

Identification of Macrophage Polarization-Related Genes as Biomarkers for Tendinopathy Based on Bioinformatics Analysis

Ling Ren

Sichuan University

Chengqi He

Sichuan University

Yuxi Qin

qinyuxi_12345@163.com

Affiliated Sichuan Provincial Rehabilitation Hospital of Chengdu University of TCM

Article

Keywords: Tendinopathy, Macrophage polarization, Biomarkers, Machine learning, WGCNA, Immune infiltration, Nomogram, Bioinformatics

Posted Date: August 11th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10345905/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 20th, 2026.
See the published version at <https://doi.org/10.1038/s41598-026-75576-z>.

Identification of Macrophage Polarization-Related Genes as Biomarkers for Tendinopathy Based on Bioinformatics Analysis

Ling Ren^{1,2} Chengqi He^{1,2} and Yuxi Qin^{3,4,5}

1 Department of Rehabilitation Medicine and Institute of Rehabilitation Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan, PR China

2 Key Laboratory of Rehabilitation Medicine in Sichuan Province, Chengdu, Sichuan, PR China

3 Affiliated Sichuan Provincial Rehabilitation Hospital of Chengdu University of TCM, Chengdu, Sichuan, PR China

4 Affiliated Sichuan Ba-Yi Rehabilitation Center of Chengdu University of TCM, Chengdu, Sichuan, PR China

5 Clinical Medicine School for Rehabilitation of Chengdu University of TCM, Chengdu, Sichuan, PR China

Corresponding author: Chengqi He, Email: hxfhcq2015@126.com and Yuxi Qin, Email: qinyuxi_12345@163.com

Ling Ren and Yuxi Qin have contributed equally to this work and share first authorship

Abstract

Background:

Tendinopathy is a common overuse-induced musculoskeletal disorder characterized by chronic pain, collagen disorganization, and immune cell infiltration. Macrophage polarization plays a critical role in tendon inflammation and tissue repair, but its diagnostic potential in tendinopathy remains largely unexplored.

Methods:

We utilized two publicly available gene expression datasets, GSE26051 and GSE180836, sourced from the GEO database, serving as the training and validation cohorts. Macrophage polarization-related genes (MPRGs) were integrated with weighted gene co-expression network analysis (WGCNA) and differential expression analysis to identify differentially expressed MPRGs (DE-MPRGs). Three machine learning algorithms—LASSO, SVM-RFE, and Boruta—were employed to screen for candidate biomarkers. A diagnostic nomogram was constructed and validated using receiver operating characteristic (ROC) and calibration curves. Additionally, functional enrichment, immune infiltration, Gene Set Enrichment Analysis (GSEA)/Gene Set Variation Analysis (GSVA), and drug-gene interaction analyses were conducted.

Results:

Three biomarkers (PTCHD1, NEURL1, and ST8SIA5) were identified and validated across both

datasets. All biomarkers were significantly downregulated in tendinopathy samples and negatively correlated with immune cell infiltration. The nomogram model achieved an AUC > 0.7 with good calibration. GSEA revealed their involvement in key signaling pathways, including PI3K-Akt and ECM-receptor interaction. A total of 16 potential therapeutic compounds were identified via drug-gene network analysis.

Conclusion

Our study identifies three robust macrophage polarization-related biomarkers for tendinopathy and provides new insights into its immunopathogenesis and therapeutic targeting.

Keywords

Tendinopathy; Macrophage polarization; Biomarkers; Machine learning; WGCNA; Immune infiltration; Nomogram; Bioinformatics

Introduction

Tendinopathy is a complex tendon pathology resulting from overuse and repetitive tensile stress on muscle fibers, which leads to degenerative changes in tendon collagen fibers. Common types of overuse-induced tendinopathy include those affecting the rotator cuff, medial and lateral epicondyles of the elbow, patellar tendon, gluteal tendon, and Achilles tendon(Baldwin et al., 2022). This condition is characterized by pain, functional decline, and reduced exercise tolerance. Its primary histological and molecular features encompass collagen fiber disorganization, increased microvascularization and sensory innervation, disruption of extracellular matrix (ECM) homeostasis, accumulation of immune cells, elevated levels of inflammatory mediators, and increased apoptosis(Millar et al., 2021; Schulze-Tanzil et al., 2022). Degenerative changes are recognized as risk factors for acute tendon rupture. Since the early 21st century, the global prevalence of tendinopathy has steadily increased, resulting in long-term or permanent functional impairments among both athletic and non-athletic individuals(Kan et al., 2023; Qiao et al., 2025; Wearing et al., 2024). Diagnosis of tendinopathy primarily relies on clinical symptoms, pain assessments, and imaging examinations, while treatment methods include exercise and surgical interventions. For instance, studies by Jefferson et al. have demonstrated pain relief and improved shoulder function following therapeutic exercises in patients diagnosed with rotator cuff tendinopathy. However, the efficacy of these treatments remains uncertain. Currently, the pathogenesis and effective treatment options for tendinopathy are not fully elucidated, with only a

