Table 1. In addition, the human gene expression profile data used in this study came from the GEO database. According to Article 32(1) and (2) of the Measures for the Ethical Review of Life Science and Medical Research Involving Human Subjects (the People's Republic of China, effective February 18, 2023) (https://www.gov.cn/zhengce/zhengceku/2023-02/28/content_5743658.htm), this part of the study only used publicly available de-identified data, and therefore no ethical review was required. This part of the study did not require Institutional Review Board approval.

LCD-associated genes (LCDGs) consist of genes involved in pyroptosis, necroptosis, and ferroptosis. A total of 366 pyroptosis-related genes²² and 159 necroptosis-related genes²³ were extracted from relevant published literature. Additionally, 701 ferroptosis-related genes were obtained from the FerrDb database (<http://www.zhounan.org/ferrdb/index.html/>). After the removal of duplicate entries, 600 unique LCDGs were identified. In addition, oxidative stress-associated genes were retrieved from the GeneCards database (<https://www.genecards.org/>). Genes with association scores ≥ 10 were included with reference to previous evidence²⁴ to ensure that the included genes were closely related to the oxidative stress pathway. For a complete list of LCDGs and oxidative stress-associated genes, please refer to [Supplementary Table S1](#).

Preprocessing of the Gene Expression Profile

The gene expression profiles were preprocessed using the “limma” package (v.3.60.6) in RStudio (version 4.1.3, <https://www.rstudio.com/>).²⁵ Conversion between probe IDs and gene symbols was carried out based on the respective platform files. In cases where multiple probes corresponded to the same gene symbol, the average expression value was calculated.

Construction of Weighted Gene Co-Expression Network Analysis (WGCNA) and Identification of the Central Module

The “WGCNA” package (v.1.7) was used to construct the weighted gene co-expression network and identify the central module using all genes from the GSE100927 dataset, which included normal controls and AS.²⁶ The following steps were performed: (1) Selection of genes with a standard deviation greater than 0.7 between samples as the input dataset; (2) Removal of outliers using the “goodSamplesGenes” function; (3) Determination of an appropriate soft threshold using the “pickSoftThreshold” function; (4) Transformation of the correlation matrix into a topological overlap matrix (TOM); (5) Identification of modules through dynamic tree-cutting and determination of their relationship with AS; (6) Selection of the module with the highest Pearson correlation coefficient as the central module.

Table 1 Basic Information About the Dataset Used in Present Study.

ID	Samples	Tissue	Species	Platforms
GSE100927	35 Normal and 69 AS	Peripheral arteries	Homo sapiens	GPL17077
GSE57691	10 Normal and 9 AS	Abdominal aorta	Homo sapiens	GPL10558
GSE43292	32 Early AS and 32 Advanced AS	Carotid atheroma	Homo sapiens	GPL6244
GSE28829	13 Early AS and 16 Advanced AS	Carotid atheroma	Homo sapiens	GPL570
GSE66360	13 Normal and 56 ACS	Circulating endothelial cells	Homo sapiens	GPL570
GSE16561	13 Normal and 56 IS	Peripheral whole blood	Homo sapiens	GSE16883
GSE21545	97 AS	Carotid atheroma	Homo sapiens	GPL570
GSE155512	3 AS	Carotid atheroma	Homo sapiens	GPL24676
GSE155513	6 AS	Carotid atheroma	Mus musculus	GPL24247

Abbreviations: AS, Atherosclerosis; ACS, ACS Acute Coronary Syndrome; IS, Ischemic Stroke.

Identification of LCD-Associated Differentially Expressed Genes (DEGs)

The GSE100927 dataset was used to identify DEGs. DEGs were filtered using the “limma” package with an adjusted $P < 0.05$ and a $|\log_2 \text{Fold Change}| > 1$.²⁷ To obtain LCD-associated DEGs (LCDEGs) for further analysis, the intersection between DEGs and LCDGs was determined using the “venn” package (<https://cran.r-project.org/web/packages/venn/index.html/>). Volcano plots and heatmaps were generated using the “ggplot2” package (v.3.5.1).

Functional Enrichment Analysis

The “clusterProfiler” package (v.4.12.6) was used for conducting enrichment analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO), while the “ggplot2” package and “enrichplot” package (v.1.24.4) was used for visualizing the results.²⁸ Functional enrichment analysis was performed on the genes of the central module obtained by WGCNA. Similarly, functional enrichment analysis was performed on DEGs and LCDEGs. The statistical significance of enrichment was determined using an adjusted P value threshold of < 0.05 . To explore the variation in pathway activity between the two atherosclerotic subtypes, gene set variation analysis (GSVA), a nonparametric and unsupervised analysis method, was carried out using the “GSVA” package (v.1.24.2).²⁹ The gene sets “h.all.v7.5.1.symbols.gmt” “c2.cp.kegg.v7.5.1.symbols.gmt” and “c5.go.v7.5.1.symbols.gmt” were derived from the Molecular Signature Database (MSigDB, version 7.5).

Protein–Protein Interaction (PPI) Network Construction

A PPI network for LCDEGs was constructed using the STRING database (<https://string-db.org/>), with *Homo sapiens* as a species limitation and a confidence score threshold of > 0.6 .³⁰ To identify genes that exhibit functional associations with the target genes, the GeneMANIA database (<https://genemania.org/>) was used. In this study, we used GeneMANIA to investigate the correlation of LCDEGs with other potentially relevant genes.³¹

Screening LCD-Associated Signature Genes of AS via Machine Learning

The least absolute shrinkage and selection operator (LASSO) and random forest (RF) algorithms were used to identify characteristic genes associated with LCD in AS. The LASSO algorithm was implemented using the “glmnet” package (v.4.1–8; parameters: family = “binomial”, alpha = 1, type.measure = “deviance”, nfolds = 10).³² The lambda parameter was selected via 10-fold cross-validation to minimize deviance, resulting in an optimal lambda value of 0.015. RF was implemented using the “randomForest” package (v.4.7–1.2; parameters: ntree = 500,³³ mtry = sqrt(p), where p is the number of features), with mtry optimized through a grid search over values ranging from 1 to sqrt(p). An artificial neural network (ANN) was constructed using the “neuralnet” (v.1.44.2; parameters: hidden = 5, learning rate = 0.01) and “neuralnettools” packages (v.1.5.3), with the hidden layer size optimized based on model convergence and performance on a validation set. The eigengenes obtained from the previous method were used as input, and the gene score was used for training the ANN. For data splitting, the GSE100927 dataset (n=100 samples) was divided into 70% training (n=70) and 30% testing (n=30) sets using stratified sampling to maintain the proportion of control and AS groups. The GSE57691, GSE43292, and GSE28829 datasets served as independent test sets. Model performance was evaluated using receiver operating characteristic (ROC) curves generated by the “pROC” package (v.1.18.5), along with additional metrics including precision, recall, and F1-score, calculated using the “caret” package (v.6.0–94). In GSE100927 and GSE57691, samples were divided into control and AS groups, while in GSE43292 and GSE28829, samples were categorized into early and advanced AS groups. The specific results of all the ROC curves for the predictive indicators are shown in [Supplementary Table S2](#).

Processing of Single-Cell RNA-Sequencing (scRNA-Seq) Data and Identification of Cell Subpopulations

Three human AS samples from the GSE155512 dataset were integrated for analysis, along with six mouse samples from the GSE155513 dataset. The mouse samples were categorized into early, middle, and late groups of AS based on experimental conditions. The single-cell data was processed using the “Seurat” package (v.5.1.0) with the following

criteria: each gene had to be expressed in at least three cells, and each cell had to express at least 200 genes.³⁴ The percentages of mitochondria and rRNA were calculated using the “PercentageFeatureSet” function. Cells were required to express more than 200 genes, have fewer than 7000 genes, have less than 20% mitochondrial content, and have at least 100 unique molecular identifiers per cell. The “Harmony” package (v.1.2.1) was used for batch corrections.³⁵ The data was then normalized using the “lognormalization” function, and highly variable genes were identified using the “FindVariableFeatures” function. The gene expression values were scaled using the “ScaleData” function, and principal component analysis (PCA) was performed for dimension reduction. Subsequently, the “FindNeighbors” and “FindClusters” functions were applied (with a resolution of 0.6) to cluster the cells and assign them annotations. DEGs between cell clusters were identified using the “FindAllMarkers” function (log-fold change threshold of 0.25).

Cell annotations were further verified by combining the “SingleR” package (v.2.6.0) annotation and marker genes identified in previous literature.^{36,37} For the mouse sample data, the same processing steps as mentioned above were applied. Additionally, the macrophage clusters were extracted separately for dimensional clustering analysis to obtain and annotate macrophage subpopulations. The 12 selected LCDGs from the previous screening were used for subsequent scoring at the single-cell level. We used two methods for scoring: (1) the “AUCell”³⁸ package (v.1.2.0) was used to score each cell, and (2) the “AddModuleScore” function of the “Seurat” package was used for per cell scoring. The obtained scores were mapped to Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP), and clusters of active cells were visualized using the “ggplot2” package. The “ggpubr” package (v.0.6.0) was used for visualization and statistical analysis. Functional enrichment analysis at the single-cell level was performed using the “singleseqset” package (v.0.1.2). DEGs among macrophage subtypes were identified using the “FindAllMarkers” function (log-change threshold of 0.25) and used as a background gene set. The CIBERSORT algorithm³⁹ was used to determine the enrichment score of different macrophage subtypes.^{40,41}

Consensus Clustering

Consensus clustering is an unsupervised technique widely used in the analysis of expression profiling data. It is aimed at identifying clinically significant disease subtypes that may elude detection through conventional testing methods.^{42,43} Consensus clustering analysis was conducted using the “ConsensusClusterPlus” package (v.1.68.0), based on the expression profiling of LCDEGs.⁴⁴ The determination of the number of clusters and their stability was achieved through a consensus clustering algorithm consisting of 1,000 iterations. PCA was performed using the “prcomp” function, and a categorical dot plot was generated to validate the clustering results.⁴⁵ To ensure the reproducibility of the subtypes, we also extracted gene expression profiles of LCDEGs from the GSE43292 and GSE28829 datasets for clustering. Furthermore, the relationship between various subtypes and both early and advanced stages of AS was analyzed using the chi-square test. The analysis was conducted using the “ggstatsplot” package (version 1.12.4).

Immune Infiltration Landscape Analysis

The single-sample gene set enrichment analysis (ssGSEA) is a digital cytometry algorithm that is based on a rank model.^{46,47} Briefly, ssGSEA can take gene expression profiles as input to assess the relative proportion of immune cells in a specific sample. In this study, ssGSEA, based on the “GSVA” package (v.1.42.0), was used to analyze the immune infiltration landscape of the two subtypes of AS. The analysis was conducted using the LM22 immune cell gene signature matrix (Newman et al, 2015).¹⁶ The LM22 matrix includes 22 immune cell type signatures, including naive B cells, memory B cells, plasma cells, CD8+ T cells, CD4+ naive T cells, regulatory T cells (Tregs), T follicular helper cells, NK cells (resting and activated), monocytes, macrophages (M0, M1, and M2), dendritic cells (resting and activated), mast cells (resting and activated), eosinophils, and neutrophils. LM22 immune cell characteristics can be obtained in the study of Newman et al.¹⁶ Additionally, the “rstatix” package (v.0.7.2) was used to compare the expression levels of chemokines and immune checkpoints between the two subtypes.

Cumulative Hazard Rate of Ischemic Events

The analysis was based on the gene expression profile of GSE21541, which encompassed 97 samples of peripheral blood mononuclear cells (pbmcs). The occurrence of an ischemic event was defined as the outcome event. The cumulative hazard rate was calculated using the “survival” (v.3.7) and “survminer” packages (v.0.4.9).

