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 Bioinformatics and Experimental Validation Identifies Macrophage Toll Like Receptor 1 as a Therapeutic Target in Carotid Atherosclerosis

Meng Zhang
The Second Affiliated Hospital, Hengyang Medical School, University of South China
Junhui Xiao
The Second Affiliated Hospital, Hengyang Medical School, University of South China
Kai Hou
The Second Affiliated Hospital, Hengyang Medical School, University of South China
Shili Xu
xushili2022@126.com
The Second Affiliated Hospital, Hengyang Medical School, University of South China

Article

Keywords: Carotid atherosclerosis, Integrative Bioinformatics, Toll Like Receptor 1, Macrophage, Inflammatory

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11374110/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Integrative Bioinformatics and Experimental Validation Identifies Macrophage Toll Like Receptor 1 as a Therapeutic Target in Carotid Atherosclerosis

Running Title: TLR1 and Carotid Atherosclerosis Risk through Macrophages

Meng Zhang[1†], Junhui Xiao[1†], Kai Hou[2], Shili Xu[3*]

[1]Department of Pharmacy, The Second Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, 421001, Hunan, China

[2]The Centre of Clinical Research of The Second Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, 421001, Hunan, China

[3]Hunan Provincial Key Laboratory of Basic and Clinical Pharmacological Research of Gastrointestinal Cancer, Institute of Pharmacy and Pharmacology, the Second Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, 421001, Hunan, China

[†]These authors contributed equally in this work.

*Correspondence address: E-mail: xushili2022@126.com (S.X.).

Abstract

Carotid atherosclerosis remains a major contributor to cerebrovascular events, with macrophage-driven inflammation playing a pivotal role in disease progression. This study employed an integrative multi-omics approach combining bulk RNA sequencing and weighted gene correlation network analysis to identify critical differential expression genes in atherogenesis. Machine learning algorithms (Least Absolute Shrinkage and Selection Operator, random forest and artificial neural network) prioritized Toll Like Receptor 1 as the most significant risk gene. Single-cell sequencing of carotid plaques demonstrated TLR1 was predominantly distributed in macrophages. Experimental validation in mouse bone marrow primary macrophages confirmed that coriander inhibits TLR1-mediated macrophage inflammatory response in carotid plaques. Immune infiltration analysis showed that TLR1 was closely related to neutrophil and macrophage activation. Our findings establish TLR1 as a macrophage-specific diagnostic biomarker and suggest that ingestion of coriander supplementation inhibits the inflammatory amplification circuit in TLR1-mediated advanced atherosclerosis.

Keywords

Carotid atherosclerosis, Integrative Bioinformatics, Toll Like Receptor 1, Macrophage, Inflammatory

1 Introduction

Carotid atherosclerosis, a pathological thickening of the arterial wall at the carotid bifurcation, is a leading cause of ischemic stroke due to plaque rupture and thromboembolism[1, 2]. Characterized by lipid accumulation, chronic inflammation, and macrophage infiltration, unstable plaques exhibit features such as thin fibrous caps and intraplaque hemorrhage. Epidemiologically, it affects approximately 20% of adults over 60 years globally, contributing to 15-20% of acute cerebrovascular events[3]. Current therapeutic strategies, including statins and surgical interventions like endarterectomy, primarily target lipid reduction and mechanical stabilization but fail to address the persistent inflammatory microenvironment driven by macrophages[4-6]. Residual inflammatory risk, as indicated by elevated hs-CRP levels, remains a critical challenge in 35% of patients, highlighting the urgent need for novel biomarkers and therapies targeting macrophage-mediated pathways.

Recent advances in single-cell transcriptomics have revealed macrophage heterogeneity as a key determinant of plaque vulnerability, with specific subsets like TREM2+ macrophages linked to rupture-prone phenotypes[7-10]. While Toll-like receptor (TLR) signaling, particularly TLR4, has been extensively studied in atherosclerosis, the role of TLR1 remains controversial—showing pro-inflammatory effects in cardiovascular diseases but the mechanism remaining unclear[11-16]. Network pharmacology has emerged as a powerful tool to identify plant-derived compounds with therapeutic potential, such as berberine's modulation of cholesterol efflux[17-20]. However, existing studies lack integration of carotid-specific multi-omics data, spatial resolution of macrophage subpopulations, and machine learning-driven biomarker discovery, leaving TLR1's role in carotid atherosclerosis underexplored.

This study employs an integrative approach combining carotid plaque-specific transcriptomics, network pharmacology, and artificial intelligence to unravel macrophage-specific drivers of plaque instability. Machine learning algorithms, including the first application of artificial neural networks (ANN) in carotid plaque analysis, validated TLR1's predictive value for plaque vulnerability. Experimental validation further explored coriander-derived compounds' ability to modulate TLR1-mediated inflammatory cascades, bridging computational predictions with biological relevance.

Our findings identify TLR1 as a macrophage-specific diagnostic biomarker. Mechanistically, we link TLR1 activation to macrophage crosstalk through an inflammatory response, which is a key pathological basis for carotid plaque rupture. The identification of coriander as a potent TLR1 inhibitor by inverse network pharmacology, coupled with its efficacy in inhibition of NF-κB phosphorylation in primary macrophages, provides a dietary intervention strategy for high-risk patients. This work advances precision medicine for atherosclerosis by unraveling carotid artery-specific molecular networks and providing phytochemistry-based therapeutic pathways to alleviate inflammation-driven plaque progression.

2 Methods

2.1 Data provenance and preprocessing

The data information included in the analysis in this study is shown in Table 1. Microarray and high-throughput sequencing data were obtained from the GEO database of the National Center for Biotechnology Information (URL: https://www.ncbi.nlm.nih.gov/geo/), and the dataset information is described in Table 1. All of these figures are from the European population.

Table1 Datasets included in the study

Database	Datasets	Year	Population	N (samples)
GEO	GSE28829	2011	European	Case = 16, Control = 13
GEO	GSE41571	2012	European	Case = 5, Control = 6
GEO	GSE43292	2013	European	Case = 32, Control = 32
GEO	GSE100927	2018	European	Case = 69, Control = 35
GEO	GSE253903	2024	European	Case = 6, Control = 6

After deduplication of the raw counting matrix (Limma packet), a log2 (x+1) transformation was performed on the median expression of the repeating gene probe and the raw data without the logarithm to compress the dynamic range and generate a standardized matrix of unique gene identifiers. The normalizeBetweenArrays function was used for inter-quantile normalization to eliminate technical bias between samples and ensure cross-sample expression spectrum comparability. The expression matrix was reconstructed according to the clinical phenotype (early atherosclerosis group and advanced group), and the corrected data file was finally output. The ggplot2 package was used to plot the grouped box plots to systematically evaluate the distribution consistency and normalization effect between samples.

2.2 Weighted gene correlation network analysis

A scale-free topological network was constructed based on the standardized gene expression profile, and the top 25% genes with the highest expression coefficient of variation were screened for subsequent analysis. Average linkage hierarchical clustering was used to eliminate outlier samples based on dynamic shearing algorithm. The soft threshold power was determined by the scale-free topology fitting index (≥0.9) and the average connectivity curve inflection point method, and finally the adjacency

matrix was constructed with β = 8. Gene similarity was quantified by topological overlap matrix, and co-expression modules were divided based on dynamic mixed tree segmentation. The module characteristic genes were extracted by principal component analysis, and the correlation between them and atherosclerosis was determined by Pearson correlation test ($p<0.05$). Gene Significance (GS) is defined as the absolute value of the correlation coefficient between gene expression and phenotype, Module Membership (MM) is the correlation coefficient between gene expression and module characteristic genes, and the core gene screening criteria are: $|MM|\geq0.7$ and $|GS|\geq0.3$。The visualization of key modules is displayed in heat maps and scatter plots.

2.3 Differential expression analysis

A linear model framework was constructed based on standardized gene expression profiles for analysis of differences between groups. Empirical Bayesian contraction (eBayes function) was used to estimate the gene-specific variance, and the threshold for defining significantly different genes was: absolute log-fold change ($|log2FC| \geq 0.585$ (corresponding to 1.5-fold expression difference) and P value < 0.05. The Benjamini-Hochberg method was used to correct the false discovery rate for multiple hypothesis testing, and the hierarchical clustering heat map (pheatmap package) and volcano map (ggplot2 package) were plotted for visualization.

2.4 Enrichment analysis

GO, KEGG, and DO enrichment analyses were performed based on differentially expressed gene sets (clusterProfiler package), and association was assessed using hypergeometric tests. Gene set enrichment analysis (GSEA) was performed using the KEGG gene sets from MSigDB. The significance threshold was set at a P-value of <0.05 and a false discovery rate FDR of <1 (Benjamini-Hochberg correction). Build interactive multi-level axes with the GGH4X package and use bubble maps to illustrate the most significant items, pathways, and diseases.