few known biomarkers, such as MACROD1, MIR133A1HG, and C10orf71-AS1. Identifying novel biomarkers for tendinopathy is crucial for enhancing treatment efficiency and prognosis(Kan et al., 2023; Suresh et al., 2022).

Macrophage polarization refers to the process by which macrophages transition from a resting state to an activated state in response to external stimuli, differentiating into M1 and M2 macrophages(Agborbesong et al., 2023; Jia et al., 2025; Tang et al., 2020). Recent research indicates that macrophage polarization plays a critical role in various pathophysiological processes of diseases, such as inflammation, tumorigenesis, and tissue repair. Increasing evidence suggests that macrophages are key players in the pro-inflammatory response, inflammation resolution, and tissue healing in tendinopathy. For instance, Zhu et al. found that macrophages are involved in balancing inflammation and tenosynovial formation during the healing of tendinopathy(L. Zhu et al., 2023). Furthermore, Zheng et al. demonstrated that extracellular vesicles derived from adipose-derived mesenchymal stem cells promote macrophage phagocytosis and M2 polarization in the treatment of tendinopathy(Zheng and Hou, 2021). In mouse models, ASC-Exos-mediated balance of M1/M2 macrophage polarization improved chronic rotator cuff tendinopathy, indicating that modulating M1/M2 macrophage polarization may serve as a novel therapeutic target for chronic rotator cuff degeneration(Quintero et al., 2023; Tang et al., 2024; Yuan et al., 2024). Given the pivotal role of macrophages in tendon repair, investigating macrophage polarization-related genes (MPRGs) specific to tendinopathy presents significant potential as a new target for disease immunotherapy.

This study aims to identify diagnostic and therapeutic biomarkers associated with macrophage polarization in tendinopathy using bioinformatics methods. By further analyzing the functions and molecular mechanisms of these biomarkers, this research provides insights into their roles in disease progression, ultimately contributing to the diagnosis and treatment of tendinopathy.

Materials and Methods

2.1 Data Sources

Training Set: The tendinopathy dataset GSE26051 was obtained from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>), sequenced on the GPL570 platform, containing 46 samples with a distribution of disease (Lesional) to control (Normal) samples at 23:23.

Validation Set: The GSE180836 dataset was obtained from the GEO database, sequenced on the

GPL24676 platform, including 31 samples with a distribution of degenerative to traumatic samples at 15:16.

Macrophage Polarization-Related Genes (MPRGs): A total of 35 MPRGs were downloaded from the MsigDB database (<https://www.gsea-msigdb.org/gsea/msigdb>), including ADAM8, ARSK, ATG4D, CRIPT, EFHD2, FAM241A, IGHMBP2, LIN7C, MYO1F, PINK1, PRDX2, PTGER2, QRIC1, SH2B2, SNHG6, SOAT1, TMEM267, ACSS1, CBX6, COMT, ENO2, FAM135A, FGD2, GLO1, HAUS2, HMGN3, LCORL, LYPLAL1, MAF, MALAT1, OCEL1, RAB10, SFI1, SLC35A1, and ZNF85. Among these, 22 genes were matched to the GSE26051 dataset.

5 supraspinatus tendinopathy (ST) patient samples and 5 control samples were collected with approval from the Hospital Ethics Committee of Sichuan Ba Yi Rehabilitation Center (approval no. CKLL-20231103) and registered at the Chinese Clinical Trial Center (ChiCTR2400083469). All subjects signed informed consent.

2.2 Analysis Methods

2.2.1 Weighted Gene Co-expression Network Analysis(WGCNA) screening of MPRGs module-related genes

WGCNA analysis was conducted on the GSE26051 dataset with MPRG expression as the trait to build a co-expression network, segment modules, and analyze Pearson correlation with MPRGs. The module with the highest correlation was selected as the MPRG-related module(Li et al., 2023; Liu et al., 2020; Shan et al., 2026).