Discovery of Potential Drugs for Treating AS Based on LCDEG

The Drug-Gene Interaction Database (DGIdb v4.2, <https://dgidb.org/>) was used to identify potential LCDEG-based drugs for the treatment of AS.⁴⁸

Animal Model

Six male C57BL/6J mice (SPF grade, six weeks old) and 12 male ApoE knockout (ApoE^{-/-}) mice of the same grade and age were obtained from the Experimental Animal Center of Dalian Medical University. Following a week of adaptive feeding, the C57BL/6J mice were maintained on a standard diet for a further 12 weeks. At the same time, the ApoE^{-/-} mice were randomly divided into two groups: the early group (n=6) and the advanced group (n=6). The advanced group was fed a high-fat diet for 12 weeks, while the early group was fed a standard diet for six weeks and then switched to a high-fat diet for six weeks. Twelve weeks later, the mice were anesthetized via intraperitoneal injection with Ketamine (100 mg/kg) and Diazepam (5 mg/kg), and blood samples were collected from the tail vein. Following the collection of blood samples, the mice were euthanized with sodium pentobarbital (150 mg/kg) administered intraperitoneally, and the aorta was extracted and divided into two parts: (1) The aortic arch (from the ascending aorta to the origin of the thoracic aorta) was placed at 4 °C and fixed in 4% paraformaldehyde for 24 hours. (2) The rest of the aorta was immediately transferred to liquid nitrogen and stored at -80 °C for detection of relevant proteins and gene expression. All animal experiments were approved by the Dalian Medical University Animal Care and Ethics Committee and conducted in accordance with the National Institutes of Health (NIH) Guide for the Care and Use of Laboratory Animals (8th Edition).⁴⁹ The experimental procedures adhered to the principles outlined in this guideline to ensure the welfare of laboratory animals. Additionally, the reporting of in vivo experiments followed the ARRIVE guidelines to ensure transparency and reproducibility.

Cytokines and Plasma Lipid Analysis

Mouse blood was centrifuged at 1500 × g for 10 min to obtain serum. The serum levels of total cholesterol (TC), triglycerides (TG), high-density lipoprotein (HDL-C), low-density lipoprotein (LDL-C), CYBB, IL-1β, TNF-α, and IL-6 were determined using a triglyceride assay kit (E-TSEL-H0025, Elabscience), total cholesterol assay kit (E-BC-K109-M, Elabscience), HDL-C assay kit (E-BC-K221-M, Elabscience), LDL-C assay kit (E-BC-K205-M, Elabscience), Mouse Cybb/Cytochrome b-245 heavy chain (CYBB) ELISA Kit (E-AB-70273, Elabscience), Mouse Interleukin 1 Beta (IL-1β) ELISA Kit (E-EL-M0037c, Elabscience), Mouse Tumor Necrosis Factor Alpha (TNF-α) ELISA Kit (E-EL-M3063, Elabscience), and Mouse Interleukin 6 (IL-6) ELISA Kit (E-AB-F1207UD, Elabscience) according to instructions.

Histology and Immunohistochemistry

The mouse heart and a segment of the aorta were excised, fixed using paraformaldehyde, and subsequently embedded. Uniform serial sections were prepared at the level of the aortic sinus. These frozen sections were then stained with hematoxylin and eosin (HE) and anti-CYBB (GB112362-50, Servicebio) in accordance with the manufacturer’s protocol, and images were captured using a light microscope. Image-Pro Plus 6.0 software was used to quantify plaque area and plaque area/lumen cross area (%).

Western Blotting

Proteins were isolated from murine aortic tissues, and their concentrations were quantified with a BCA Protein Concentration Assay Kit (P0010, Beyotime Biotechnology, China). Subsequently, the protein samples were transferred onto a polyvinylidene difluoride membrane (ISEQ00010, Millipore, USA) and subjected to electrophoresis using a 10%

sodium dodecyl sulfate-polyacrylamide gel (P0690, Beyotime Biotechnology, China). Following a 15-minute sealing period, the primary antibody (Anti-NOX2/gp91phox antibody, EPR28415-13, Abcam) was incubated overnight at 4 °C. Subsequently, the secondary antibody (Goat Anti-Rabbit IgG H&L (HRP), ab6721, Abcam) was incubated for 2 hours at room temperature. Chemiluminescence analysis was conducted using an ECL kit (180–5001, Tanon, China), followed by image acquisition. Protein expression levels were normalized using β -actin as a reference and the Anti-beta Actin antibody (ab8227, Abcam). Relative protein quantification was performed using ImageJ software.

Statistical Analysis

Statistical analyses were conducted using R software (v4.1.3) through RStudio (v4.1.3). Correlation coefficients were assessed using Spearman's rank correlation analysis, with p-values adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) method to control the type I error rate. Chi-square tests were conducted using the "ggstatsplot" package (v0.9.0), with FDR correction applied to account for multiple comparisons. One-way ANOVA was used for normally distributed data, and Kruskal–Wallis tests were used for nonparametric comparisons among three or more groups, with post-hoc pairwise comparisons adjusted using the Bonferroni method. All statistical tests were two-sided, and p-values < 0.05 were considered statistically significant unless otherwise specified.

Results

Co-Expression Network Creation and Hub Module Identification

WGCNA identified 16 modules that were each assigned a different color. Among these modules, the turquoise module exhibited the most significant correlation with AS ($r = 0.72$, $P = 4E-18$) ([Supplementary Figure 1A–D](#)). Consequently, our focus was on understanding the biological functions of the genes within the turquoise module using KEGG and GO analysis. In KEGG pathway analysis, a majority of the genes were found to be enriched in the chemokine signaling pathway, NOD-like receptor signaling pathway, and leukocyte transendothelial migration pathway, all of which are associated with immune regulation and cell death ([Figure S1E](#)). GO annotation results revealed that these genes were primarily enriched in pathways pertaining to positive regulation of cytokine production and lytic vacuole membrane ([Figure S1F](#)). Based on these findings, we propose that the pathological mechanism of AS is intricately associated with the regulation of cell lysis and immune regulatory pathways.

Identification of LCDEGs

Following preprocessing and normalization of the gene profiles, differential expression analysis was conducted. A total of 437 DEGs were identified using adjusted P value < 0.05 and $|\log_2 FC| > 1$ as the screening criteria. These DEGs were then intersected with a set of 600 LCDGs, yielding 12 overlapping genes, as illustrated in the Venn diagram ([Figure 2A and B](#)). These 12 genes were further consolidated into a new gene set referred to as LCDEGs. Our subsequent analysis aimed to compare the mRNA expression profiles of these 12 genes in AS and control cases. Among the 12 LCDEGs, 11 genes were up-regulated, while a single gene was down-regulated ([Figure 2D and E](#)). The interplay among the LCDEGs was investigated using the Spearman correlation test. Notably, HLPDA expression showed a negative correlation with other genes, while there was a positive correlation among the remaining genes ([Figure 2C](#)).

Functional Enrichment and PPI Analyses

The pathway enrichment results of DEGs indicated that these genes were actively involved in several biological processes such as regulation of T cell activity and regulation of inflammatory responses, and were closely related to pathways such as NOD-like receptor signaling, antigen processing, and presentation ([Figure 2F and G](#)). In the network relationship analysis of DEGs and enrichment results, DEGs were simultaneously enriched in multiple signaling pathways. Notably, seven LCDEGs, particularly CYBB, PYCARD, and IL1B, were repeated in the GO and KEGG enrichment results ([Figure 2H and I](#)).

Subsequent GO enrichment analysis for LCDEGs showed that these genes were prominently enriched in pathways related to the positive regulation of cytokine production, cellular response to biotic stimuli, I-kappaB kinase/NF-kappaB

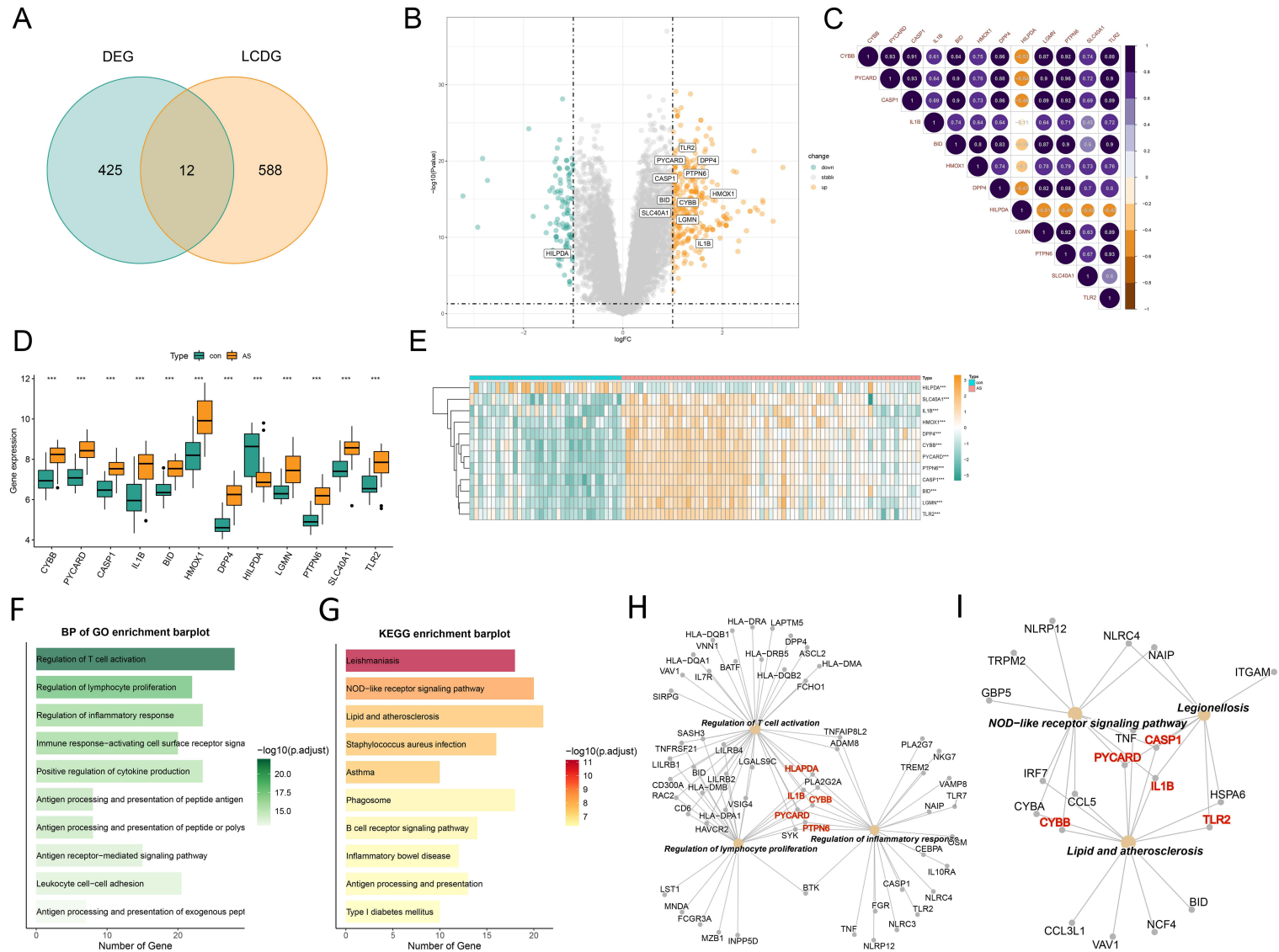


Figure 2 Identification of LCD-associated differentially expressed genes (LCDEGs). **(A)** Venn diagram showing the overlap between DEGs and LCDGs. **(B)** Volcano plot illustrating the LCDEGs. Each dot represents a gene, with Orange dots representing upregulated genes and green dots representing downregulated genes. **(C)** Correlations between LCDEGs in AS. **(D)** The difference in the gene expression profile of LCDEGs between AS and normal control. **(E)** Heatmap visualizing the LCDEGs. Rows represent genes, and columns represent samples. **(F)** The Gene Ontology (GO) enrichment barplot of DEGs. **(G)** The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment barplot of DEGs. **(H)** Network relationships between DEGs and GO enrichment. **(I)** Network relationships between DEGs and KEGG enrichment. Red represent LCDEGs.*** indicates $P < 0.001$ compared to the control group.

signaling, the inflammasome complex, the vacuolar lumen, and other pathways (Figure S2A). KEGG enrichment analysis demonstrated that 12 LCDEGs displayed significant enrichment in several pathways, including lipid and AS, necroptosis, NOD-like receptor signaling, ferroptosis, and the Toll-like receptor signaling pathway (Figure S2B). The protein interaction network, created using the STRING database, is displayed in Figure S2C. Furthermore, Figure S2D presents the predicted interactors associated with LCDEGs according to the GeneMANIA database.