2.5 Machine learning

Based on the standardized gene expression profile, the minimum absolute contraction and selection operator (LASSO) regression model (glmnet package) and random forest algorithm (Random Forest, randomForest package) were used to screen

the core genes related to atherosclerosis. The sample was used as the observation unit, gene expression was the characteristic variable, and the clinical group was divided into dichotomous response variables. The LASSO regression model was used to construct a penalty logistic regression model, and the regularization parameter λ was optimized through ten-fold cross-verification, and the non-zero gene with λmin correspondence coefficient was selected as the characteristic gene. RF constructs an ensemble model of 500 decision trees, evaluates the performance of the model through OOB error, and dynamically selects the optimal number of trees to reconstruct the model. The Gini index was used to quantify the importance of features, the top 10 key genes were screened and the importance bubble map was drawn by the viridis color scale, the point size and color depth mapped the importance score, the error curve plot showed the convergence trend of OOB error with the number of trees, and the feature ranking results were verified by ten folds for robustness test. Random seeds (seed = 12345) were fixed throughout the analysis to ensure reproducibility of results.

2.6 Artificial neural networks

Based on the binary feature matrix of characteristic genes, a multi-layer feedforward neural network (ANN) was constructed for atherosclerosis classification prediction: the median threshold binarization (the median coding of the expression value ≥ the median was 1 and 0 otherwise) was used for the up-regulated genes, and the down-regulated genes were inversely encoded to generate a binary feature input matrix. The neuralnet package was used to establish a single hidden layer network (5 neurons, sigmoid activation function), the number of nodes in the input layer corresponded to the number of characteristic genes, the output layer was a binary sorting node (Control/ASD), and a random seed (seed=12345678) control parameter was set to initialize. The backpropagation algorithm (Rprop+) was used for model training, the classification performance was evaluated by the confusion matrix, the group-specific accuracy was calculated, the network topology map was visualized with the weight connection strength, and the key feature contribution was analyzed by the Garson algorithm

2.7 Gene set enrichment analysis

Based on the standardized gene expression profile, Gene set enrichment analysis (GSEA) was used to analyze the GO annotation and KEGG pathway regulatory effects

of the target genes. The Gaussian kernel density estimation was used to calculate the single-sample enrichment score, and the 0-1 normalization process was performed. The samples were divided into high/low expression groups based on the median gene expression threshold, and the differential activation pathways between the groups were screened by independent sample t-test ($P<0.05$ was significant).

2.8 Immune infiltration correlation analysis

Based on the standardized gene expression profile, the Cibersort algorithm was used to deconvolute the proportion of immune cells, and the relative abundance of 22 immune cell subtypes was analyzed by the LM22 feature matrix. The Wilcoxon rank-sum test was used to compare the proportion of cells in the advanced and early stages of carotid atherosclerosis, and the Spearman rank correlation analysis was used to screen the regulatory association between characteristic genes and 22 immune subtypes. The significance threshold was set at $P<0.05$, and the Benjamini-Hochberg method was used to correct multiple hypothesis tests.

2.9 Single-cell sequencing analysis

Single-cell RNA sequencing data from carotid atherosclerotic plaques (GSE253903) were analyzed using 12 samples (6 symptomatic vs. 6 asymptomatic patients). Raw counts were processed through the Seurat (v5.0) and SCP pipelines. Quality control excluded cells with <500 genes, >7,000 genes, or mitochondrial content >15%. Data normalization employed LogNormalize scaling (scale factor=10,000), followed by identification of 2,000 highly variable genes. Batch effects were corrected via Harmony integration (dims=1:22), with dimensionality reduction using PCA and UMAP (dims=1:22). Clustering at resolution=0.15 identified 13 distinct cell populations through the Leiden algorithm.

Differentially expressed genes ($|logFC|>0.25$, $p<0.05$) were identified using Wilcoxon rank-sum tests. Cell type annotation combined SingleR (HumanPrimaryCellAtlasData) and easybio (marker-based validation).

2.10 Immunohistochemical staining analysis

The human tissue samples used by IHC were obtained from the Department of Pathology, the Second Affiliated Hospital of the University of South China. Atherosclerosis and control arterial tissue were preserved by paraffin embedding. The

sections and primary antibody were incubated overnight at 4°C, and the secondary antibody and DAPI were incubated the next day and inspected using the Thermo Fisher Scientific EVOS M7000 Intelligent Imaging System. Data analysis was performed using Image J and Prism software.

2.11 Extraction and culture of mouse bone marrow primary macrophages

Primary macrophages were isolated from bone marrow of 8–12-week-old C57BL/6 mice. Briefly, femurs and tibias were dissected, and bone marrow cells were flushed with ice-cold PBS supplemented with 2% FBS. Erythrocytes were lysed using ACK buffer (155 mM NH4Cl, 10 mM KHCO3, 0.1 mM EDTA) for 5 min at 4°C. Cells were resuspended in RPMI-1640 medium containing 10% FBS, 1% penicillin-streptomycin, and 30% L-929 cell-conditioned medium (as a source of M-CSF) and seeded at 1×10^6 cells/mL. Non-adherent cells were replenished with fresh medium on day 3, and mature macrophages were harvested between days 7–10 post-isolation. C57BL/6 mice were deeply anesthetized with sodium pentobarbital (100–150 mg/kg, i.p.), and the depth of anesthesia was confirmed by the absence of the pedal withdrawal and corneal reflexes. Under deep anesthesia, euthanasia was performed by terminal cardiac puncture followed by exsanguination. Death was confirmed by the cessation of heartbeat and respiration, loss of the corneal reflex, and no signs of recovery during a 5-min observation period. All procedures were performed in accordance with the AVMA Guidelines for the Euthanasia of Animals.

3 Results

3.1 Identification of the carotid atherosclerotic plaque characterization module

Based on standardized gene expression profiles, WGCNA was used to identify modules associated with the progression of carotid atherosclerotic plaques. Construct an adjacency matrix based on the soft threshold (β = 8) (Figure 1A). Seven gene modules (blue, brown, green, gray, red, turquoise, and yellow) were outlined by the dynamic tree cutting algorithm, and the green and red modules were the most strongly correlated with carotid atherosclerotic plaques, with correlation coefficients of 0.26 (Pgreen = 0.008, Pred = 0.009) (Figure 1B-D). Therefore, the green and red modules are identified as critical modules and are identified according to the | MM|≥0.7 and | GS|≥0.3 screened 91 module core genes (Figure 1E).

3.2 Identification of differentially expressed genes in carotid atherosclerotic plaques

A boxplot was constructed to observe the distribution of the normalized samples, and the results showed that the median expression of all samples was evenly distributed within the IQR range of 1.5 times the interquartile range between groups, and no outliers to be excluded were detected, confirming the validity of batch correction (Figure 2A). Differential analysis identified 605 differentially expressed genes (DEGs, including 373 up-regulated and 223 down-regulated genes) between the advanced and early stages. Use the volcano map to show $|\log2FC| \geq 0.585$ and $P < 0.0$ for differentially expressed genes, while showing a cluster heatmap of the expression of the top 20 differentially expressed genes (Figure 2B-C).

After intersecting the core genes of the module with DEGs, 16 candidate genes were finally screened (Figure 2D).

3.3 Functional enrichment analysis of candidate genes

Gene Ontology (GO) and pathway (KEGG) annotation analyses were performed on 16 candidate genes to explore their potential biological functions (Figure 3A-D). In terms of biological process (BP), candidate genes are mainly involved in cellular stress. In terms of cellular component (CC), candidate genes are associated with cell adhesion junctions. At the molecular function (MF) level, candidate genes are enriched in Toll-like receptor binding and other functions. Pathway enrichment analysis showed that the

candidate genes were related to Toll-like receptor signaling pathways. In addition, disease ontology (DO) enrichment analysis revealed that candidate genes were associated with a variety of inflammatory diseases (Figure 3E).

3.4 Machine learning screens for characteristic genes

Use a variety of machine learning algorithms to further screen reliable and critical biomarkers. The LASSO regression model identified 5 potential diagnostic biomarkers, the RF algorithm identified 10 diagnostic genes, and the SVM screened 9 key genes (Figure 4A-D). Mapping the Venn plot revealed that 3 key genes (*TLR1*, *C14orf132*, and *SVIL*) overlapped and were identified as reliable biomarkers (Figure 4E). An ANN diagnostic model was constructed based on gene weights, and the training group and control group were divided 7:3 for further validation (Figure. 4F). The information of the three characteristic genes is as follows (Table 2).

Table2 Characteristic gene infromation

Symbol	Description	Chromosome	Length	Mass (Da)
TLR1	Toll-like receptor 1	4p14	786	90,291
C14orf132	chromosome 14 open reading frame 132	14q32.2	83	8,830
SVIL	Supervillin	10p11.23	2,214	247,746

3.5 Characteristic gene validation and functional analysis

Testing the model using ROC curves revealed that all three genes had good diagnostic performance (Figure 5A-B). Continued GESAGO and GESAKEGG annotation analysis of the unique risk gene TLR1 showed that it was significantly associated with inflammatory response and was mainly enriched in the Toll-like receptor signaling pathway (Figure 5C-D). Immunohistochemical staining based on clinical samples and analysis of external validation sets also revealed that TLR1 is a risk gene for atherosclerosis (Figure 5E-F).