2.2.2 Differential Expression Analysis and DE-MPRG Identification

Differentially expressed genes (DEGs) between disease and control groups were identified in both training and validation sets using the R “limma” package, with screening criteria of $|\log_2\text{-fold change}| (|\log_2\text{FC}|) > 0.5$ and $p\text{-value} < 0.05$ (J. Zhou et al., 2025). The “ggplot2” and “pheatmap” packages were used to create a volcano plot and heatmap, respectively, to visualize differential gene expression(Peng et al., 2025; N. Zhang et al., 2025; Y. Zhang et al., 2025). The intersection of DEGs from the training and validation sets with MPRG module-related genes was taken to obtain DE-MPRGs.

2.2.3 Functional Enrichment Analysis of DE-MPRGs

DE-MPRGs were analyzed for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and

Genomes (KEGG) pathway enrichment using the R "clusterProfiler" package, with a significance threshold of $P < 0.05$.

2.2.4 Machine Learning for Candidate Gene Screening

The "glmnet" package in R was used to conduct Lasso regression analysis on DE-MPRGs, with three-fold cross-validation applied. The "e1071" package in R was then used to implement the SVM-RFE method, determining the importance and ranking of each gene by significance. This process included calculating the error rate and accuracy of each iterative combination, with the combination yielding the lowest error rate selected as the optimal set, followed by extraction of the corresponding genes. The "Boruta" package in R was subsequently applied for Boruta analysis. Finally, the ggvenn package in R was used to compute the intersection of genes identified by the three machine learning algorithms, with intersecting genes designated as candidate genes.

2.2.5 Candidate Gene Expression Validation

Candidate gene expression levels were visualized and validated in both the GSE26051 and GSE180836 datasets using box plots.

2.2.6 Construction of Diagnostic Nomogram for Tendinopathy

A diagnostic nomogram was constructed using the R "rms" package based on biomarker expression levels, and calibration curves and ROC curves were generated to validate the model.

2.2.7 Immune Infiltration Analysis of Biomarkers

The ssGSEA algorithm in the R "GSVA" package was used to calculate immune cell scores for 28 immune cells in all samples, followed by Pearson correlation analysis with biomarkers.

Visualization using the R package "ggcor".

2.2.8 Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA)

We performed Spearman correlation analysis between each biomarker and all genes individually. Biomarkers were ranked by correlation, and Gene Set Enrichment Analysis (GSEA) was conducted using the R package clusterProfiler, with selection criteria of $|\text{NES}| > 1$ and $P_{\text{adjust}} < 0.05$. Additionally, we applied the R packages GSVA and limma to perform Gene Set Variation Analysis (GSVA) based on the MSigDB "KEGG" background gene set (c2.cp.kegg.v7.0.symbols.gmt) to compare enrichment between disease and control groups.

2.2.9 TF-miRNA-mRNA Prediction

Transcription factors (TFs) were predicted for biomarkers using the ChEA3 database, and

miRNAs were predicted using the miRDB database. Only miRNAs with a Target Score greater than 80 were selected. Subsequently, TF-mRNA and miRNA-mRNA networks were constructed using Cytoscape software.

2.2.10 Drug-Gene Interaction Network

Drug compounds interacting with biomarkers were identified through the DSigDB database (<https://dsigdb.tanlab.org/DSigDBv1.0/>), and a network visualization was created using Cytoscape.

2.2.11 RT-qPCR Validation of Biomarker Expression

The expression of biomarkers was further validated via reverse transcription-quantitative polymerase chain reaction (RT-qPCR). Tendon tissue samples were collected from five patients diagnosed with supraspinatus tendinopathy (ST) and five control participants at the Sichuan Ba Yi Rehabilitation Center. Informed consent was obtained from all participants, and the study protocol was approved by the Hospital Ethics Committee of Sichuan Ba Yi Rehabilitation Center (Sichuan Rehabilitation Hospital), under ethical approval number CKLL-20231103.. Total RNA was isolated from each sample individually using TRIzol reagent (Ambion, Austin, TX, USA), strictly adhering to the manufacturer's instructions. Reverse transcription of total RNA into complementary DNA (cDNA) was conducted using the SureScript First-Strand cDNA Synthesis Kit (Servicebio, Wuhan, China), following the standard operating procedure provided(Shao et al., 2025). RT-qPCR reactions were performed using the 2×Universal Blue SYBR Green qPCR Master Mix (Servicebio, Wuhan, China). The specific primer sequences for target genes are listed in Supplementary Table1. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) served as the internal reference gene. Relative gene expression levels were calculated using the 2^{-ΔΔCt} method, and statistical analysis along with data visualization were conducted using GraphPad Prism version 10.1.2(Liu et al., 2025).