Identifying LCDEG Features and Building the ANN Model via Machine Learning

To determine the diagnostic significance of LCDEGs, the LASSO and RF algorithms were used to identify characteristic genes associated with both AS and LCDEGs. By integrating the outcomes of these two algorithms, a total of eight characteristic genes related to LCD were identified: CASP1, CYBB, DPP4, HMOX1, IL1B, PTPN6, PYCARD, and HILPDA. These eight genes were used to construct a neural network (Figure 3A) and further evaluated using the ROC curve. In the training set (GSE100927) and the test set (GSE57691), the area under the curve (AUC) values were 97.3% and 78.9%, respectively, highlighting the high accuracy of the ANN model (Figure 3B and Figure S4B). In the other two test sets (GSE43292 and GSE28829), the AUC values were 79.2% and 82.7%, respectively (Figure 3C, Figure S5B). These results demonstrate that the ANN model, based on LCDGEGs, can accurately diagnose various stages of AS. Consequently, these eight LCDGEGs were designated as signature genes for AS. To further assess the individual diagnostic value of these characteristic genes identified through the machine learning approach, ROC curves were constructed and the AUC was calculated. The AUC values ranged from 0.771 to 0.953, indicating good diagnostic performance (Figure 3D–K). Additionally, external validation was performed using the GSE57691, GSE43292, and GSE28829 datasets to assess the gene expression profiles of the eight AS signature genes (Figures S3A–H, S4A and S5A). Subsequently, ROC curves were constructed and the AUC values ranged from 0.567 to 0.956, 0.545 to 0.849, and 0.683 to 0.976 for the three test sets, respectively (Figures S3I–P, S4C–J, and S5C–J). These results confirm the excellent diagnostic performance of these genes. The model performance correlation index is shown in Supplementary Table S2.

Clustering Based on LCDEGs and Clinical Significance of Subtypes

By using the consensus clustering method, we clustered samples based on the expression profiles of 69 AS. Two optimal subtypes were determined based on the analysis of the CDF plot, the relative change of the area under the CDF curve, the tracking plot, and the consensus matrix. Using the gene expression profiles of 12 LCDEGs, we divided the AS samples into two clusters, namely cluster A and cluster B (Figure 4A–C). There were statistically significant differences in the expression of these 12 LCDEGs between the two groups ($P < 0.05$) (Figure 4D and E). The PCA plot (Figure 4F) clearly displays the distinct distributions between the two clusters. Furthermore, the results of immune infiltration demonstrated that the levels of activated B cells, activated CD4 T cells, activated CD8 T cells, activated dendritic cells, CD56 bright natural killer cells, eosinophils, immature dendritic cells, MDSC, macrophage, mast cells, monocytes, natural killer cells, plasmacytoid dendritic cells, regulatory T cells, T follicular helper cells, type 1 T helper cells, and type 17 T helper cell significantly increased in cluster B ($P < 0.05$) (Figure 4G). The correlation between gene expression and immune cell infiltration is shown in Figure 4H. We examined the differential expression of three gene categories: HLA genes, immune checkpoint genes, and chemokine genes. In cluster B, the expression of most chemokines, such as CXCL16, CXCL9, CXCL11, CCL5, CCR5, CXCR3, and CXCR6, as well as most immune checkpoints, such as LAG3, CTLA4, TNFRSF9, ICOS, CD80, PDCD1LG2, and TIGIT, were significantly higher compared to cluster A ($P < 0.05$) (Figure S6A and S6C). Additionally, the oxidative stress score of cluster B was significantly higher than that of cluster A ($P < 0.05$) (Figure S6B). The GSVA enrichment results for the two clusters are presented in Figure S6D–E. Notably, cluster B exhibited significantly higher scores for functionally enriched pathways related to immune response compared to cluster A. These pathways were associated with peroxisomes, reactive oxygen species pathway, interferon-gamma response, interferon alpha response, IL6-JAK-STAT3 signaling, inflammatory response, mTOR signaling, lysosome, antigen processing and presentation, Toll-like receptor signaling pathway, natural killer cell-mediated cytotoxicity, NOD-like receptor signaling pathway, and B cell receptor signaling pathway. Based on these findings, cluster A has been classified as the non-immune subtype, C1, whereas cluster B has been identified as the immune subtype, C2. Furthermore, a high expression of LCDEGs was associated with the C2 subtype, whereas the C1 subtype exhibited the opposite pattern. The

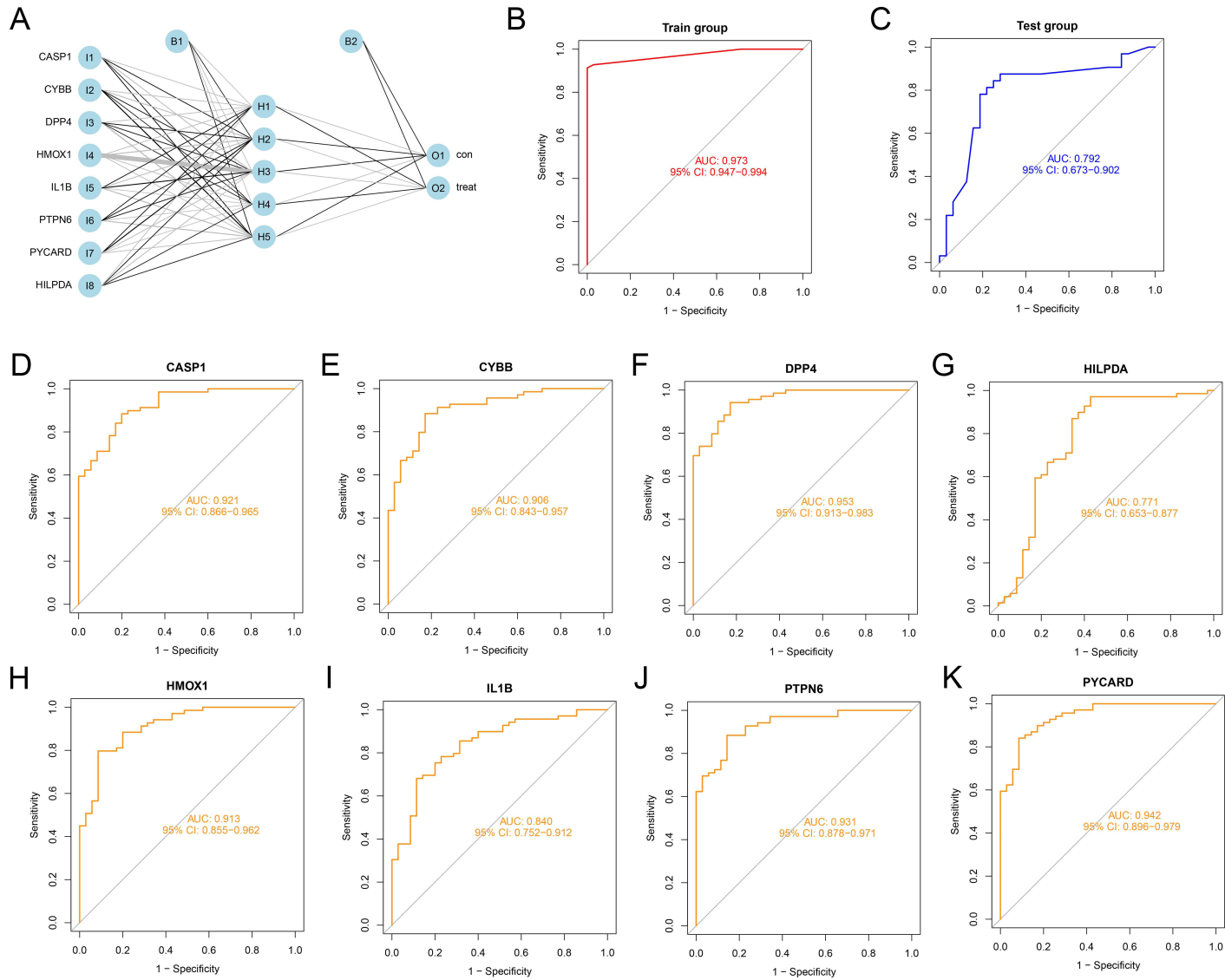


Figure 3 Diagnostic value of LCDEGs in AS. **(A)** Structure of the ANN model: I1–I8 represent the input layers (scores and weights of eight AS signature genes), H1–H5 represent the hidden layers, and O1–O2 represent the output layers (sample attributes). **(B and C)** Receiver operating characteristic (ROC) curves evaluating the diagnostic performance of the neural network model in the GSE100927 (training set) and GSE43292 (testing set). **(D–K)** ROC curves for the AS signature genes.

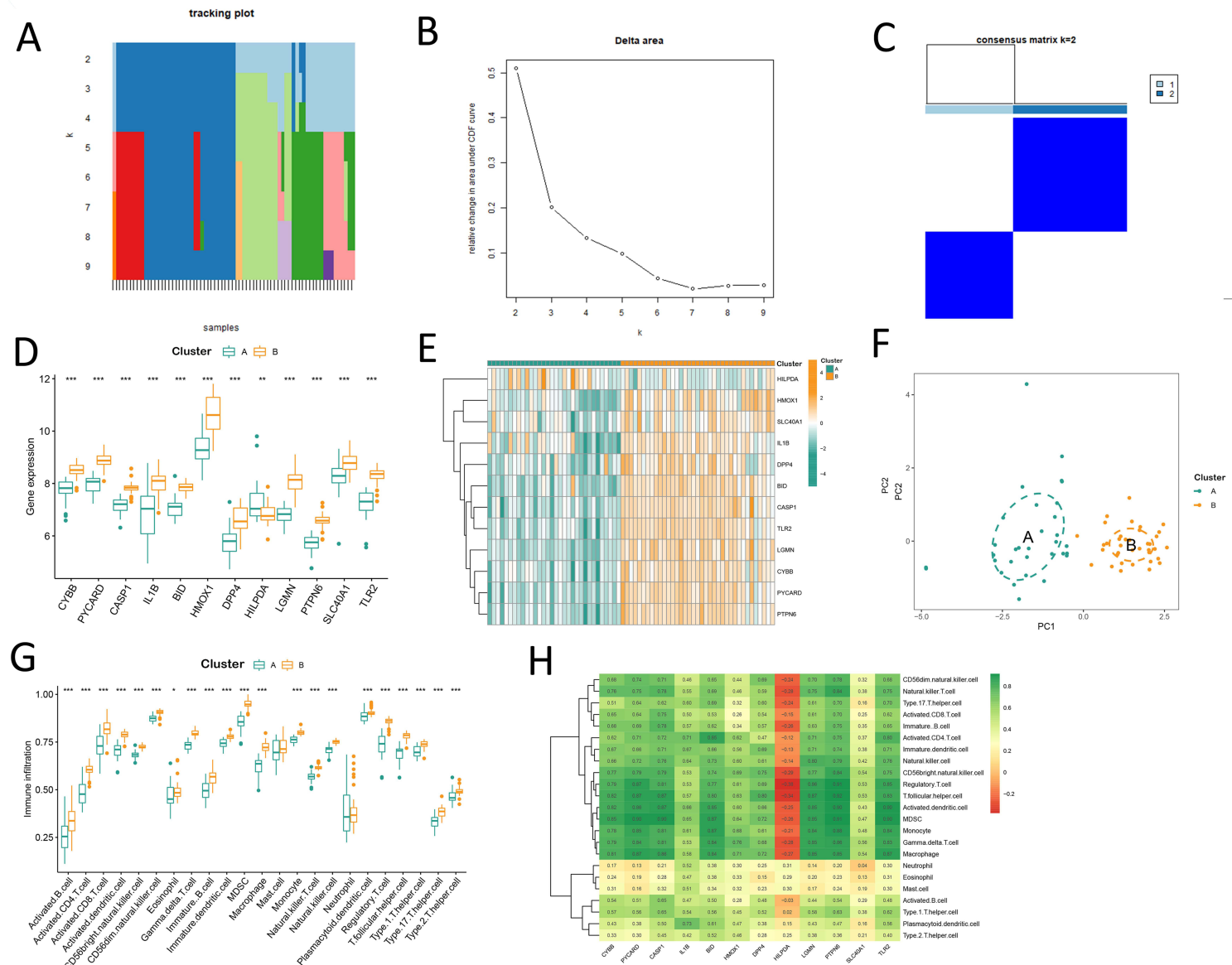


Figure 4 Construction of AS clusters based on LCDEGs in GSE100927. Unique immune and functional pathway signatures of the clusters. **(A)** Tracking plot displaying sample classification for $k = 2-9$. **(B)** Consensus cumulative distribution function (CDF) plot for $k = 2-9$. **(C)** Consensus matrix heatmap for $k = 2$. **(D)** Boxplots illustrating the expression of LCDEGs between the clusters. **(E)** Heatmap showing the expression of LCDEGs between the clusters. **(F)** Principal component analysis (PCA) plot demonstrating the classification of AS into two clusters based on the gene expression profile of LCDEGs. **(G)** Differences in the abundance of immune cells between the clusters. **(H)** Heatmap showing the correlation of LCDEGs with immune cells. * indicates $P < 0.05$ compared to cluster A, ** indicates $P < 0.01$ compared to cluster A, and *** indicates $P < 0.001$ compared to cluster A.

classification results based on LCDEGs for the early and late AS datasets were in line with previous results (Figure 5A and B). The majority of early AS samples fell into the C1 subtype category, whereas advanced AS samples predominantly belonged to the C2 subtype (Figure 5C and D). These findings were statistically significant based on the results of the chi-square test ($P < 0.05$).