3.6 Immune infiltration analysis

The CIBERSORT algorithm was used to explore the immune infiltration of advanced carotid atherosclerotic plaque, and it was found that the proportion of macrophages in the advanced stage increased significantly compared with the early stage (Figure 6A), while the proportions of B cells, monocytes and NK cells decreased

297 significantly, and were mainly related to the increase of TLR1 expression (Figure 6B-
298 D).
299
300 **3.7 Single-cell sequencing analysis of carotid atherosclerotic plaques**
301 SingleR and easybio were used to annotate the cells of carotid atherosclerotic
302 vessels (Figure 7A-B), and it was found that TLR1 was mainly distributed in
303 macrophages, SVIL was mainly distributed in fibroblasts and endothelial cells, and
304 C14orf132 was mainly distributed in fibroblasts (Figure 7C).
305
306 **3.8 The TLR1-activated macrophage inflammatory response is inhibited in**
307 **carotid** atherosclerotic mice ingested with coriander
308 A mouse model of carotid atherosclerosis was constructed, bone marrow-derived
309 macrophages were extracted and induced into M1 macrophages using LPS and IFN-γ.
310 Subsequently, Western blot technology was used to detect the increase in the expression
311 of Tlr1 and inflammatory factors (Figure 8A-B), and the increase was greatly decreased
312 in the mouse model fed coriander, indicating that coriander could inhibit the
313 inflammatory response of carotid atherosclerotic mice (Figure 8C-D).
314

4 Conclusion and Discussion

The diagnosis and treatment of atherosclerosis has always been a research hotspot in the field of cardiovascular diseases[21-23]. However, the screening of core genes for carotid atherosclerotic plaque progression and its underlying mechanisms remain a challenge for current research[24]. In this review, we screened out TLR1 and other genes through systematic bioinformatics analysis, combined with machine learning, ANN and SHAP, and explored the potential role of coriander in inhibiting macrophage inflammation.

Toll-like receptor 1 (TLR1), encoded on chromosome 4p14 as a 786-amino-acid transmembrane protein (90.3 kDa), functions as a key pattern recognition receptor that partners with TLR2 to form heterodimers essential for detecting bacterial triacylated lipopeptides. Its leucine-rich repeat extracellular domain facilitates pathogen-associated molecular pattern recognition, while the cytoplasmic Toll/IL-1 receptor (TIR) domain initiates MyD88-dependent signaling cascades, driving NF-κB and MAPK activation. Recent evidence extends TLR1's role beyond innate immunity, implicating its serine/threonine phosphorylation motifs (e.g., Tyr-775, Ser-787) in sterile inflammatory processes, particularly through crosstalk with JAK-STAT pathways.

In atherosclerosis, our single-cell transcriptomics identified TLR1 as a marker of pro-inflammatory macrophages within human carotid plaque shoulders—anatomic sites prone to rupture. Structural analysis reveals its molecular mass corresponds to post-translationally glycosylated forms critical for ligand binding and membrane trafficking. Notably, the TLR1 locus (4p14) overlaps with cardiovascular risk variants (e.g., rs4833095) identified in genome-wide association studies, suggesting regulatory polymorphisms may influence its expression in vascular inflammation.

Mechanistically, TLR1 amplifies IL-17 signaling via STAT3 phosphorylation (Tyr-705), creating a feed-forward loop that enhances neutrophil extracellular trap formation and matrix metalloproteinase activity, as evidenced by spatial colocalization with MMP9+ regions (Pearson's $r = 0.67$). These findings position TLR1 as a dual-functional mediator, bridging pathogen sensing and sterile inflammation, with therapeutic potential highlighted by network pharmacology predictions of coriandrin as a TLR1 modulator.

There are also some limitations to this study. First, our analysis was based on multiple independent cohorts in a public database, and the European population needed

to expand the sample size and data of different ethnic groups to further improve the accuracy and broad applicability of the model. Second, although we have screened key genes through a variety of methods, their specific molecular mechanisms and mechanisms of action in carotid atherosclerosis still need to be verified by molecular biology experiments. In summary, this study screened out potential biomarkers of carotid atherosclerosis progression such as TLR1 through bioinformatics analysis, and revealed the specific expression patterns and potential molecular regulatory mechanisms of these genes in macrophages. These findings provide a new theoretical basis for the personalized treatment of carotid atherosclerosis progression, and also lay a foundation for subsequent experimental verification and clinical application.

Declarations

Data availability statement

The datasets(GEO accession numbers: GSE28829, GSE41571, GSE43292, GSE100927, and GSE253903)presented in this study can be found in online repositories. Data were obtained from the GEO database of the National Center for Biotechnology Information (URL: https://www.ncbi.nlm.nih.gov/geo/).

The publicly available gene expression datasets analysed in this study were obtained from the Gene Expression Omnibus (GEO; **https://www.ncbi.nlm.nih.gov/geo/**) under accession numbers GSE28829 (**https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28829**), GSE41571, GSE43292, GSE100927 and GSE253903. The experimental data generated during this study are included in the article and its Supplementary Information files. No original KEGG pathway map images are reproduced. Any additional data is available from the corresponding author upon reasonable request.

Ethics statement

The study involving human was reviewed and approved by Clinical Research Ethics Committee of the Second Affiliated Hospital of University of South China (approval number: 2026-015). This study was conducted in accordance with a protocol approved by the Ethics Committee and adhered to the principles of the Declaration of Helsinki, as well as the ethical principles, moral standards, and applicable laws, regulations, norms, and institutional policies of China. The ethics committee waived the requirement of written informed consent for participation.

The study involving animals was reviewed and approved by Ethics Committee of University of South China (approval number: 11000076). All procedures were performed in accordance with the AVMA Guidelines for the Euthanasia of Animals.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions Statement

MZ and JX: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing– original draft. SX: Conceptualization, Investigation, Project administration, Resources, Supervision. KH: Con ceptualization, Writingreview& editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by Natural Science Foundation of Hunan Province (2025JJ81016,2025JJ81000,2026JJ82274) and the Foundation of Hunan Provincial Key Laboratory (2023TP1014).

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

References

[1]. Saba L., Cau R., Vergallo R., et al., *Carotid artery atherosclerosis: mechanisms of instability and clinical implications.* Eur Heart J, 2025. **46**(10): p. 904-921.

[2]. Chen W. H., Jin W., Lyu P. Y., et al., *Carotid Atherosclerosis and Cognitive Impairment in Nonstroke Patients.* Chin Med J (Engl), 2017. **130**(19): p. 2375-2379.

[3]. Song P., Fang Z., Wang H., et al., *Global and regional prevalence, burden, and risk factors for carotid atherosclerosis: a systematic review, meta-analysis, and modelling study.* Lancet Glob Health, 2020. **8**(5): p. e721-e729.

[4]. Zheng H., Li H., Wang Y., et al., *Rosuvastatin Slows Progression of Carotid Intima-Media Thickness: The METEOR-China Randomized Controlled Study.* Stroke, 2022. **53**(10): p. 3004-3013.

[5]. Yaneva-Sirakova T., Traykov L., Karamfiloff K., et al., *Neurotrophins in carotid atherosclerosis and stenting.* Ann Med, 2023. **55**(1): p. 335-341.

[6]. Saba L., Cau R., Murgia A., et al., *Carotid Plaque-RADS: A Novel Stroke Risk Classification System.* JACC Cardiovasc Imaging, 2024. **17**(1): p. 62-75.

[7]. Ahmad H., Gopakumar J., Nachun D. C., et al., *Single cell RNA sequencing of haematopoietic cells in fresh and frozen human atheroma tissue.* Cardiovasc Res, 2025. **121**(3): p. 396-404.

[8]. Zernecke A., Erhard F., Weinberger T., et al., *Integrated single-cell analysis-based classification of vascular mononuclear phagocytes in mouse and human atherosclerosis.* Cardiovasc Res, 2023. **119**(8): p. 1676-1689.

[9]. Slysz J., Sinha A., DeBerge M., et al., *Single-cell profiling reveals inflammatory polarization of human carotid versus femoral plaque leukocytes.* JCI Insight, 2023. **8**(17).

[10]. Xue Y., Luo M., Hu X., et al., *Macrophages regulate vascular smooth muscle cell function during atherosclerosis progression through IL-1β/STAT3 signaling.* Commun Biol, 2022. **5**(1): p. 1316.

[11]. Ji H., Dong Z., Yang Y., et al., *Neixiao-ruanmai decoction No 2 improves carotid atherosclerosis by modulating gut microbiota and inhibiting TLR4/NF-κB pathway activation.* Phytomedicine, 2025. **142**: p. 156775.

[12]. Huang R., Hu Z., Chen X., et al., *The Transcription Factor SUB1 Is a Master Regulator of the Macrophage TLR Response in Atherosclerosis.* Adv Sci (Weinh), 2021. **8**(19): p. e2004162.

[13]. Zhou W., Li X., Zhou H., et al., *TNF-α/IL-1β/IL-1α/IL-12 inflammatory cytokine axes coupled with TLR1/TLR3/TLR5/MYD88 immune signaling pathway over-activation contribute to simultaneous carotid and coronary artery and occlusion in elderly patients.* Cytokine, 2025. **185**: p. 156808.

[14]. Chaudhari A., Axelsson C., Mattsson Hultén L., and Rotter Sopasakis V., *Toll-like Receptors 1, 3 and 7 Activate Distinct Genetic Features of NF-κB Signaling and γ-Protocadherin Expression in Human Cardiac Fibroblasts.* Inflammation, 2025.