Table1 . PCR primers sequence

Primer		Sequence
PTCHD1	F	GCTGCTGACCTGACTTCTGG
PTCHD1	R	CCAGGTTACAGCAGGTAGC
NEURL1	F	TGGAGAACCGGAAGATGACC
NEURL1	R	GGTGCTGTTGCTGATGTTGG

ST8SIA5	F	CAGCAGTTCACCAACCTGAC
ST8SIA5	R	AGGTCAGGTTGGTGAAGTGC
internal reference H-GAPDH	F	ATGGGCAGCCGTTAGGAAAG
internal reference H-GAPDH	R	AGGAAAAGCATCACCCGGAG

2.2.12 Statistical analysis

All analyses were conducted in R version 4.2.1. For significance testing of various values (e.g., gene expression levels, infiltration proportions, and various characteristic values), the Wilcoxon rank-sum test or T-test was used to compare differences between two sample groups. In the visualizations, "ns" denotes $P > 0.05$, * indicates $P < 0.05$, ** indicates $P < 0.01$, *** indicates $P < 0.001$, and **** indicates $P < 0.0001$.

Results

3.1 Screening of MPRG-Related Genes through WGCNA

3.1.1 Screening of MPRG-Related Genes through WGCNA

Sample clustering was conducted on the GSE26051 dataset to identify and exclude outliers, ensuring analytical accuracy (Figure 1-A).

3.1.2 Determination of Soft Threshold

To ensure that gene-gene interactions closely adhere to a scale-free distribution, we used the expression matrix of all genes as the input data, applied computational analysis, and determined the soft threshold. As shown in Figure 1-B, when the scale-free topology R^2 on the y-axis approaches the threshold of 0.85 (indicated by the red line) and the mean connectivity on the right plot trends toward zero, the network approximates a scale-free distribution. At this point, the optimal soft threshold β is 13; therefore, we selected 13 as the optimal soft threshold (β).

3.1.3 Construction of Co-Expression Network

The core of constructing a co-expression matrix is to categorize tens of thousands of genes from the input expression matrix into a few or several dozen modules. This is achieved by calculating the adjacency between genes, followed by determining gene similarity based on this adjacency, and subsequently deriving the dissimilarity coefficient between genes, which forms the basis for generating a gene dendrogram. Following the criteria of the dynamic tree cutting algorithm, the minimum number of genes per module was set to 80, resulting in a total of 9 modules, as shown in Figure 1-C.

3.1.4 Key Module Selection

Using the expression levels of MPRGs as traits, we performed a Pearson correlation analysis between these traits and the nine modules obtained from WGCNA analysis. As shown in Figure 1-D, we identified the magenta module as the key module due to its highest positive correlation with MPRG expression levels, with a correlation coefficient of 0.8 ($P < 0.001$). This magenta module contains 865 genes and is defined as the MPRG-associated module.