Single-Cell Sequencing Results of Human Atherosclerotic Tissue Samples

Cells were grouped into 18 distinct clusters using the FindClusters function (Figure 6A). The results of cell subpopulation annotations are presented in Figure 6B. Important markers for these cells are illustrated in Figure 6C. AUCell scoring results were obtained and are presented in Figure 7A. The scoring outcomes of the AddModuleScore can be observed in Figure 7B. Notably, both of these scores exhibited significantly higher values in macrophages compared to other cell types, indicating that LCDEGs exhibited heightened transcriptional activity or substantial expression in macrophages. Among the 12 LCDEGs analyzed, macrophages demonstrated relatively high expression levels for CASP1, IL1B, CYBB, HMOX1, SLC40A1, and LGMN (Figure 7C). Violin plots qualitatively visualize the expression of 12 LCDEGs in various cell subtypes, and most of the LCDEGs are highly expressed in macrophages. The tSNE plots for each cell subtype are shown in Supplementary Figure 7. FOLR2, MS4A7, TRPV4, ADAP2, C3, ACSM5, SIGLEC1, C2, IGF1, EBI3, etc. can be used as cell markers of the first subtype of macrophages (Macrophages 1). Genes such as MYCL, S100Z, CCR2, LINC00877, AL034397.3, LUCAT1, LILRA1, C19orf38, C15orf481, LILRA5, and PRAM1 are used as cell markers of the second subtype of macrophages (Macrophages 2). Supplementary Table S3 shows the list of marker genes annotated for each subpopulation.

Single-Cell Sequencing Results of Mouse Atherosclerotic Tissue Samples

The marker profile of macrophage subsets in a previously reported single-cell sequencing data set of atherosclerotic mouse tissue is visualized in Figure 8A. GSEA of cell subpopulations was conducted using the “singleseqset” package, with results displayed in Figure 8B. The visualization results for the annotation of cell subpopulations are presented in Figure 8C and D. We examined the ratio of macrophage subtypes at the early, middle, and late stages (Figure 8E). We used the AddModuleScore to obtain scoring results for macrophage subtypes, as shown in Figure 8F. The IM subtype corresponds to blood-derived inflammatory macrophages and exhibits pathway enrichment results associated with inflammation, such as IL6-JAK-STAT3 and TNF- α . The RLM subtype represents tissue-resident inflammatory macrophages, with GSEA pathway enrichment results reflecting the inflammatory response and interferon alpha response. FM subtype macrophages are foam macrophages involved in lipid phagocytosis and matrix formation. The corresponding GSEA pathway enrichment results include protein secretion and cholesterol homeostasis. SM subtype macrophages mainly originate from smooth muscle cells (SMCs), with GSEA pathway enrichment results indicating signaling pathways associated with epithelial–mesenchymal transition. Box plots display the scoring results for the four macrophage subtypes, and statistical results are demonstrated in Figure 8F. Notably, the IM and RLM subtypes, which are closely associated with inflammation, exhibited higher scores, while the FM and SM subtypes demonstrated lower scores (all $P < 0.05$). Furthermore, by calculating individual enrichment scores for different macrophage subtypes in two mutually independent study cohorts, we confirmed that cluster B showed higher enrichment of inflammation-associated macrophage subtypes (IM, RLM), whereas non-inflammatory subtypes (FM, SM) were more enriched in cluster A ($P < 0.05$) (Figure 8G and H). Overall, pro-inflammatory macrophages had higher lysogenic cell death scores, and their infiltration and proportion increased with the progression of AS and may be a key factor in inducing differences in AS subtypes.

LCDEGs as Potential Biomarkers of Ischemic Events

The samples were divided into two groups based on the median expression level of each LCDEG. The cumulative hazard rate of ischemic events was performed on all 12 LCDEGs using the “survival” and “survminer” packages. Consequently, the cumulative hazard rate of CYBB was found to be statistically significant at $P < 0.05$ (Figure 9A). The results suggest a correlation between the high expression of CYBB and the occurrence of ischemic events in patients with AS. Figure 9B and C displays the expression and ROC curve of CYBB in pbmc samples after acute coronary syndrome. Notably, there

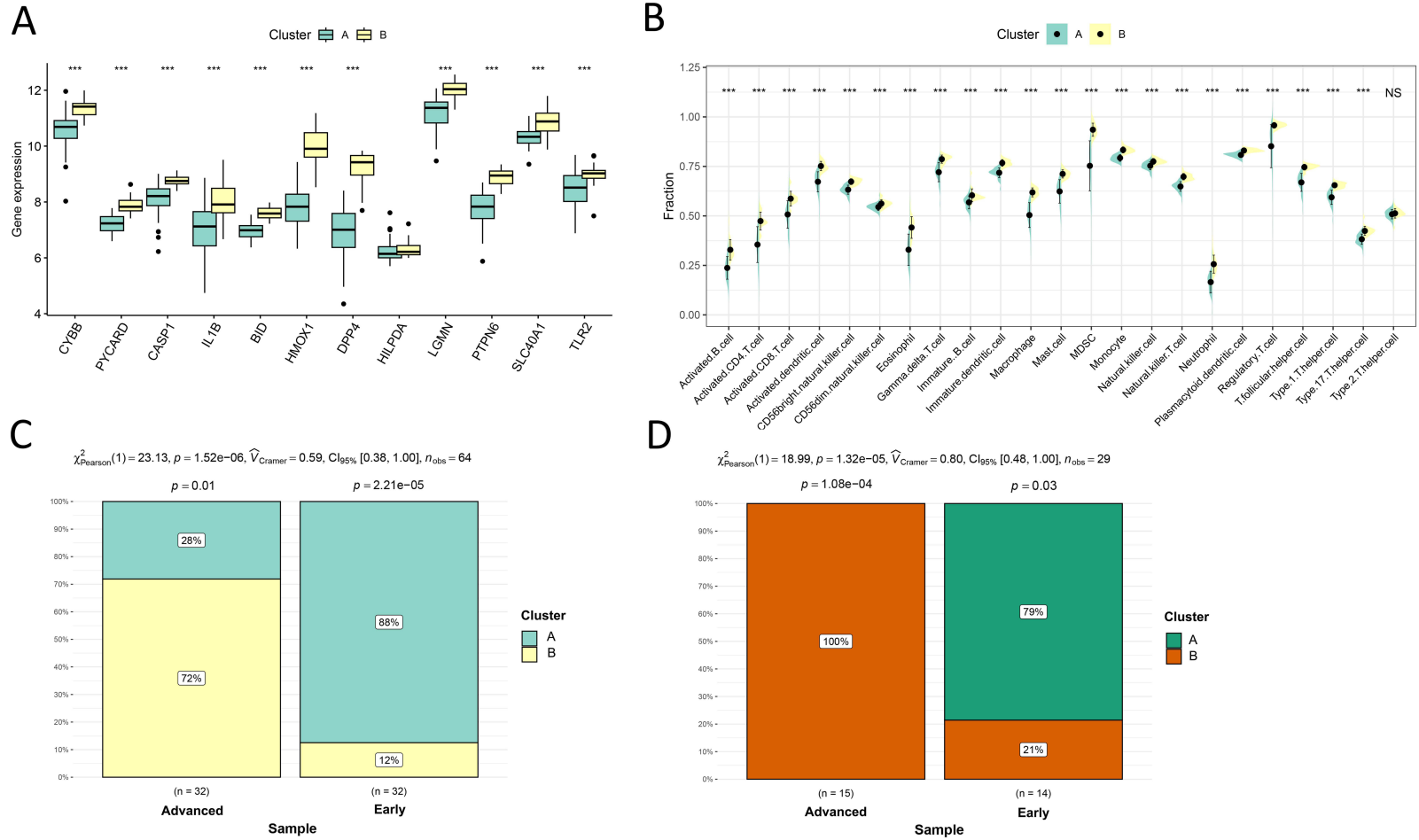


Figure 5 Construction of AS clusters based on LCDEGs in GSE43292 and GSE28829. **(A)** Boxplots illustrating the expression of LCDEGs between the two clusters. **(B)** Differences in the abundance of different immune cells between the two clusters. **(C)** Chi-square tests comparing the association of subtype samples with early and advanced samples. **(D)** Chi-square tests comparing the association of subtype samples with early and advanced samples. * indicates $P < 0.05$ compared to cluster A, ** indicates $P < 0.01$ compared to cluster A, and *** indicates $P < 0.001$ compared to cluster A.

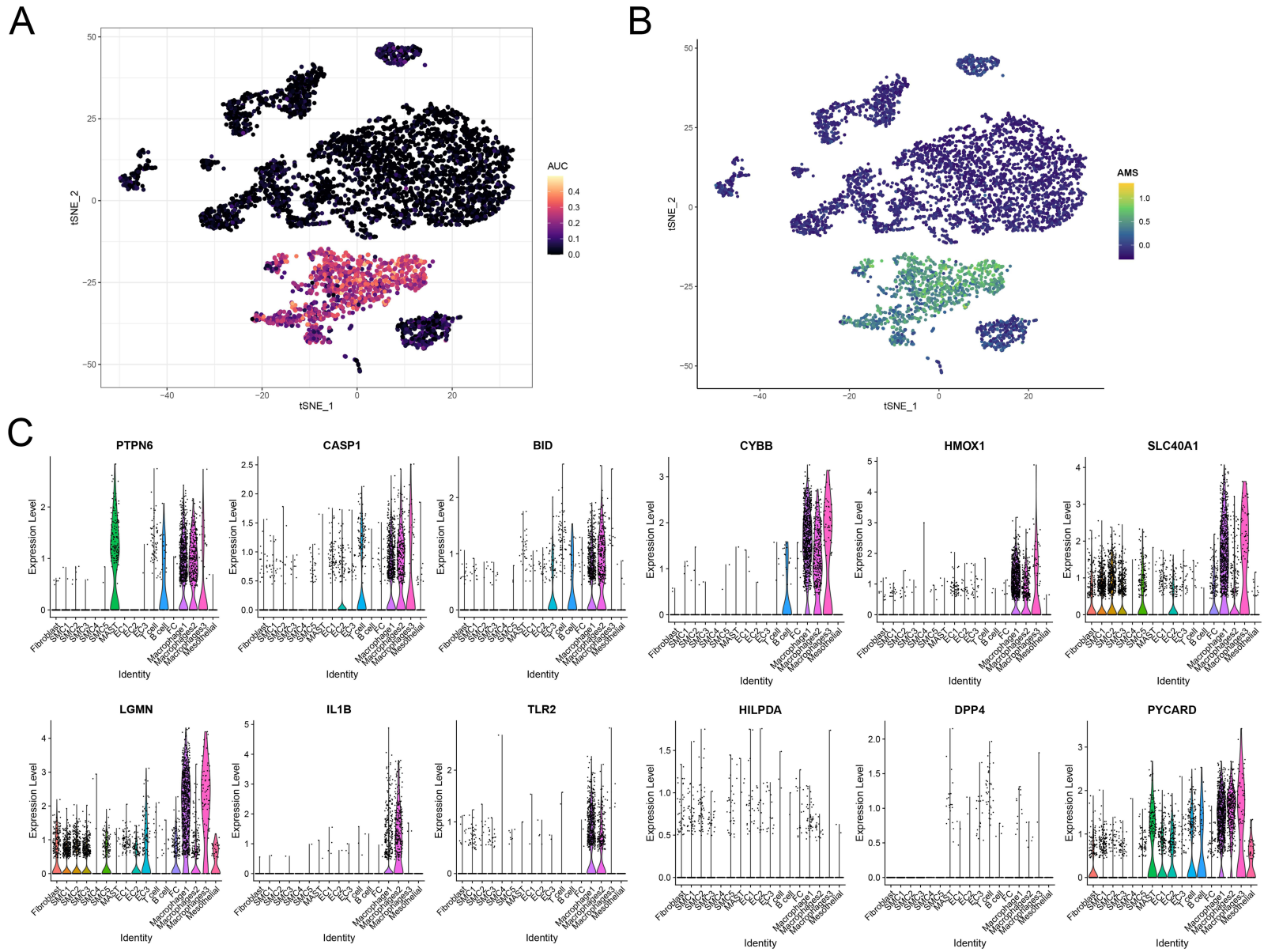
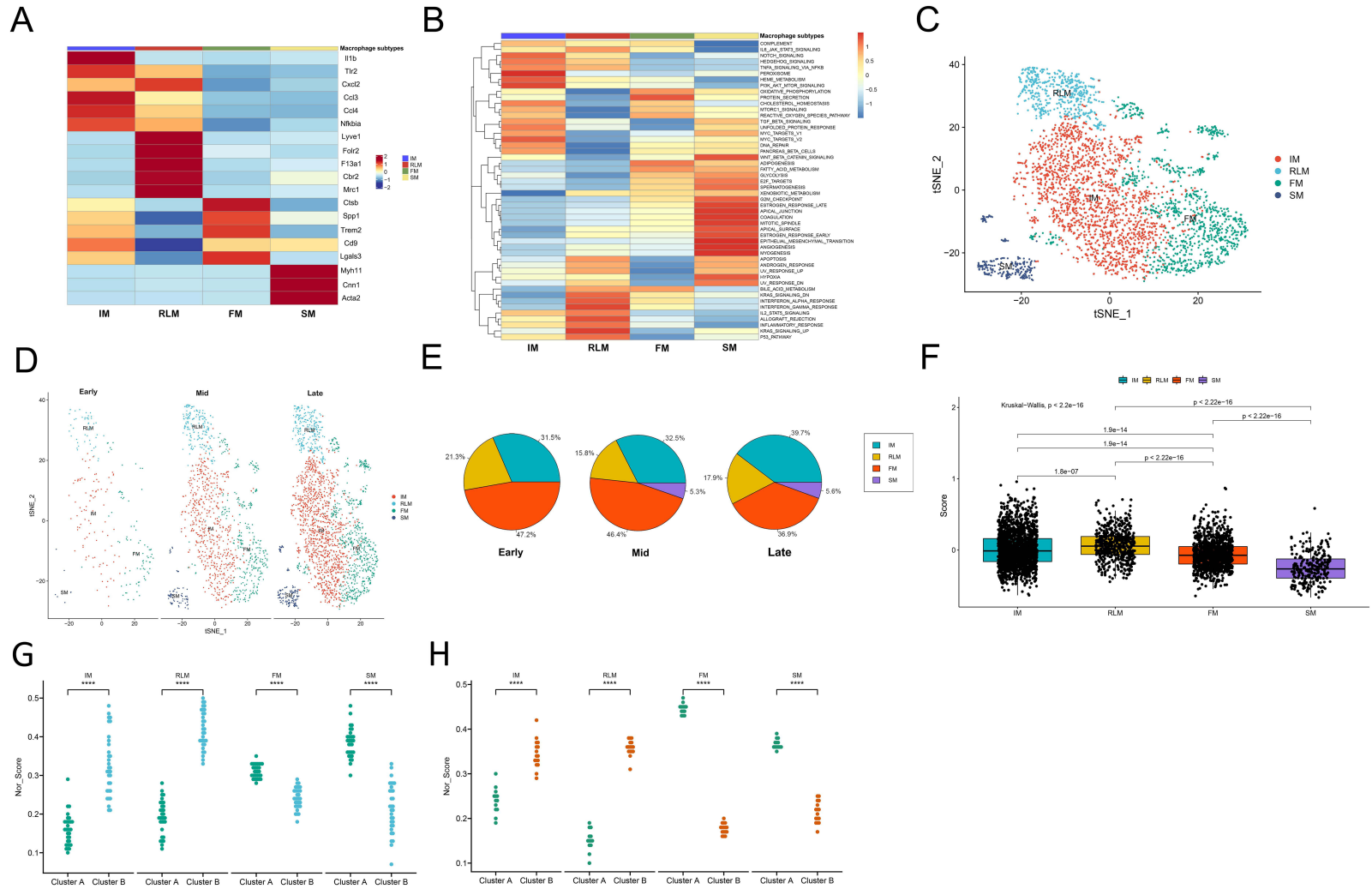