[15]. Tian Y., Huang B., Li J., et al., *Identification of the Association Between Toll-Like Receptors and T-Cell Activation in Takayasu's Arteritis.* Front Immunol, 2021. **12**: p. 792901.

[16]. Rotter Sopasakis V., Sandstedt J., Johansson M., et al., *Toll-like receptor-mediated inflammation markers are strongly induced in heart tissue in patients*

with cardiac disease under both ischemic and non-ischemic conditions. Int J Cardiol, 2019. **293**: p. 238-247.
[17]. Goleij P., Sanaye P. M., Alam W., et al., *Unlocking daidzein's healing power: Present applications and future possibilities in phytomedicine.* Phytomedicine, 2024. **134**: p. 155949.
[18]. Cao Y., Yao W., Yang T., et al., *Elucidating the mechanisms of Buyang Huanwu Decoction in treating chronic cerebral ischemia: A combined approach using network pharmacology, molecular docking, and in vivo validation.* Phytomedicine, 2024. **132**: p. 155820.
[19]. Arab Z. N., Khayatan D., Razavi S. M., et al., *Phytochemicals as Modulators of Paraoxonase-1 in Health and Diseases.* Antioxidants (Basel), 2022. **11**(7).
[20]. Wang S., Ren H., Zhong H., et al., *Combined berberine and probiotic treatment as an effective regimen for improving postprandial hyperlipidemia in type 2 diabetes patients: a double blinded placebo controlled randomized study.* Gut Microbes, 2022. **14**(1): p. 2003176.
[21]. Porsch F. and Binder C. J., *Autoimmune diseases and atherosclerotic cardiovascular disease.* Nat Rev Cardiol, 2024. **21**(11): p. 780-807.
[22]. Xu S., Ilyas I., Little P. J., et al., *Endothelial Dysfunction in Atherosclerotic Cardiovascular Diseases and Beyond: From Mechanism to Pharmacotherapies.* Pharmacol Rev, 2021. **73**(3): p. 924-967.
[23]. Burnett J. R., Hooper A. J. and Hegele R. A., *Remnant Cholesterol and Atherosclerotic Cardiovascular Disease Risk.* J Am Coll Cardiol, 2020. **76**(23): p. 2736-2739.
[24]. Bonati L. H., Jansen O., de Borst G. J., and Brown M. M., *Management of atherosclerotic extracranial carotid artery stenosis.* Lancet Neurol, 2022. **21**(3): p. 273-283.

Figure Legend

Figure 1. WGCNA analysis of significant modules in AS patients

A: Selection of soft threshold for AS patients; B: Gene cluster tree diagram of AS patients (dynamic tree cutting algorithm); C: Heat map of each module and clinical features; D: Pearson correlation test for the most significant modules (green and red); E: Overlapping genes of the green and red modules.

Figure 2. Identification of DEGs in patients with AS

A: Principal component analysis results of AS patients (GSE28829, GSE41571, and GSE43292) before and after de-batching; B: Volcano diagram of all differential genes, blue indicates down-regulated differential genes, and red represents up-regulated differential genes; C: Heat map of the top 20 differential genes; D: Venn plot identifies genes common to WGCNA and differential analysis.

Figure 3. Feature enrichment analysis

A: GO enrichment analysis of common differential genes in biological process; B: GO enrichment analysis of common differential genes in cellular component; C: GO enrichment analysis of common differential genes in molecular function; D: KEGG enrichment analysis of common differential genes; E: DO enrichment analysis of common differential genes.

Figure 4. Identify characteristic genes through machine learning

A-B: Gene expression cluster analysis of AS patients based on LASSO regression model; C-D: Based on the random forest algorithm, the relationship between the total number of trees and the error rate of the dataset, and the ranking of the relative importance score of genes; E: Venn diagram identifies the common characteristic genes identified by three machine learning algorithms. F: Artificial neural network (ANN) based on three characteristic genes.

Figure 5. Verification and analysis of characteristic genes

A: Box plot shows the expression distribution of characteristic genes in the training set of AS patients; B: ROC curves of characteristic genes based on standardized expression profiles; C-D: GSEA analysis showed the results of GO and KEGG enrichment analysis when TLR1 was highly expressed. E: IHC showed the differential expression of TLR1, SVIL and C14orf132 in the carotid artery tissue of AS patients. F: Boxplot showing the expression distribution of characteristic genes in the validation set (GSE100927) of AS patients.

Figure 6. Immune infiltration analysis of global and characteristic genes

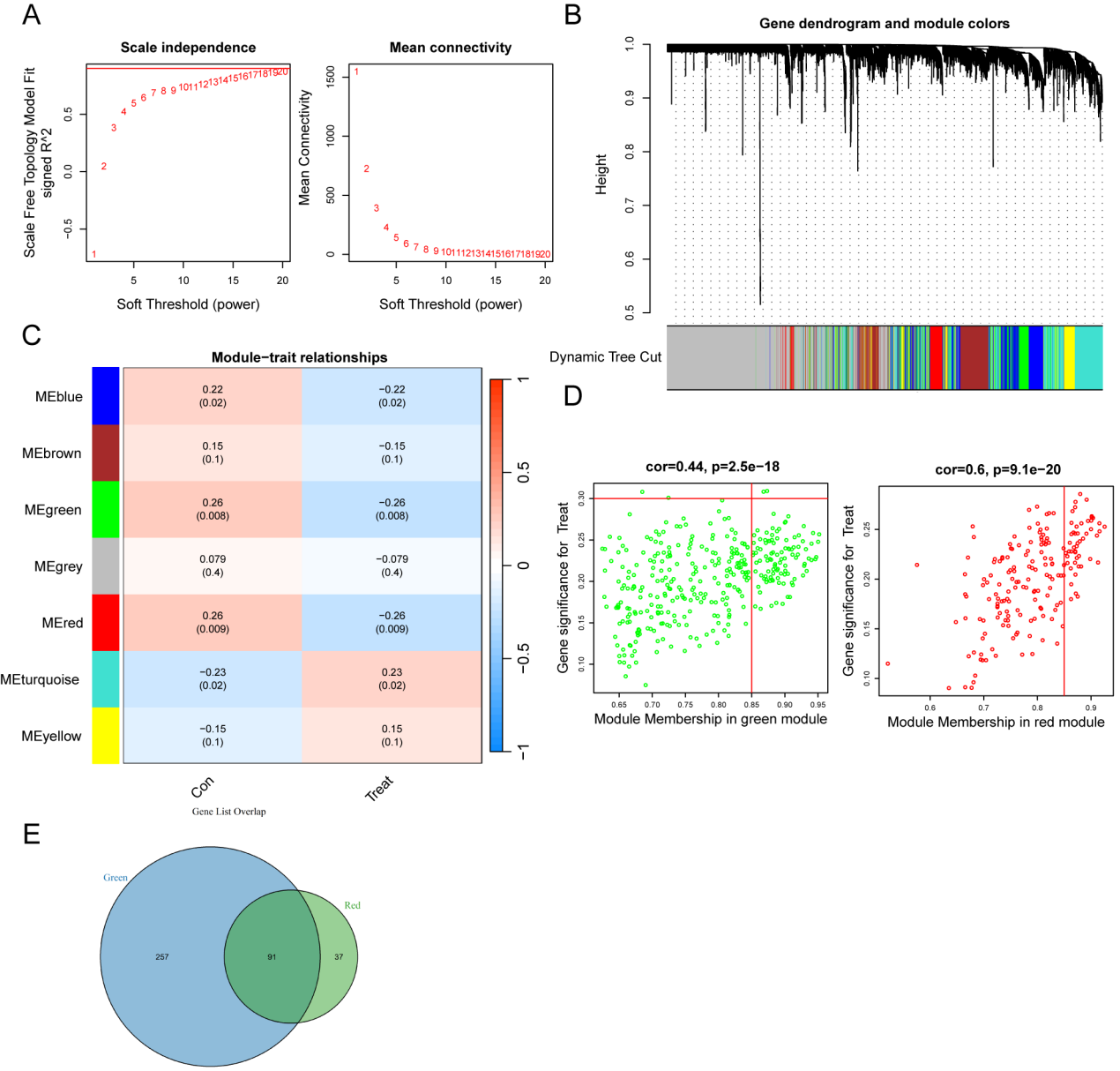
A: Proportion of immune cells in AS patients and control groups; B: Association between C14orf132 and immune cells in AS patients; C: Association between SVIL and immune cells in AS patients; D: Association between TLR1 and immune cells in AS patients.