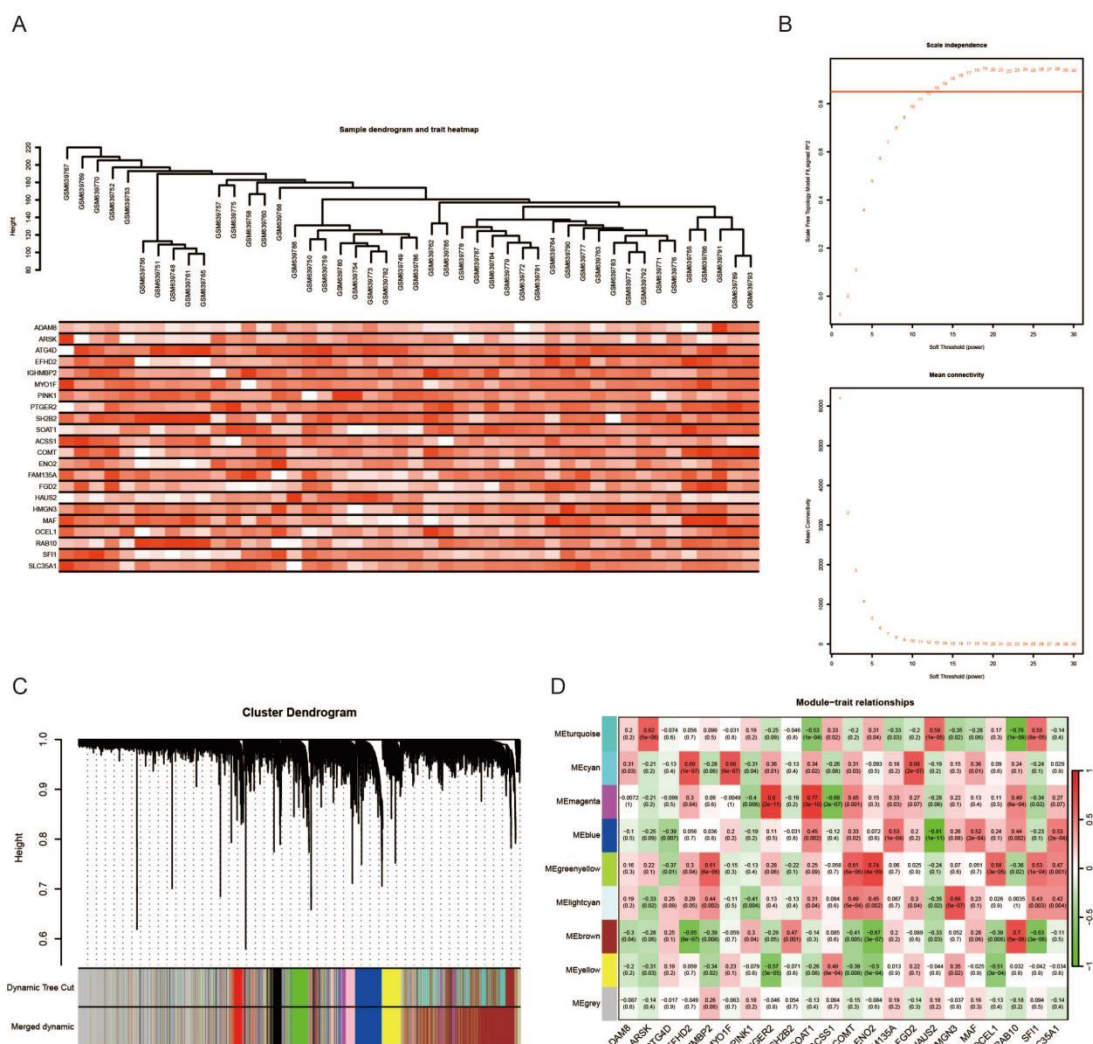


Figure 1. Screening of MPRG-related genes by WGCNA. (A) Sample clustering dendrogram (based on GSE26051 dataset). (B) Determination of the soft-thresholding power (β). In all panels, the horizontal axis represents the weight parameter, power. In the left plot, the vertical axis indicates the scale-free fit index, signed R^2 , where a higher squared correlation coefficient suggests a closer approximation to a scale-free network. In the right plot, the vertical axis represents the mean adjacency function of all genes within the corresponding gene modules. (C) Dynamic dendrogram of co-expression modules. The upper portion below the dendrogram,

labeled "Dynamic Tree Cut," represents modules before merging, while the lower portion, labeled "Merged Dynamic," shows modules after merging. (D) Heatmap of module–trait correlations.

3.2 Differential Expression Analysis

3.2.1 Differential Expression Analysis of GSE26051 Dataset

Using the "limma" package in R, we compared the gene expression levels between Lesional and Control samples in the GSE26051 dataset (Lesional vs. Control). Differentially expressed genes (DEGs) were identified with the following criteria: $|\log_2FC| > 0.5$ and $P\text{-value} < 0.05$. A total of 1,781 DEGs were identified, including 864 upregulated and 917 downregulated genes. Figure 2-A displays a volcano plot of these differential genes, created using the "ggplot2" package in R, illustrating the differences between the groups. Figure 2-B presents a heatmap, generated using the "pheatmap" package in R, showing the top 20 upregulated and top 20 downregulated DEGs, ranked by $|\log_2FC|$, across Control and Lesional samples.

3.2.2 Differential Expression Analysis of GSE180836 Dataset

Using the "limma" package in R, we analyzed the differential gene expression levels between Degenerative and Control samples in the GSE180836 dataset (Degenerative vs. Control). Differentially expressed genes (DEGs) were identified based on the criteria $|\log_2FC| > 0.5$ and $P\text{-value} < 0.05$, resulting in a total of 364 DEGs, including 214 upregulated and 150 downregulated genes. Figure 2-C displays a volcano plot created with the "ggplot2" package in R, illustrating the differential expression between groups. Figure 2-D shows a heatmap generated with the "pheatmap" package in R, highlighting the expression of the top 20 upregulated and top 20 downregulated DEGs, ranked by $|\log_2FC|$, across Control and Degenerative samples.