Figure 7 Assessment of LCDEGs activity. **(A)** t-SNE plot of scores obtained using the AUCCell method. **(B)** t-SNE plot of scores obtained using the AddModuleScore method. **(C)** Violin plot of LCDEGs expression in different cell types.



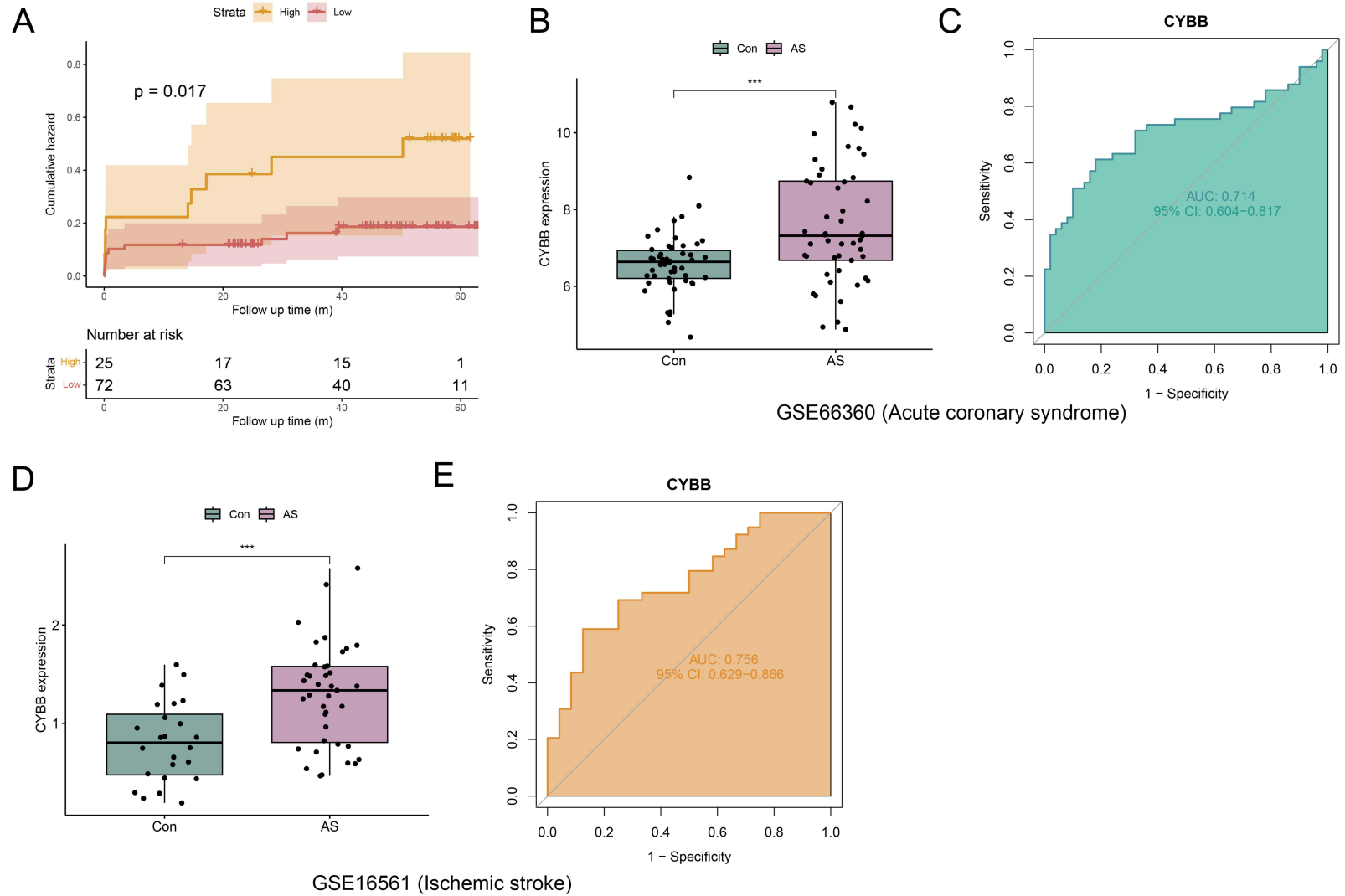


Figure 9 CYBB as a biomarker for ischemic events. **(A)** Comparison of AS patients with high and low expression of CYBB gene in the GSE21545. **(B)** Differences in the expression of CYBB in the GSE66360. **(C)** ROC curves evaluated the diagnostic efficacy of CYBB in the GSE66360. **(D)** Differences in the expression of CYBB in the GSE16561. **(E)** ROC curves evaluated the diagnostic efficacy of CYBB in the GSE16561. *** indicates $P < 0.001$ compared to cluster A.

was a significant difference in the expression of CYBB between the samples of acute coronary syndrome and normal samples ($P < 0.001$). Furthermore, the AUC value of 0.756 indicates that CYBB exhibits high diagnostic accuracy for acute coronary syndrome. The expression and ROC curve of CYBB in pbmc samples of ischemic stroke are presented in [Figure 9C](#) and [D](#). Similarly, there was a significant difference in the expression of CYBB between ischemic stroke and normal samples ($P < 0.001$). Additionally, the AUC value of 0.714 suggests that CYBB has a high diagnostic accuracy for ischemic stroke ([Figure 9E](#)).

Experimental Validation of CYBB as a Potential Biomarker for AS

[Figure 10A](#) illustrates that mice in the early disease group exhibited significantly elevated serum TC, TG, and LDL-C levels, alongside reduced HDL-C levels, in comparison to the normal group. Furthermore, these lipid profile alterations were more pronounced in the late disease group relative to the early disease group ($P < 0.001$), indicating a progressive exacerbation of dyslipidemia with the advancement of the disease. [Figure 10B](#) demonstrates that serum levels of CYBB, IL-1 β , TNF- α , and IL-6 were significantly elevated in the early disease group compared to the normal group, and these levels were further increased in the late disease group compared to the early disease group ($P < 0.001$). These findings demonstrate that CYBB levels were elevated with the advancement of AS, accompanied by an increase in inflammatory factors. As illustrated in [Figure 10C](#) and [D](#), histological examination via HE staining revealed that the early disease group exhibited plaque formation on the vessel wall and partial lumen narrowing compared to the normal group. The advanced disease group showed a further increase in plaque area and more pronounced lumen narrowing relative to the early disease group. As the AS advances, there is a marked intensification in the immunohistochemical staining of CYBB, which is particularly pronounced in the surface region of the plaque on the luminal side of the vessel ([Figure 10E](#)). [Figure 10F](#) illustrates that the protein expression level of CYBB is significantly elevated in the early disease cohort compared to the normal cohort ($P < 0.01$), and is further increased in the late disease cohort relative to the early disease cohort ($P < 0.05$). [Figure 10G](#) and [H](#) show that the plaque area and plaque area/luminal cross-sectional area (%) of the early plaque group were larger than that of the normal group, and the plaque area of the advanced plaque group was larger than that of the early group ($P < 0.001$).

LCDG-Based Drug Prediction

Based on the analysis of the 12 LCDGEs, potential therapeutic drugs were identified using the gene-drug interaction method available in the DGIdb database ([Supplementary Table S4](#)). A total of 99 drugs were found to target seven specific genes. Notably, CASP1, IL1B, and DPP4 exhibited a relatively high number of targeted drugs and are considered potential therapeutic targets for AS. Among these, chrysin, apigenin, and DPP4 inhibitors emerged as potential therapeutic agents. Notably, several of these agents have shown clinical efficacy in the prevention of ischemic events.