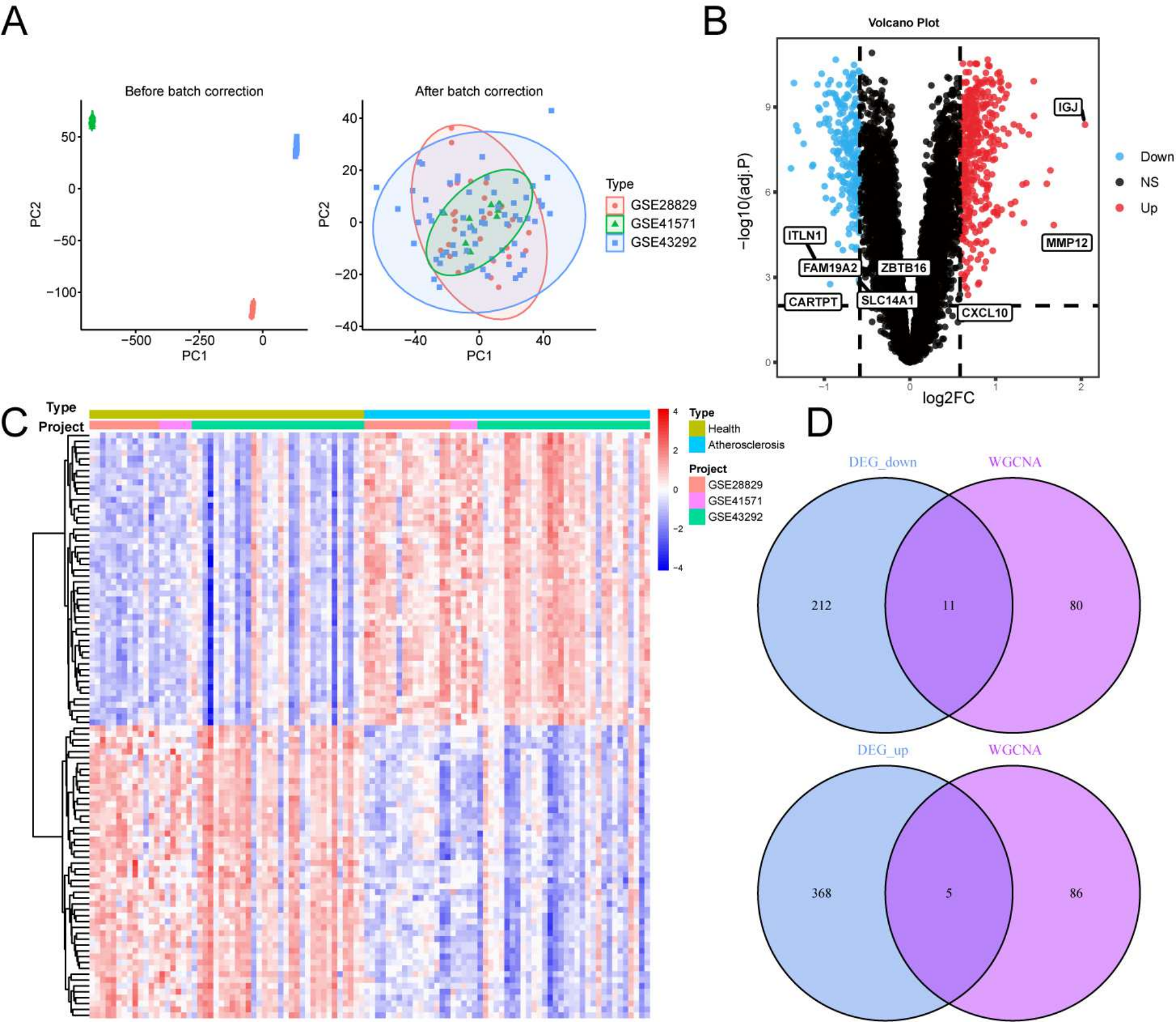
Figure 7. Single-cell sequencing analysis of characteristic genes

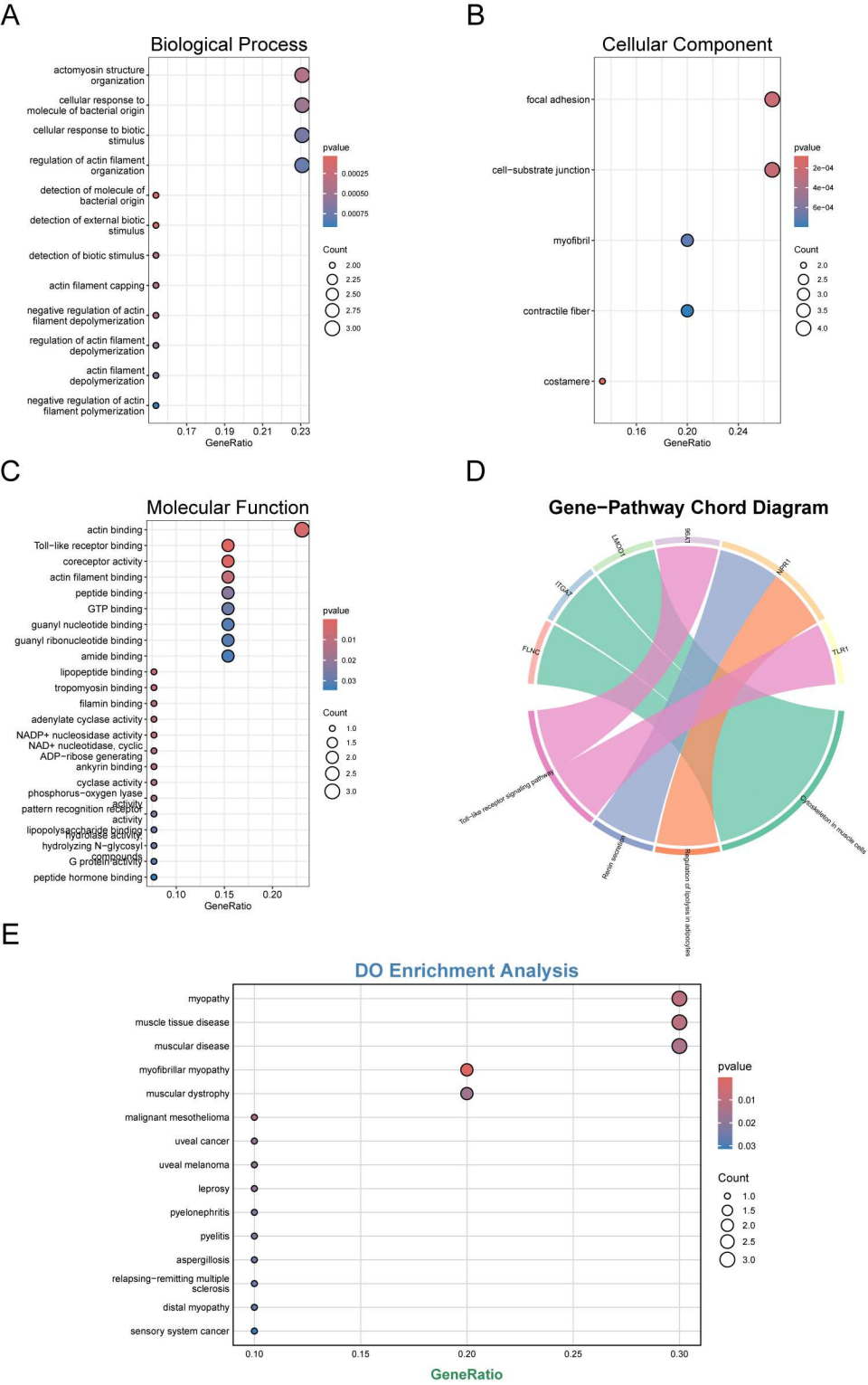
A: SingleR-based single-cell cluster annotation; B: Single-cell cluster annotation based on easybio; C: Cell distribution of TLR1, SVIL and C14orf132.

Figure 8. In vitro cell model validation

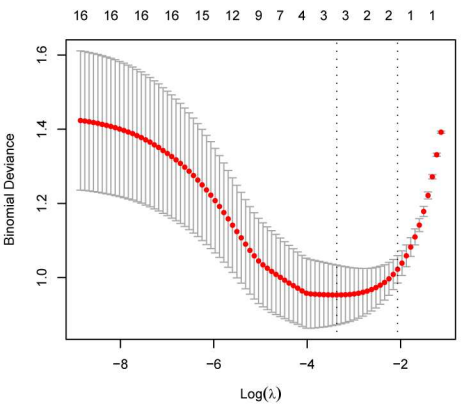
A: Expression of TLR1 and inflammatory cytokines in inflammatory macrophages; B: WB quantitative analysis results; C: The expression of TLR1 and inflammatory factors in inflammatory macrophages induced by coriander extract; D: WB quantitative analysis results.