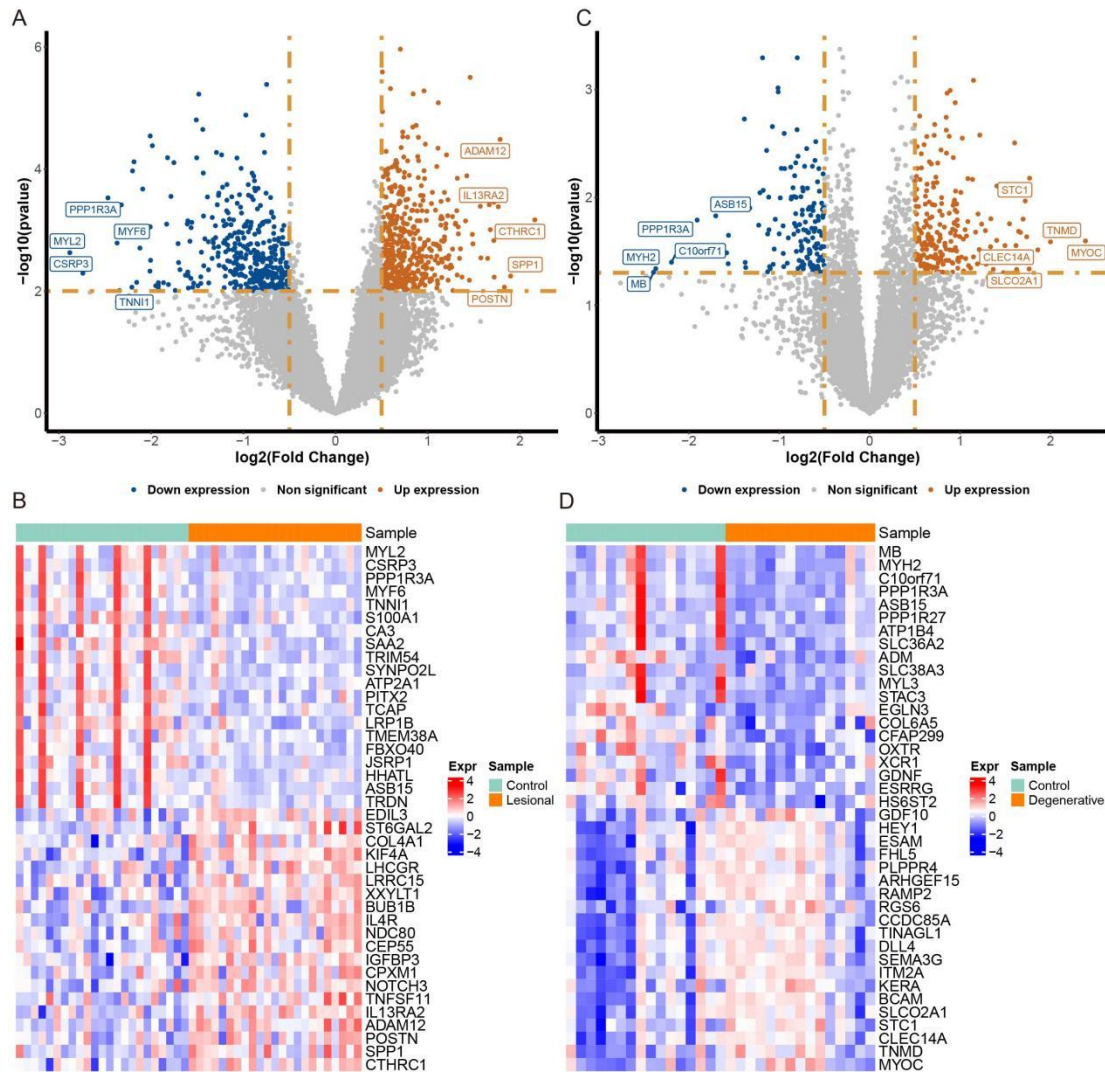


Figure2. Differential expression analysis of the GSE26051 and GSE180836 datasets.(A) Volcano plot of DEGs in GSE26051 (Lesional vs. Control).(B) Heatmap of the top 20 DEGs in GSE26051.(C) Volcano plot of DEGs in GSE180836 (Degenerative vs. Control).(D) Heatmap of the top 20 DEGs in GSE180836.