Discussion

On the activation of the LCD process, a cell may experience plasma membrane rupture, leading to the release of large molecular inflammatory substances known as DAMPs, as well as the formation of oligomeric pores on the cell membrane, resulting in the release of small molecular inflammatory substances such as IL1B.^{50,51} Increasing evidence from rigorous studies suggests that LCD plays a driving role in the development of AS. For example, a study by Soehnlein et al demonstrated that the interaction between histone H4 and SMCs promotes LCD, contributing to plaque instability, and neutralizing this interaction can prevent SMC LCD and improve AS.⁵² Another study revealed that macrophages can promote the progression of AS by undergoing a RIPK3-MLKL pathway-dependent LCD process, which is significantly reversed by RIPK3 inhibition.⁵³ Consequently, understanding the role of LCD in AS is of paramount importance for the development of future prevention and treatment strategies. Nonetheless, the LCD pathway implicated in AS is intricate and may encompass multiple cell death processes. Thus, identifying the key players of LCD in AS is crucial. Fortunately, bioinformatics methods, when combined with expression profile data, excel at identifying differential genes among thousands of genes, and machine learning algorithms can effectively screen disease-related characteristic genes due to their powerful classification capabilities.⁵⁴ Therefore, in this study, we used an integrative

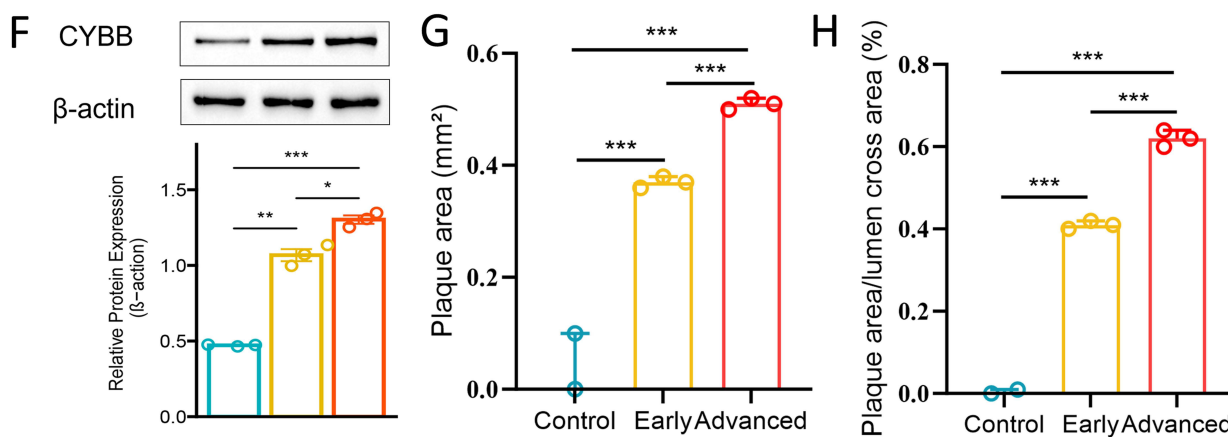


Figure 10 Experimental validation of CYBB as a new biomarker for AS. (A) TG, TC, HDL-C, and LDL-C levels in mouse serum. (B) CYBB, IL-1 β , TNF- α and IL-6 levels in mouse serum. (C) HE staining at the aorta (scale bar: 500 μ m). (D) HE staining at the aorta (scale bar: 100 μ m). (E) Immunohistochemical staining of CYBB (scale bar: 100 μ m). (F) Western blotting results of CYBB in mouse aorta tissue. (G) Plaque area assessed by HE staining of aortic valve cross section. (H) Plaque area/luminal cross-sectional area assessed by HE staining of aortic valve cross section. * indicates $P < 0.05$, ** indicates $P < 0.01$, and *** indicates $P < 0.001$.

This study discovered several key findings, including the following: (1) The WGCNA revealed that the genes in the most relevant modules of AS primarily influence the immunity pathway, specifically the chemokine signaling pathway and the T cell receptor signaling pathway, as well as the cell death pathway, particularly the NOD-like receptor signaling pathway. (2) By performing differential expression analysis, we identified 12 AS-related LCDEGs and confirmed their distinct molecular mechanisms that regulate cell death and immunity using functional enrichment analysis. (3) Eight characteristic genes associated with AS (CASP1, CYBB, DPP4, HILPDA, HMOX1, IL1B, PTPN6, and PYCARD) were identified through the application of two machine learning algorithms, LASSO and RF. (4) An ANN was developed as a diagnostic model using characteristic genes. Our study has demonstrated strong diagnostic capabilities for identifying individuals with and without AS, as well as distinguishing between early and advanced stages of AS. Based on the findings from the first four analyses, it is preliminarily concluded that characteristic genes hold great potential as biomarkers for AS, while the ANN model shows promise as a diagnostic tool. Moreover, these results suggest a connection between immune regulation and the underlying mechanism of LCD in AS. (5) Two molecular subtypes were identified using the consensus clustering method, and their immune landscapes were examined by considering indicators such as immune cell infiltration, cell chemokine genome, and immune checkpoint expression. Based on the immune landscape analysis, these subtypes were categorized as a non-immune subtype, C1, and an immune subtype, C2. Subsequent findings from 5 to 6 revealed that the newly identified AS subtype effectively differentiates between early (low-risk) and advanced (high-risk) atherosclerotic plaques, presenting a potential tool for stratifying plaque risks. (7) Single-cell level analysis demonstrated that LCDEGs exhibited higher expression levels, particularly within macrophages, and were more prevalent in the “inflammatory” subtype of macrophages. These findings strongly indicate a significant connection between macrophages and LCD, with macrophage-based LCD potentially playing a crucial role in AS development. (8) The elevated expression of CYBB, one of the LCDEGs, is correlated with a higher incidence of ischemic events, as revealed by cumulative hazard rate analysis. Additionally, CYBB shows significant expression in blood samples from individuals with acute coronary syndrome and stroke, indicating its potential as a novel prognostic marker for AS. (9) DGIdb was used to investigate potential drugs that target LCDEGs. Our study predicted the effectiveness of select potential drugs for AS treatment, some of which have already been validated in preclinical studies or clinical trials.^{55,56}

By integrating the aforementioned findings, it is essential to explore the following key indicators in depth. Among the eight signature genes identified in our study, CASP1, PYCARD, IL1B, CYBB, DPP4, PTPN6, and HMOX1 were significantly up-regulated in AS, whereas HILPDA was significantly down-regulated. In the subsequent sections, we will describe the functions of the proteins corresponding to these genes in AS, based on findings in previous studies. Caspase 1 (CASP1) is a proteolytic enzyme known to cleave GSDMD, thereby initiating LCD. It also participates in the

proteolysis of IL1B and promotes the secretion of IL1B.^{57,58} PYD and CARD domain containing PYCARD plays a crucial role as a mediator of inflammation by acting as an essential adapter in the assembly of various inflammasomes, such as NLRP1, NLRP2, and NLRP3. Upon recruitment and activation of caspase-1, PYCARD leads to the release of pro-inflammatory cytokines.⁵⁹ Interleukin 1 beta (IL1B) is a highly potent pro-inflammatory cytokine that not only promotes T cell activation but also stimulates cytokine production.⁶⁰ Additionally, the activation of NLRP3 by cholesterol crystals, an endogenous danger signal, is a critical initiating factor in the development of AS.⁶¹ The outcomes of several clinical randomized controlled trials have demonstrated the effectiveness of long-term administration of colchicine in reducing the incidence of cardiovascular events in patients with AS. One important pharmacological mechanism of colchicine is its ability to impede the assembly of NLRP3.^{62,63} Moreover, clinical studies have indicated that canakinumab, a monoclonal antibody that targets IL-1 β , significantly decreases the recurrence rate of cardiovascular events when compared to a placebo in the treatment of AS.⁶⁴ Based on the results of our drug prediction analysis and in conjunction with previous research, it is undeniable that CASP1, PYCARD, and IL1B play crucial roles in the occurrence and development of AS. Therefore, pharmacological inhibition of these proteins may present a promising anti-inflammatory treatment strategy for AS. In this study, the potential of CYBB as an AS marker was further confirmed by joint experimental verification using multiple bioinformatics methods. CYBB, alias NADPH oxidase 2 (NOX2), is an important bactericidal oxidase of phagocytes that promotes the production of reactive oxygen species and is associated with oxidative stress levels involved in AS, platelet aggregation, myocardial infarction, and stroke.⁶⁵ Previous studies have reported that atorvastatin inhibits platelet NOX2 in a dose-dependent manner, leading to reduced levels of platelet isoprostanes and thromboxane A(2), thereby effectively reducing both oxidative stress and platelet activation.⁶⁶ Dipeptidyl peptidase (DPP4), also referred to as CD26, undergoes cleavage from the membrane and subsequent release into the circulation.⁶⁷ The use of DPP4 inhibitors in the treatment of type 2 diabetes is considered to significantly prevent diabetes-induced cardiovascular events, including AS.⁶⁸ Protein tyrosine phosphatase non-receptor type 6 (PTPN6) is a signaling molecule that plays a crucial role in the vitamin D3 (VitD3)-mediated attenuation of AS by inhibiting autophagy and preventing the formation of macrophage foam cells.⁶⁹ Heme oxygenase 1 (HMOX1) is recognized as a crucial regulator in the process of ferroptosis. Its overexpression promotes ferroptotic oxidative stress, thereby contributing to the progression of AS.⁷⁰ Hypoxia-induced lipid droplet association (HILPDA) is linked to the buildup of triglycerides in lipid droplets within adipocytes. According to reports, the upregulation of the HILPDA protein suppresses lipolysis, decreases the production of pro-inflammatory substances like IL-6, and stimulates lipopolysaccharide-mediated macrophage activation. Nonetheless, the role of HILPDA in AS remains uncertain and requires additional investigation.⁷¹ In summary, it is evident that LCDGEs are closely associated with the onset and progression of AS, with macrophages playing an essential role in this process.

Recent advances in molecular profiling techniques have significantly enhanced the capability to observe the transcriptome of patients affected by diseases.⁷² Machine learning models can be utilized to detect molecular signatures in expression profiles, thereby facilitating disease prediction.⁷³ Limited studies address the combination of multiple machine learning methods for establishing a predictive model for AS, this study aims to bridge this gap. In this study, we used three specific machine learning algorithms: LASSO, RF, and ANN. LASSO is a regularization method that excels in high-dimensional datasets by selecting a sparse set of significant features, thus reducing overfitting and enhancing interpretability.⁷⁴ RF is a powerful ensemble learning algorithm that can manage non-linear relationships and interactions between variables. Its robustness against overfitting, particularly when dealing with large and noisy datasets, made it an ideal choice for identifying key features associated with AS progression.⁷⁵ Although other machine learning techniques, such as support vector machine, K-nearest neighbors, extreme gradient boosting, gradient boosting, multilayer perceptron, and logistic regression, could have been applied, we opted for RF due to its high accuracy in feature selection and classification tasks involving large datasets. Moreover, ANN was selected for its ability to simulate complex neural interactions, making it highly suitable for capturing non-linear patterns in molecular data, which are often present in diseases like AS.⁷⁶⁻⁷⁸ Typically, non-neoplastic diseases rely on biomarkers for diagnosis, whereas tumors often use multi-gene expression features to prognosticate or determine drug treatment sensitivity due to their heterogeneity.⁷⁹ It is important to emphasize that substantial heterogeneity in molecular expression exists between stable and unstable plaques in AS, which greatly influences the prognosis. Consequently, using the consensus clustering method in this study to

establish new AS subtypes based on the expression traits of multiple genes is deemed both feasible and crucial. Consensus clustering aggregates results from multiple clustering runs, providing a consensus matrix that reflects the probability of sample co-clustering, thus ensuring the reproducibility and reliability of the identified subtypes.⁴⁴ In contrast to other clustering approaches, such as K-means, hierarchical clustering, or spectral clustering, consensus clustering minimizes the risk of instability that can arise from random initialization or the sensitivity to small variations in the data.^{44,80} These advantages are particularly important in high-dimensional datasets, such as transcriptomic profiles, where noise and data variability can heavily influence the clustering outcome. Furthermore, the integration of radiomics and expression profiling is emerging as an advanced approach to characterize diseases and predict outcomes.⁸¹ The incorporation of expression profiling-based analysis in this study may pave the way for future integration of multi-omics data for developing more precise and biologically interpretable models. Patients with distinct expression signatures may experience both beneficial and harmful effects from immunotherapy.⁸² The novel AS subtype could serve as a guide for cancer patients with AS who undergo immunotherapy, as well as for AS patients who are unresponsive to standard treatments and opt for immunotherapy. Adapting treatment approaches based on distinct subtypes can help avoid severe cardiovascular events associated with medications while maximizing the management of AS.

This study underscores the clinical significance of the ANN model and the identified subtypes, highlighting their potential applicability in clinical practice. Notably, the ANN model demonstrated high reliability as a diagnostic tool for AS disease, exhibiting excellent accuracy in distinguishing individuals with and without AS. Consequently, integrating this model into clinical assessments could facilitate more rapid and precise screening of potential AS patients within the population, which is of critical importance for early detection and timely intervention. Conversely, it is crucial to acknowledge that early AS typically lacks overt symptoms, whereas advanced AS is often associated with significant cardiovascular complications. The application of the ANN model in detecting patients with advanced AS is notably effective. Clinically, it is imperative to use this model in conjunction with other screening methods to identify high-risk AS patients. This integrated approach aims to prevent severe ischemic events, thereby decreasing disability and mortality rates and extending the duration of survival. Furthermore, the newly identified AS subtypes demonstrate the capability to effectively distinguish between early-stage (low-risk) and advanced-stage (high-risk) atherosclerotic plaques, potentially serving as a valuable tool for future plaque risk stratification. Previous findings indicate that these subtypes are predominantly associated with the high-risk atherosclerotic group, characterized by “inflammatory” features. The application of monoclonal antibodies targeting IL-1 β to inhibit inflammation could represent a promising strategy for the treatment of severe AS.⁶⁴ Consequently, the identification and classification of novel AS subtypes may offer valuable insights into the clinical management of patients with AS.