A
Scale independence
Scale Free Topology Model Fit
signed R^2
Soft Threshold (power)
Mean connectivity
Mean Connectivity
Soft Threshold (power)
B
Gene dendrogram and module colors
Height
Dynamic Tree Cut
C
Module−trait relationships
MEblue
MEbrown
MEgreen
MEgrey
MEred
MEturquoise
MEyellow
0.22
(0.02)
−0.22
(0.02)
0.15
(0.1)
−0.15
(0.1)
0.26
(0.008)
−0.26
(0.008)
0.079
(0.4)
−0.079
(0.4)
0.26
(0.009)
−0.26
(0.009)
−0.23
(0.02)
0.23
(0.02)
−0.15
(0.1)
0.15
(0.1)
Con
Treat
Gene List Overlap
D
cor=0.44, p=2.5e−18
Gene significance for Treat
Module Membership in green module
cor=0.6, p=9.1e−20
Gene significance for Treat
Module Membership in red module
E
Green
Red
257
91
37

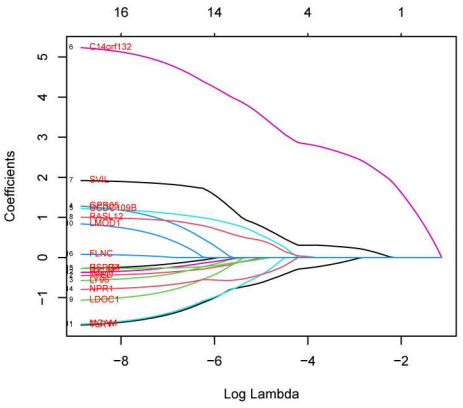

A
Before batch correction
After batch correction
PC1
PC2
Type
GSE28829
GSE41571
GSE43292
B
Volcano Plot
−log10(adj.P)
log2FC
Down
NS
Up
IGJ
MMP12
ITLN1
FAM19A2
ZBTB16
CARTPT
SLC14A1
CXCL10
C
Type
Project
Health
Atherosclerosis
GSE28829
GSE41571
GSE43292
D
DEG_down
WGCNA
212
11
80
DEG_up
WGCNA
368
5
86


A
Biological Process
actomyosin structure organization
cellular response to molecule of bacterial origin
cellular response to biotic stimulus
regulation of actin filament organization
detection of molecule of bacterial origin
detection of external biotic stimulus
detection of biotic stimulus
actin filament capping
negative regulation of actin filament depolymerization
regulation of actin filament depolymerization
actin filament depolymerization
negative regulation of actin filament polymerization
0.17
0.19
0.21
0.23
GeneRatio
pvalue
0.00025
0.00050
0.00075
Count
2.00
2.25
2.50
2.75
3.00
B
Cellular Component
focal adhesion
cell−substrate junction
myofibril
contractile fiber
costamere
0.16
0.20
0.24
GeneRatio
pvalue
2e−04
4e−04
6e−04
Count
2.0
2.5
3.0
3.5
4.0
C
Molecular Function
actin binding
Toll−like receptor binding
coreceptor activity
actin filament binding
peptide binding
GTP binding
guanyl nucleotide binding
guanyl ribonucleotide binding
amide binding
lipopeptide binding
tropomyosin binding
filamin binding
adenylate cyclase activity
NADP+ nucleosidase activity
NAD+ nucleotidase, cyclic ADP−ribose generating
ankyrin binding
cyclase activity
phosphorus−oxygen lyase activity
pattern recognition receptor activity
lipopolysaccharide binding
hydrolase activity, hydrolyzing N−glycosyl compounds
G protein activity
peptide hormone binding
0.10
0.15
0.20
GeneRatio
pvalue
0.01
0.02
0.03
Count
1.0
1.5
2.0
2.5
3.0
D
Gene−Pathway Chord Diagram
LY96
LMOD1
NPR1
ITGA5
TLR1
FLNC
Toll−like receptor signaling pathway
Actin skeleton in muscle cells
Rho GTPase
Regulation of lipolysis in adipocytes
E
DO Enrichment Analysis
myopathy
muscle tissue disease
muscular disease
myofibrillar myopathy
muscular dystrophy
malignant mesothelioma
uveal cancer
uveal melanoma
leprosy
pyelonephritis
pyelitis
aspergillosis
relapsing−remitting multiple sclerosis
distal myopathy
sensory system cancer
0.10
0.15
0.20
0.25
0.30
GeneRatio
pvalue
0.01
0.02
0.03
Count
1.0
1.5
2.0
2.5
3.0

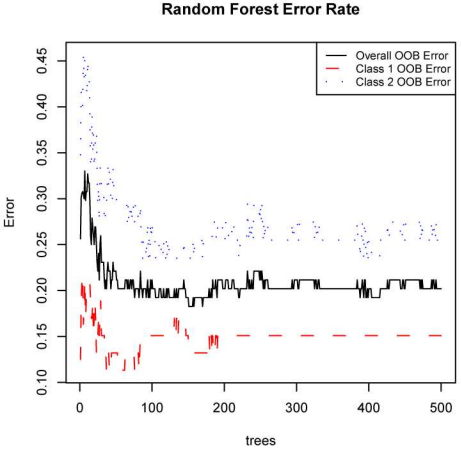
A


16 16 16 16 15 12 9 7 4 3 3 2 2 1 1
Binomial Deviance
1.0
1.2
1.4
1.6
−8
−6
−4
−2
Log(λ)

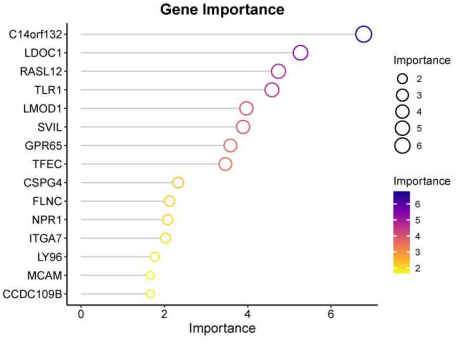
B


16 14 4 1
Coefficients
−1
0
1
2
3
4
5
6
C14orf132
SVIL
GPR65
RASL12
LDOC1
FLNC
NPR1
LDOC1
−8
−6
−4
−2
Log Lambda

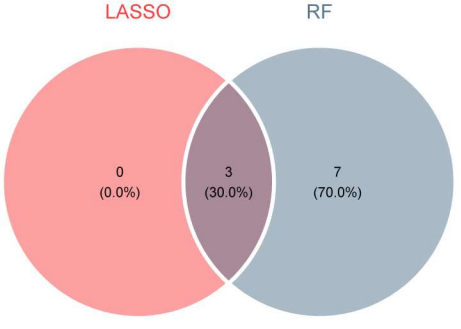
C


Random Forest Error Rate
Overall OOB Error
Class 1 OOB Error
Class 2 OOB Error
Error
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0
100
200
300
400
500
trees