3.3 Identification and Analysis of DE-MPRGs

3.3.1 Selection of DE-MPRGs

The 865 module-related genes associated with MPRGs obtained in section 2.1 were intersected with the significantly up- and down-regulated differential genes from the two datasets in section 2.2. Visualization was performed using the R package ggvenn, as shown in Figure 3-(A - B), resulting in five intersecting genes, defined as DE-MPRGs.

3.3.2 Functional Enrichment Analysis of DE-MPRGs

To understand the molecular functions and mechanisms of action of the DE-MPRGs, we

performed GO and KEGG enrichment analyses on the five DE-MPRGs using the R package clusterProfiler to identify shared gene functions and associated pathways.

The Gene Ontology (GO) system comprises three categories: Biological Processes (BP), Cellular Components (CC), and Molecular Functions (MF). Using a p-value threshold of < 0.05 , we identified 242 enriched terms, including 212 significantly enriched terms in BP, 13 in CC, and 17 in MF. Visualization of the top 8 enriched terms in each category, ranked by ascending p-value, was conducted using the R package ggpubr. The results are shown in Figure 3-C.

The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a commonly used database for pathway research. KEGG pathway analysis involves annotating pathways for all identified DE-MPRGs using the KEGG database, aiming to explore the primary metabolic and signal transduction pathways in which these DE-MPRGs are involved.

In this study, KEGG pathway enrichment analysis was performed with a p-value threshold of < 0.05 , identifying five significantly enriched pathways. The relationships among these five pathways were visualized using the R package enrichplot, with the enrichment results shown in Figure 3-D.