This study is not without limitations. First, there is heterogeneity across the various datasets and platforms utilized, potentially impacting the reliability of the findings. Second, this study is unavoidably subject to random error and selection bias. Third, despite the significant number of datasets used for analysis and validation, the incorporation of larger datasets could enhance the reliability of the research conclusions. Fourth, the analysis of single-cell sequencing data in this study indicates that LCDGEs may play a significant role in the development of AS, primarily in association with macrophages. Fifth, the reliance on publicly available GEO datasets may introduce variability due to differences in experimental conditions. Additionally, the sample sizes of some datasets, such as GSE57691, were relatively small, which may affect the generalizability of our findings. Notably, the high AUC values observed in the training set (eg, 97.3% for GSE100927) compared to lower AUCs in test sets (eg, 78.9% for GSE57691) suggest a potential risk of overfitting in our machine learning models. To mitigate this, we used LASSO regularization with 10-fold cross-validation to select an optimal lambda parameter and optimized RF hyperparameters (eg, mtry) through grid search to balance model complexity. However, the limited sample size of test sets may still constrain model performance. Sixth, the analysis of CYBB as a predictor of clinical outcomes lacks comprehensive mechanistic insights, and predictive drug discovery based on LCD genes is only speculative due to the lack of experimental validation. In addition, the direct mechanistic contribution of LCD genes to plaque rupture or thrombosis remains unclear. More in-depth studies, including clinical studies and in vivo experiments using animal models, are needed to verify the prognostic value of CYBB, confirm the therapeutic potential of predictive drugs, and clarify the functional role of LCD genes in the pathogenesis of AS. Seventh, using an unadjusted $P < 0.05$ threshold in our GO/KEGG enrichment analysis and DEG screening may

increase the risk of false positives. Future studies should apply multiple testing corrections, such as Benjamini-Hochberg FDR, and independently validate enriched pathways using experimental validation or multi-omics approaches.

Eighth, the single-cell cluster annotations obtained by SingleR lack experimental validation and may be subject to misannotation bias. Although SingleR can automatically identify cell types, marker gene-based methods are more accurate in defining cell types. Future work should combine marker gene validation or experimental techniques, such as flow cytometry or immunohistochemistry, to confirm cell identity. Ninth, LCD in this study specifically refers to lytic cell death. LCD has also been considered as lysosomal cell death in previous studies. Lysosomal cell death plays a critical role in the progression of atherosclerosis by regulating macrophage function, lipid metabolism, and inflammatory responses. Dysfunction of lysosomes in macrophages impairs the degradation of oxidized low-density lipoprotein (oxLDL), promoting foam cell formation and chronic inflammation within atherosclerotic plaques.^{61,83} In addition, regulated necrosis can also be considered to cover pyroptosis, necroptosis, and ferroptosis in the context of AS in Wim Martinet's review.⁸⁴ Obviously, the interpretation and understanding of the concept of LCD may overlap with other concepts of cell death, which may be unavoidable. Finally, this study emphasized the role of macrophage death in AS, which is also related to the findings on increased expression levels of CASP1 and IL1B in macrophages,⁸⁵ and the conclusions of studies in the early 1990s on macrophage oxLDL and lysosomal membrane permeability⁸⁶ in the context of atherosclerosis, which is conducive to a deeper understanding of previous studies. This study is also expected to deepen the understanding of the clinical relevance of macrophage markers in atherosclerosis.

Conclusion

In summary, our study used bioinformatics combined with machine learning strategies to explore the potential application and mechanism of LCD in AS, providing new insights into its diagnosis and treatment.

Abbreviations

ANN, Artificial neural network; ApoE, Apolipoprotein E; AS, Atherosclerosis; CASP1, Caspase 1; CYBB, Cytochrome B-245 β chain; DAMPs, Damage-associated molecular patterns; DEGs, Differentially expressed genes; DGIdb, Drug-gene interaction database; DPP4, Dipeptidyl peptidase; GEO, Gene expression omnibus; GB, Gradient boosting; GO, Gene ontology; GSVA, Gene set variation analysis; HDL-C, High-density lipoprotein; HILPDA, Hypoxia-induced lipid droplet association; HMOX1, Heme oxygenase 1; IL1B, Interleukin 1 beta; IL6, Interleukin 6; KEGG, Kyoto encyclopedia of genes and genomes; KNN, K-nearest neighbors; LASSO, Least absolute shrinkage and selection operator; LCD, Lytic cell death; LCDEGs, LCD-associated DEGs; LCDG, LCD-related genes; LDL-C, Low-density lipoprotein; LR, Logistic regression; MLP, Multilayer perceptron; MSigDB, Molecular signature database; NOX2, NADPH oxidase 2; PCA, Principal component analysis; PPI, Protein-protein interaction; PTPN6, Protein tyrosine phosphatase non-receptor type 6; RF, Random forest; ROC, Receiver operating characteristic; ssGSEA, Single-sample gene set enrichment analysis; scRNA-seq, Single-cell RNA-sequencing; SMCs, Smooth muscle cells; SVM, Support vector machine; TOM, Topological overlap matrix; TNF- α , Tumor necrosis factor alpha; TC, Total cholesterol; TG, Triglycerides; UMAP, Uniform manifold approximation and projection for dimension reduction; UMIs, Unique molecular identifiers; VitD3, Vitamin D3; WGCNA, Weighted gene co-expression network analysis; XGB, Extreme gradient boosting.

Data Sharing Statement

The public data used here are available in the GEO database (accession number: GSE100927, GSE57691, GSE43292, GSE28829, GSE66360, GSE16561, GSE21545, GSE155512, and GSE155513; <https://www.ncbi.nlm.nih.gov/geo/>). The patients involved in the database have obtained ethical approval. The codes can be obtained from the first author upon reasonable request. Users can download relevant data for free for research and publish relevant articles.

Author Contributions

All authors have made significant contributions to the work of the report, whether in terms of conception, research design, implementation, data acquisition, analysis and interpretation, or in all these aspects. Participated in the drafting,

revision or review of the article; Finally, the upcoming version was reviewed; An agreement has been reached on the journal to which the article will be submitted. And agree to be responsible for all aspects of the work.

Funding

This work was supported by grants from the Dalian Municipal Dengfeng Clinical Medicine Grant Support (No. 2021024).

Disclosure

The authors declare no competing interests in this work.

References

1. Yao J, Liang J, Li H. Screening for key genes in circadian regulation in advanced atherosclerosis: a bioinformatic analysis. *Front Cardiovasc Med.* 2022;9:990757. doi:10.3389/fcvm.2022.990757
2. Libby P. The changing landscape of atherosclerosis. *Nature.* 2021;592(7855):524–533. doi:10.1038/s41586-021-03392-8
3. Guaraldi G, Raggi P. Atherosclerosis in frailty: not frailty in atherosclerosis. *Atherosclerosis.* 2017;266:226–227. doi:10.1016/j.atherosclerosis.2017.09.014
4. NCD Countdown 2030 collaborators. NCD countdown 2030: worldwide trends in non-communicable disease mortality and progress towards sustainable development goal target 3.4. *Lancet.* 2018;392(10152):1072–1088. doi:10.1016/s0140-6736(18)31992-5
5. Robinson JG, Davidson MH. Can we cure atherosclerosis? *Rev Cardiovasc Med.* 2018;19(S1):S20–s24. doi:10.3909/ricm19S1S0003
6. Ibanez B, Fernández-Ortiz A, Fernández-Friera L, García-Lunar I, Andrés V, Fuster V. Progression of Early Subclinical Atherosclerosis (PESA) study: JACC focus seminar 7/8. *J Am Coll Cardiol.* 2021;78(2):156–179. doi:10.1016/j.jacc.2021.05.011
7. Mo S, Wang Y, Yuan X, et al. Identification of common signature genes and pathways underlying the pathogenesis association between nonalcoholic fatty liver disease and atherosclerosis. *Front Cardiovasc Med.* 2023;10:1142296. doi:10.3389/fcvm.2023.1142296
8. Döring Y, Manthey HD, Drechsler M, et al. Auto-antigenic protein-DNA complexes stimulate plasmacytoid dendritic cells to promote atherosclerosis. *Circulation.* 2012;125(13):1673–1683. doi:10.1161/circulationaha.111.046755
9. Gautheron J, Gores GJ, Rodrigues CMP. Lytic cell death in metabolic liver disease. *J Hepatol.* 2020;73(2):394–408. doi:10.1016/j.jhep.2020.04.001
10. Ning X, Wang Y, Jing M, et al. Apoptotic caspases suppress type I interferon production via the cleavage of cGAS, MAVS, and IRF3. *Molecular Cell.* 2019;74(1):19–31.e7. doi:10.1016/j.molcel.2019.02.013
11. Yang F, He Y, Zhai Z, Sun E. Programmed cell death pathways in the pathogenesis of systemic lupus erythematosus. *J Immunol Res.* 2019;2019:3638562. doi:10.1155/2019/3638562
12. Frank D, Vince JE. Pyroptosis versus necroptosis: similarities, differences, and crosstalk. *Cell Death Differ.* 2019;26(1):99–114. doi:10.1038/s41418-018-0212-6
13. Li M, Wang ZW, Fang LJ, Cheng SQ, Wang X, Liu NF. Programmed cell death in atherosclerosis and vascular calcification. *Cell Death Dis.* 2022;13(5):467. doi:10.1038/s41419-022-04923-5
14. Zhe-Wei S, Li-Sha G, Yue-Chun L. The role of necroptosis in cardiovascular disease. *Front Pharmacol.* 2018;9:721. doi:10.3389/fphar.2018.00721
15. Jinson S, Zhang Z, Lancaster GI, Murphy AJ, Morgan PK. Iron, lipid peroxidation, and ferroptosis play pathogenic roles in atherosclerosis. *Cardiovasc Res.* 2025;121(1):44–61. doi:10.1093/cvr/cvae270
16. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12(5):453–457. doi:10.1038/nmeth.3337
17. Wirka RC, Pjanic M, Quertermous T. Advances in transcriptomics: investigating cardiovascular disease at unprecedented resolution. *Circ Res.* 2018;122(9):1200–1220. doi:10.1161/circresaha.117.310910
18. Paik DT, Cho S, Tian L, Chang HY, Wu JC. Single-cell RNA sequencing in cardiovascular development, disease and medicine. *Nat Rev Cardiol.* 2020;17(8):457–473. doi:10.1038/s41569-020-0359-y
19. Wang A, Li Z, Sun Z, Liu Y, Zhang D, Ma X. Potential mechanisms between HF and COPD: new insights from bioinformatics. *Curr Prob Cardiol.* 2023;48(3):101539. doi:10.1016/j.cpcardiol.2022.101539
20. Liu Z, Liu L, Weng S, et al. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. *Nat Commun.* 2022;13(1):816. doi:10.1038/s41467-022-28421-6
21. Li H, Xu Y, Wang A, Zhao C, Zheng M, Xiang C. Integrative bioinformatics and machine learning approach unveils potential biomarkers linking coronary atherosclerosis and fatty acid metabolism-associated gene. *J Cardiothorac Surg.* 2025;20(1):70. doi:10.1186/s13019-024-03199-4
22. Fu XW, Song CQ. Identification and validation of pyroptosis-related gene signature to predict prognosis and reveal immune infiltration in hepatocellular carcinoma. *Front Cell Develop Biol.* 2021;9:748039. doi:10.3389/fcell.2021.748039
23. Liu F, Wei T, Liu L, et al. Role of necroptosis and immune infiltration in human Stanford type A aortic dissection: novel insights from bioinformatics analyses. *Oxid Med Cell Longev.* 2022;2022:6184802. doi:10.1155/2022/6184802
24. Yuan K, Hu D, Mo X, et al. Novel diagnostic biomarkers of oxidative stress, immune-infiltration characteristics and experimental validation of SERPINE1 in colon cancer. *Discover Oncol.* 2023;14(1):206. doi:10.1007/s12672-023-00833-w
25. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47. doi:10.1093/nar/gkv007
26. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 2008;9:559. doi:10.1186/1471-2105-9-559
27. Liang L, Sun J, Teng T, et al. Expression profile of inflammation response genes and potential regulatory mechanisms in dilated cardiomyopathy. *Oxid Med Cell Longev.* 2022;2022:1051652. doi:10.1155/2022/1051652
28. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic.* 2012;16(5):284–287. doi:10.1089/omi.2011.0118

29. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* 2013;14:7. doi:10.1186/1471-2105-14-7
30. Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 2017;45(D1):D362–d368. doi:10.1093/nar/gkw937
31. Franz M, Rodriguez H, Lopes C, et al. GeneMANIA update 2018. *Nucleic Acids Res.* 2018;46(W1):W60–w64. doi:10.1093/nar/gky311
32. Li Y, Lu F, Yin Y. Applying logistic LASSO regression for the diagnosis of atypical Crohn's disease. *Sci Rep.* 2022;12(1):11340. doi:10.1038/s41598-022-15609-5
33. Li Y, Qi D, Zhu B, Ye X. Analysis of m6A RNA methylation-related genes in liver hepatocellular carcinoma and their correlation with survival. *Int J Mol Sci.* 2021;22(3). doi:10.3390/ijms22031474
34. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell.* 2019;177(7):1888–1902.e21. doi:10.1016/j.cell.2019.05.031
35. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods.* 2019;16(12):1289–1296. doi:10.1038/s41592-019-0619-0
36. Vallejo J, Cochain C, Zerneck A, Ley K. Heterogeneity of immune cells in human atherosclerosis revealed by scRNA-Seq. *Cardiovasc Res.* 2021;117(13):2537–2543. doi:10.1093/cvr/cvab260
37. Pan H, Xue C, Auerbach BJ, et al. Single-cell genomics reveals a novel cell state during smooth muscle cell phenotypic switching and potential therapeutic targets for atherosclerosis in mouse and human. *Circulation.* 2020;142(21):2060–2075. doi:10.1161/circulationaha.120.048378
38. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods.* 2017;14(11):1083–1086. doi:10.1038/nmeth.4463
39. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* 2018;1711:243–259. doi:10.1007/978-1-4939-7493-1_12
40. Lu J, Chen Y, Zhang X, Guo J, Xu K, Li L. A novel prognostic model based on single-cell RNA sequencing data for hepatocellular carcinoma. *Can Cell Inter.* 2022;22(1):38. doi:10.1186/s12935-022-02469-2
41. Zhang Q, Liu Y, Wang X, Zhang C, Hou M, Liu Y. Integration of single-cell RNA sequencing and bulk RNA transcriptome sequencing reveals a heterogeneous immune landscape and pivotal cell subpopulations associated with colorectal cancer prognosis. *Front Immunol.* 2023;14:1184167. doi:10.3389/fimmu.2023.1184167
42. Hayes DN, Monti S, Parmigiani G, et al. Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. *J Clin Oncol.* 2006;24(31):5079–5090. doi:10.1200/jco.2005.05.1748
43. Garber ME, Troyanskaya OG, Schluens K, et al. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci USA.* 2001;98(24):13784–13789. doi:10.1073/pnas.241500798
44. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics.* 2010;26(12):1572–1573. doi:10.1093/bioinformatics/btq170
45. Yu G, Bao J, Zhan M, et al. Comprehensive analysis of m5C methylation regulatory genes and tumor microenvironment in prostate cancer. *Front Immunol.* 2022;13:914577. doi:10.3389/fimmu.2022.914577
46. Barbie DA, Tamayo P, Boehm JS, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature.* 2009;462(7269):108–112. doi:10.1038/nature08460
47. Le T, Aronow RA, Kirshtein A, Shahriyari L. A review of digital cytometry methods: estimating the relative abundance of cell types in a bulk of cells. *Briefings Bioinf.* 2021;22(4). doi:10.1093/bib/bbaa219
48. Cotto KC, Wagner AH, Feng YY, et al. DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. *Nucleic Acids Res.* 2018;46(D1):D1068–d1073. doi:10.1093/nar/gkx1143
49. National Research Council Committee for the Update of the Guide for the Care and Use of Laboratory Animals. The National Academies Collection: reports funded by National Institutes of Health. In: *Guide for the Care and Use of Laboratory Animals*. National Academies Press (US) Copyright © 2011, National Academy of Sciences; 2011.
50. Newton K, Dixit VM, Kayagaki N. Dying cells fan the flames of inflammation. *Science.* 2021;374(6571):1076–1080. doi:10.1126/science.abi5934
51. Broz P, Pelegrin P, Shao F. The gasdermins, a protein family executing cell death and inflammation. *Nat Rev Immunol.* 2020;20(3):143–157. doi:10.1038/s41577-019-0228-2
52. Silvestre-Roig C, Braster Q, Wichapong K, et al. Externalized histone H4 orchestrates chronic inflammation by inducing lytic cell death. *Nature.* 2019;569(7755):236–240. doi:10.1038/s41586-019-1167-6
53. Karunakaran D, Nguyen MA, Geoffrion M, et al. RIPK1 expression associates with inflammation in early atherosclerosis in humans and can be therapeutically silenced to reduce NF-κB activation and atherogenesis in mice. *Circulation.* 2021;143(2):163–177. doi:10.1161/circulationaha.118.038379
54. Mohapatra SK, Krishnan A. Microarray data analysis. *Methods Mol Biol.* 2011;678:27–43. doi:10.1007/978-1-60761-682-5_3
55. Basu A, Das AS, Majumder M, Mukhopadhyay R. Antiatherogenic roles of dietary flavonoids chrysin, quercetin, and luteolin. *J Cardiovasc Pharmacol.* 2016;68(1):89–96. doi:10.1097/fjc.0000000000000380
56. Yuvaraj S, Sasikumari S, Puhari SSM, et al. Chrysin reduces hypercholesterolemia-mediated atherosclerosis through modulating oxidative stress, microflora, and apoptosis in experimental rats. *J Food Biochem.* 2022;46(11):e14349. doi:10.1111/jfbc.14349
57. Kayagaki N, Warming S, Lamkanfi M, et al. Non-canonical inflammasome activation targets caspase-11. *Nature.* 2011;479(7371):117–121. doi:10.1038/nature10558
58. Gao L, Dong X, Gong W, et al. Acinar cell NLRP3 inflammasome and gasdermin D (GSDMD) activation mediates pyroptosis and systemic inflammation in acute pancreatitis. *Br J Pharmacol.* 2021;178(17):3533–3552. doi:10.1111/bph.15499
59. Yan YQ, Fang Y, Zheng R, Pu JL, Zhang BR. NLRP3 Inflammasomes in Parkinson's disease and their regulation by Parkin. *Neuroscience.* 2020;446:323–334. doi:10.1016/j.neuroscience.2020.08.004
60. Lopez-Castejon G, Brough D. Understanding the mechanism of IL-1β secretion. *Cytokine Growth Factor Rev.* 2011;22(4):189–195. doi:10.1016/j.cytogfr.2011.10.001
61. Duewell P, Kono H, Rayner KJ, et al. NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals. *Nature.* 2010;464(7293):1357–1361. doi:10.1038/nature08938

62. Tardif JC, Kouz S, Waters DD, et al. Efficacy and safety of low-dose colchicine after myocardial infarction. *New Engl J Med.* 2019;381(26):2497–2505. doi:10.1056/NEJMoa1912388
63. Deftereos SG, Beerkens FJ, Shah B, et al. Colchicine in cardiovascular disease: in-depth review. *Circulation.* 2022;145(1):61–78. doi:10.1161/circulationaha.121.056171
64. Ridker PM, Everett BM, Thuren T, et al. Antiinflammatory therapy with Canakinumab for Atherosclerotic Disease. *New Engl J Med.* 2017;377(12):1119–1131. doi:10.1056/NEJMoa1707914
65. Forte M, Nocella C, De Falco E, et al. The pathophysiological role of NOX2 in hypertension and organ damage. *High Blood Pressure Cardiovasc Prevent.* 2016;23(4):355–364. doi:10.1007/s40292-016-0175-y
66. Pignatelli P, Carnevale R, Pastori D, et al. Immediate antioxidant and antiplatelet effect of atorvastatin via inhibition of Nox2. *Circulation.* 2012;126(1):92–103. doi:10.1161/circulationaha.112.095554
67. Röhrborn D, Wronkowitz N, Eckel J. DPP4 in Diabetes. *Front Immunol.* 2015;6:386. doi:10.3389/fimmu.2015.00386
68. Liu H, Guo L, Xing J, et al. The protective role of DPP4 inhibitors in atherosclerosis. *Eur J Pharmacol.* 2020;875:173037. doi:10.1016/j.ejphar.2020.173037
69. Kumar S, Nanduri R, Bhagyaraj E, et al. Vitamin D3-VDR-PTPN6 axis mediated autophagy contributes to the inhibition of macrophage foam cell formation. *Autophagy.* 2021;17(9):2273–2289. doi:10.1080/15548627.2020.1822088
70. Wu D, Hu Q, Wang Y, Jin M, Tao Z, Wan J. Identification of HMOX1 as a critical ferroptosis-related gene in atherosclerosis. *Front Cardiovasc Med.* 2022;9:833642. doi:10.3389/fcvm.2022.833642
71. van Dierendonck X, Vrieling F, Smeehuijzen L, et al. Triglyceride breakdown from lipid droplets regulates the inflammatory response in macrophages. *Proc Natl Acad Sci USA.* 2022;119(12):e2114739119. doi:10.1073/pnas.2114739119
72. Elmarakeby HA, Hwang J, Arafeh R, et al. Biologically informed deep neural network for prostate cancer discovery. *Nature.* 2021;598(7880):348–352. doi:10.1038/s41586-021-03922-4
73. Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B. Definitions, methods, and applications in interpretable machine learning. *Proc Natl Acad Sci USA.* 2019;116(44):22071–22080. doi:10.1073/pnas.1900654116
74. Hu JY, Wang Y, Tong XM, Yang T. When to consider logistic LASSO regression in multivariate analysis? *Eur J Surg Oncol.* 2021;47(8):2206. doi:10.1016/j.ejso.2021.04.011
75. Roy MH, Larocque D. Prediction intervals with random forests. *Stat Methods Med Res.* 2020;29(1):205–229. doi:10.1177/0962280219829885
76. Kriegeskorte N, Golan T. Neural network models and deep learning. *Curr Biol.* 2019;29(7):R231–r236. doi:10.1016/j.cub.2019.02.034
77. Sahu A, Mishra J, Kushwaha N. Artificial Intelligence (AI) in Drugs and Pharmaceuticals. *Comb Chem High Throughput Screening.* 2022;25(11):1818–1837. doi:10.2174/1386207325666211207153943
78. Wu W, Wang J, Cheng M, Li Z. Convergence analysis of online gradient method for BP neural networks. *Neural Networks.* 2011;24(1):91–98. doi:10.1016/j.neunet.2010.09.007
79. Koncina E, Haan S, Rauh S, Letellier E. Prognostic and predictive molecular biomarkers for colorectal cancer: updates and challenges. *Cancers.* 2020;12(2). doi:10.3390/cancers12020319
80. Zhang JZ, Wang C. A comparative study of clustering methods on gene expression data for lung cancer prognosis. *BMC Res Notes.* 2023;16(1):319. doi:10.1186/s13104-023-06604-8
81. Binczyk F, Prazuch W, Bozek P, Polanska J. Radiomics and artificial intelligence in lung cancer screening. *Transl Lung Cancer Res.* 2021;10(2):1186–1199. doi:10.21037/tlcr-20-708
82. Vuong JT, Stein-Merlob AF, Nayeri A, Sallam T, Neilan TG, Yang EH. Immune checkpoint therapies and atherosclerosis: mechanisms and clinical implications: JACC State-of-the-Art Review. *J Am Coll Cardiol.* 2022;79(6):577–593. doi:10.1016/j.jacc.2021.11.048
83. Tabas I. Macrophage death and defective inflammation resolution in atherosclerosis. *Nat Rev Immunol.* 2010;10(1):36–46. doi:10.1038/nri2675
84. Puylaert P, Zurek M, Rayner KJ, De Meyer GRY, Martinet W. Regulated Necrosis in Atherosclerosis. *Arteriosclerosis Thrombosis Vasc Biol.* 2022;42(11):1283–1306. doi:10.1161/atvbaha.122.318177
85. Jiang X, Wang F, Wang Y, et al. Inflammasome-driven Interleukin-1 α and Interleukin-1 β production in atherosclerotic plaques relates to hyperlipidemia and plaque complexity. *JACC.* 2019;4(3):304–317. doi:10.1016/j.jacbs.2019.02.007
86. Yuan XM, Li W, Olsson AG, Brunk UT. The toxicity to macrophages of oxidized low-density lipoprotein is mediated through lysosomal damage. *Atherosclerosis.* 1997;133(2):153–161. doi:10.1016/s0021-9150(97)00094-4