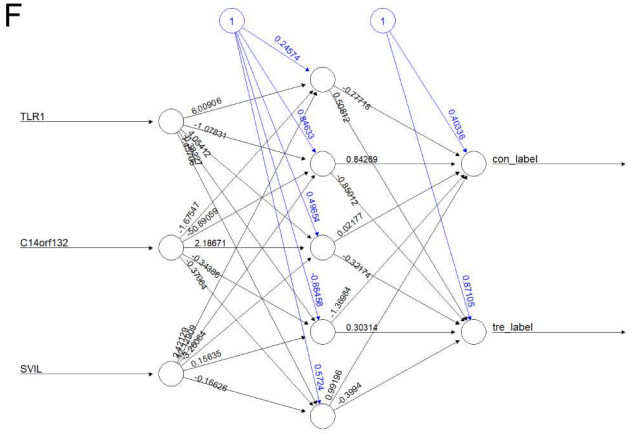
D

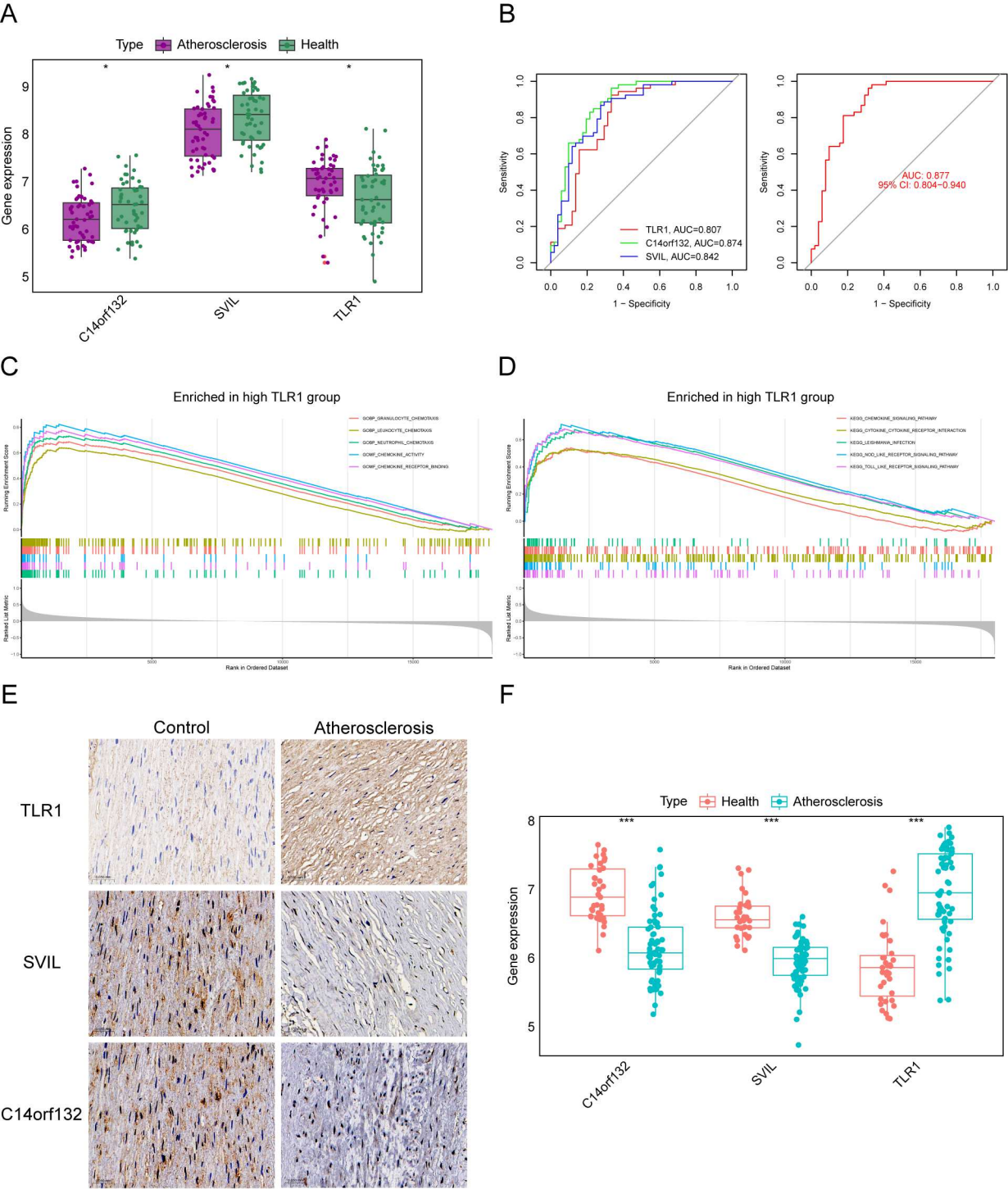

Gene Importance
C14orf132
LDOC1
RASL12
TLR1
LMOD1
SVIL
GPR65
TFEC
CSPG4
FLNC
NPR1
ITGA7
LY96
MCAM
CCDC109B
0
2
4
6
Importance
Importance
2
3
4
5
6
Importance
6
5
4
3
2

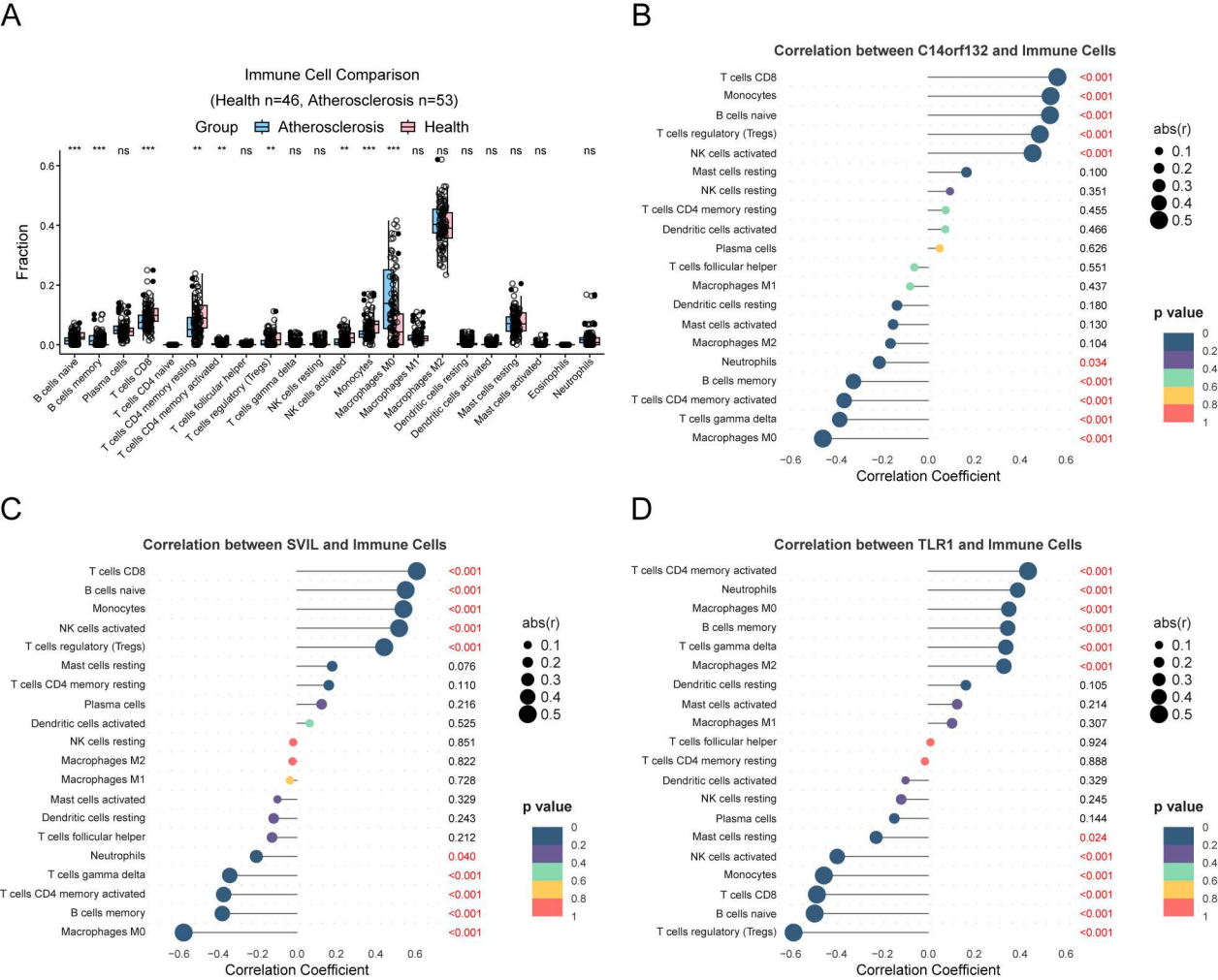
E

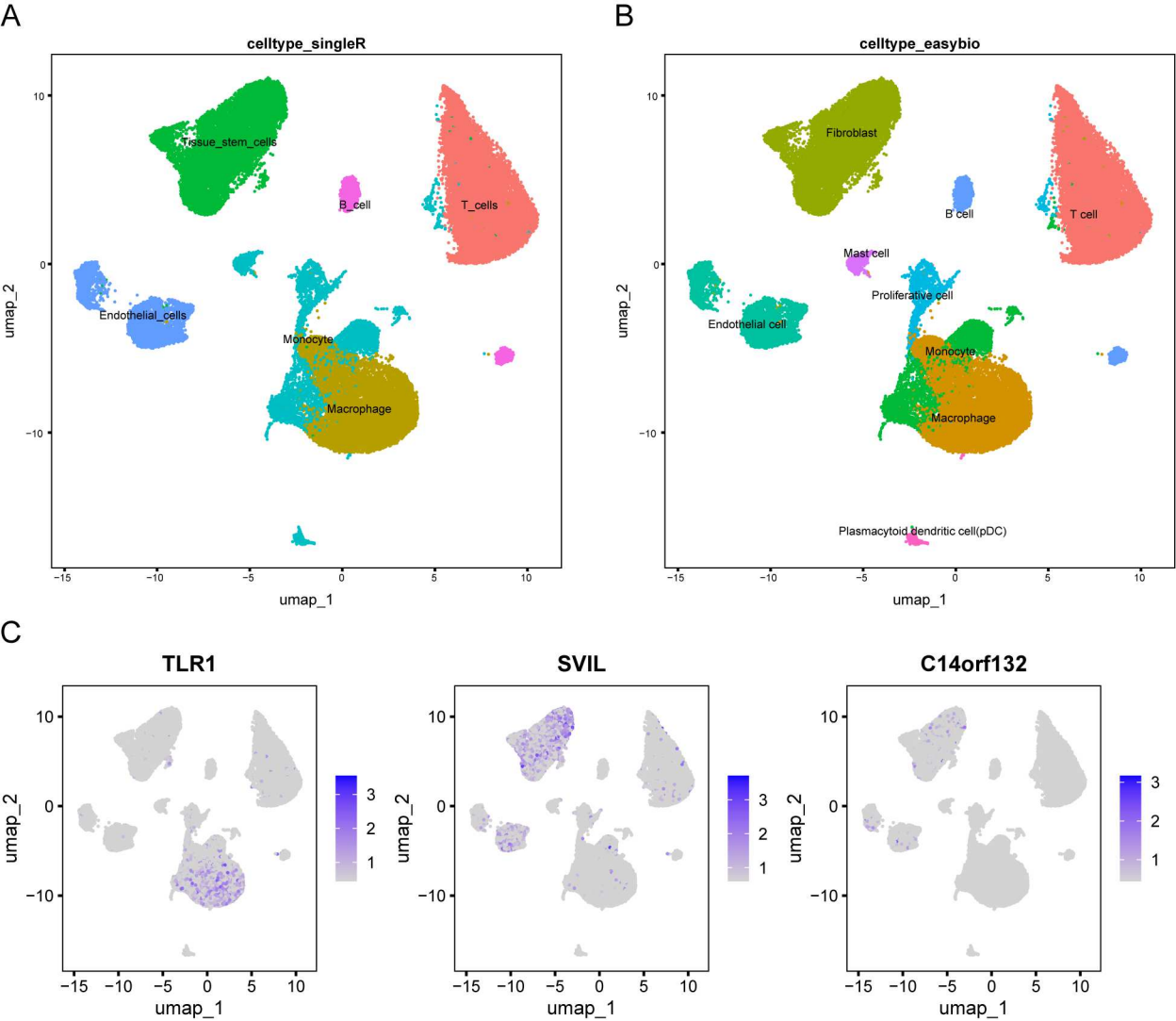

LASSO
RF
0
(0.0%)
3
(30.0%)
7
(70.0%)

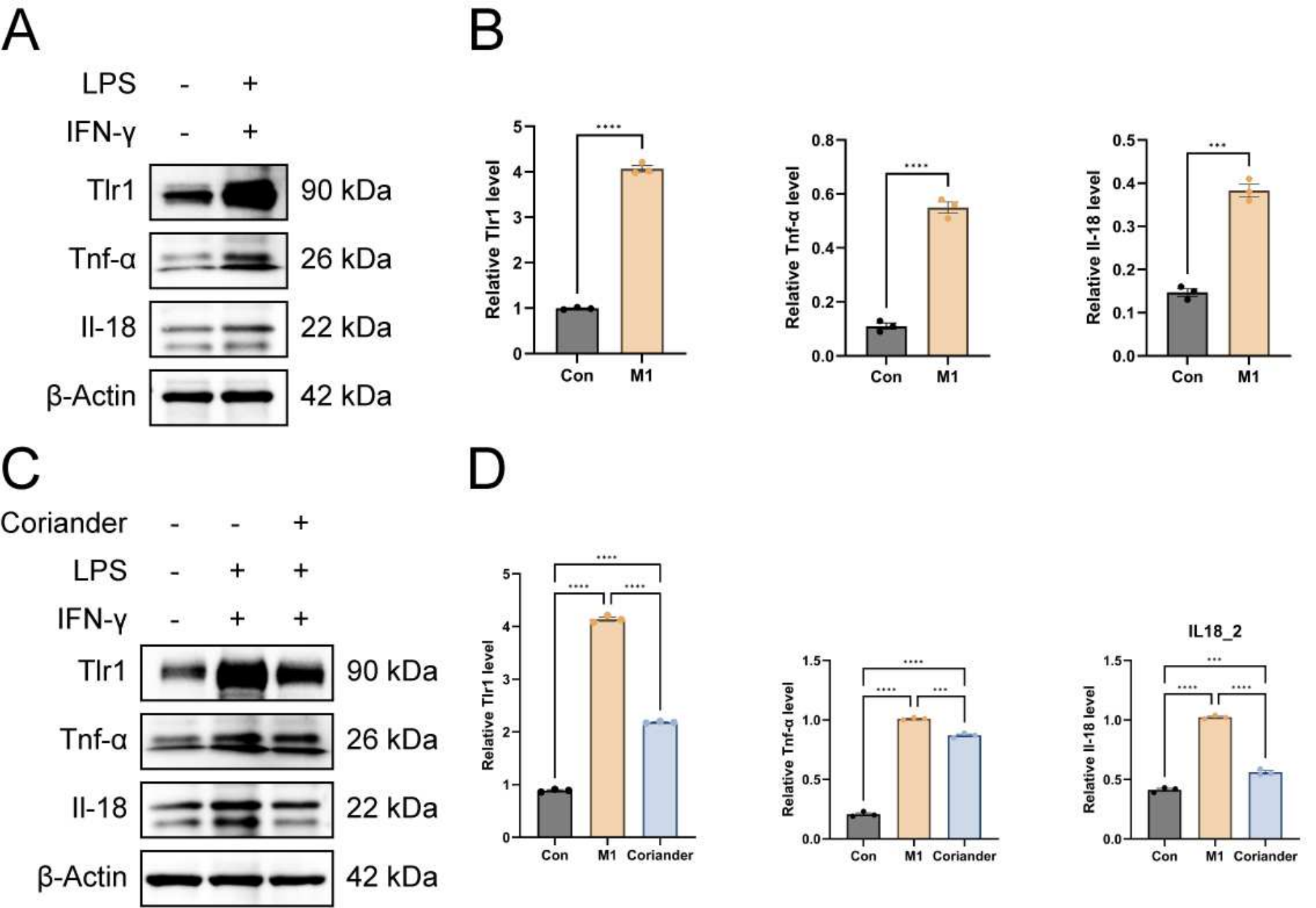
F


1
1
TLR1
C14orf132
SVIL
con_label
tre_label
0.24514
6.09906
-1.07831
-3.40947
-3.58907
0.84633
-0.27716
-0.54812
0.84289
0.40356
-0.89572
1.67647
-50.89069
2.18671
-0.34496
-0.37984
0.49854
0.02177
-0.32174
-0.86456
1.39894
0.87105
0.30314
0.15635
-0.16626
0.5724
0.99196
-0.3894


A
Type
Atherosclerosis
Health
Gene expression
C14orf132
SVIL
TLR1
B
Sensitivity
1 − Specificity
TLR1, AUC=0.807
C14orf132, AUC=0.874
SVIL, AUC=0.842
AUC: 0.877
95% CI: 0.804−0.940
C
Enriched in high TLR1 group
GOBP_GRANULOCYTE_CHEMOTAXIS
GOBP_LEUKOCYTE_CHEMOTAXIS
GOBP_NEUTROPHIL_CHEMOTAXIS
GOMF_CHEMOKINE_ACTIVITY
GOMF_CHEMOKINE_RECEPTOR_BINDING
Running Enrichment Score
Ranked List Metric
Rank in Ordered Dataset
D
Enriched in high TLR1 group
KEGG_CHEMOKINE_SIGNALING_PATHWAY
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION
KEGG_LEISHMANIA_INFECTION
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY
Running Enrichment Score
Ranked List Metric
Rank in Ordered Dataset
E
Control
Atherosclerosis
TLR1
SVIL
C14orf132
F
Type
Health
Atherosclerosis
Gene expression
C14orf132
SVIL
TLR1


A
Immune Cell Comparison
(Health n=46, Atherosclerosis n=53)
Group Atherosclerosis Health
Fraction
B
Correlation between C14orf132 and Immune Cells
T cells CD8 <0.001
Monocytes <0.001
B cells naive <0.001
T cells regulatory (Tregs) <0.001
NK cells activated <0.001
Mast cells resting 0.100
NK cells resting 0.351
T cells CD4 memory resting 0.455
Dendritic cells activated 0.466
Plasma cells 0.626
T cells follicular helper 0.551
Macrophages M1 0.437
Dendritic cells resting 0.180
Mast cells activated 0.130
Macrophages M2 0.104
Neutrophils 0.034
B cells memory <0.001
T cells CD4 memory activated <0.001
T cells gamma delta <0.001
Macrophages M0 <0.001
Correlation Coefficient
abs(r)
p value
C
Correlation between SVIL and Immune Cells
T cells CD8 <0.001
B cells naive <0.001
Monocytes <0.001
NK cells activated <0.001
T cells regulatory (Tregs) <0.001
Mast cells resting 0.076
T cells CD4 memory resting 0.110
Plasma cells 0.216
Dendritic cells activated 0.525
NK cells resting 0.851
Macrophages M2 0.822
Macrophages M1 0.728
Mast cells activated 0.329
Dendritic cells resting 0.243
T cells follicular helper 0.212
Neutrophils 0.040
T cells gamma delta <0.001
T cells CD4 memory activated <0.001
B cells memory <0.001
Macrophages M0 <0.001
Correlation Coefficient
abs(r)
p value
D
Correlation between TLR1 and Immune Cells
T cells CD4 memory activated <0.001
Neutrophils <0.001
Macrophages M0 <0.001
B cells memory <0.001
T cells gamma delta <0.001
Macrophages M2 <0.001
Dendritic cells resting 0.105
Mast cells activated 0.214
Macrophages M1 0.307
T cells follicular helper 0.924
T cells CD4 memory resting 0.888
Dendritic cells activated 0.329
NK cells resting 0.245
Plasma cells 0.144
Mast cells resting 0.024
NK cells activated <0.001
Monocytes <0.001
T cells CD8 <0.001
B cells naive <0.001
T cells regulatory (Tregs) <0.001
Correlation Coefficient
abs(r)
p value


A
celltype_singleR
Tissue_stem_cells
B_cell
T_cells
Endothelial_cells
Monocyte
Macrophage
umap_1
umap_2
B
celltype_easybio
Fibroblast
B cell
T cell
Mast cell
Proliferative cell
Endothelial cell
Monocyte
Macrophage
Plasmacytoid dendritic cell(pDC)
umap_1
umap_2
C
TLR1
SVIL
C14orf132
umap_1
umap_2


A
LPS - +
IFN-γ - +
Tlr1
90 kDa
Tnf-α
26 kDa
Il-18
22 kDa
β-Actin
42 kDa
B
Relative Tlr1 level
Relative Tnf-α level
Relative Il-18 level
Con
M1
C
Coriander - - +
LPS - + +
IFN-γ - + +
Tlr1
90 kDa
Tnf-α
26 kDa
Il-18
22 kDa
β-Actin
42 kDa
D
Relative Tlr1 level
Relative Tnf-α level
IL18_2
Relative Il-18 level
Con
M1
Coriander