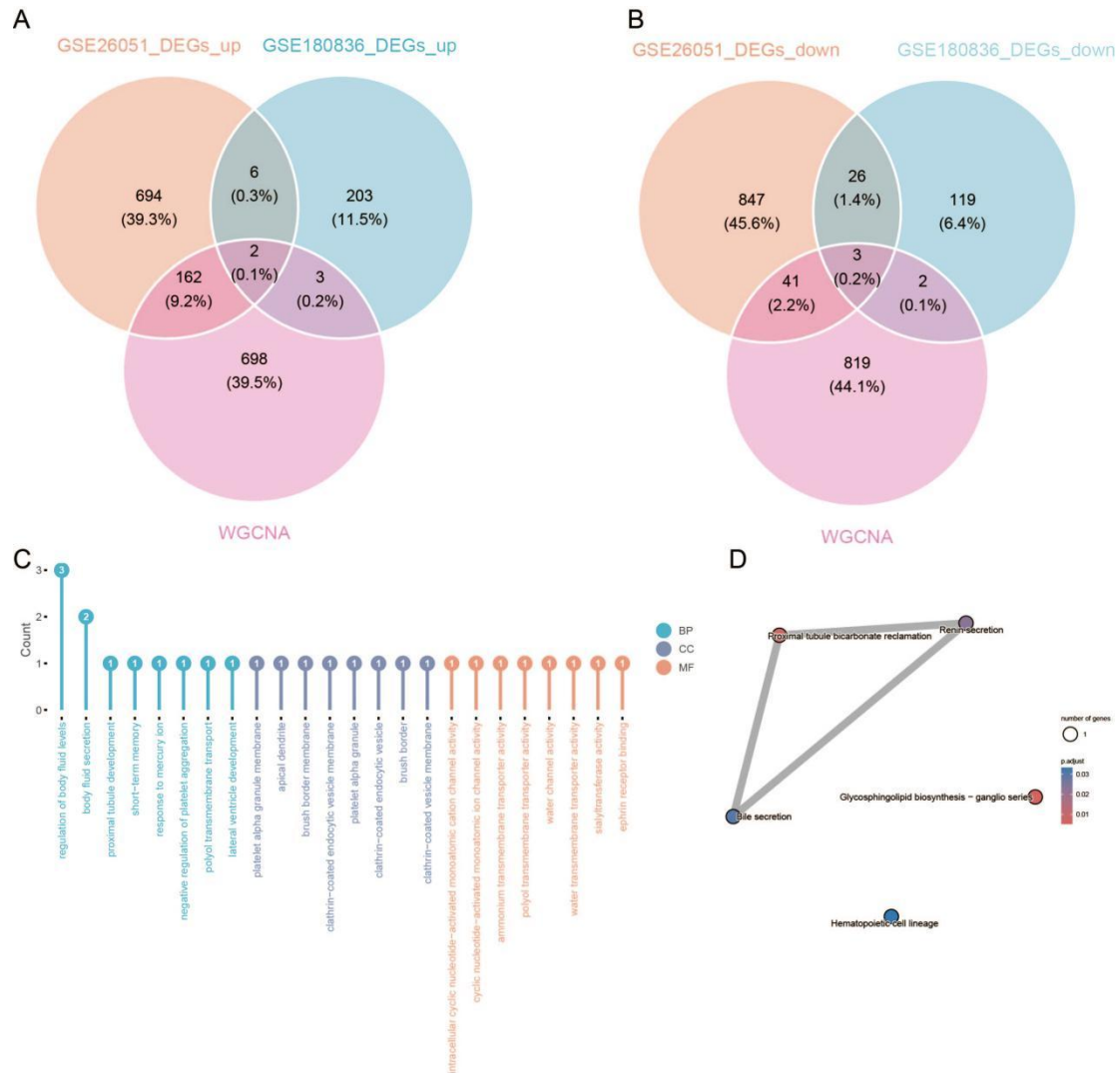


Figure 3. Identification and functional analysis of DE-MPRGs. (A-B) Venn diagram of DE-MPRGs. Intersection of 865 MPRG-associated module genes (from WGCNA) with differentially expressed genes (DEGs) from GSE26051 and GSE180836. The five overlapping genes are defined as DE-MPRGs.(C) GO enrichment analysis of DE-MPRGs. The horizontal axis shows GO terms (top 8 terms by ascending p-value in each category: Biological Processes, BP; Cellular Components, CC; Molecular Functions, MF). The vertical axis (Count) indicates the number of genes associated with each term.(D) KEGG pathway enrichment analysis of DE-MPRGs. Each square represents a significantly enriched pathway (p < 0.05). Square size reflects the number of enriched genes in the pathway; color indicates p-value (darker = smaller p-value).(The analysis was based on the KEGG (Kyoto Encyclopedia of Genes and Genomes) database. Pathway images were obtained from the KEGG database (<http://www.kegg.jp>).Citation: Kanehisa, M., et al. (2016); Kanehisa, M. and Goto, S. (2000))

3.4 Machine Learning for Candidate Gene Selection

Biomarker selection based on DE-MPRGs was conducted using the Lasso, Boruta, and Support Vector Machine-Recursive Feature Elimination (SVM-RFE) algorithms. The intersection of genes identified by these three algorithms was taken to obtain the final biomarkers.

3.4.1 Lasso Regression

Lasso analysis was conducted using the glmnet package to perform logistic regression, with a three-fold cross-validation approach to determine the optimal turning/penalty parameter. LASSO analysis of the DE-MPRGs was performed using the R package glmnet. Three-fold cross-validation results are shown in Figure 4-(A-B) : the dashed line on the left in the left panel indicates the position with the minimum cross-validation error, and the corresponding x-coordinate ($\log(\text{Lambda})$) was identified based on this point (Lambda.min). The top x-axis shows the number of genes in the model, allowing identification of the optimal $\log(\text{Lambda})$ value, and the right panel displays the selected genes with their coefficients, as well as the proportion of residual variance explained by the model. The Lasso model with the minimum error was selected as the result of the Lasso analysis, yielding five genes through Lasso regression.

3.4.2 Boruta Analysis

The Boruta algorithm is a wrapper method based on the random forest classification algorithm. Random forest is a relatively fast classification algorithm that typically requires minimal parameter tuning and provides numerical estimates of feature importance. This ensemble method classifies data by voting among multiple unbiased weak classifiers—decision trees. Importance can be measured using the Z-score, calculated by dividing the mean decrease in loss by the standard deviation. Boruta applies the same principles as the random forest classifier by incorporating randomness into the system, collecting results from random sample sets to reduce the misleading effects of random fluctuations and correlation.

Using the Boruta package in R, we analyzed the DE-MPRGs, resulting in the identification of four genes: AQP1, PTCHD1, NEURL1, and ST8SIA5, as shown in Figure 4-C.

3.4.3 SVM-RFE

Support Vector Machine Recursive Feature Elimination (SVM-RFE) is an effective feature selection technique. This method is a sequential backward selection algorithm based on the SVM's maximum-margin principle. It involves training a model on sample data, ranking each feature

based on a scoring criterion, and eliminating the feature with the lowest score. The model is then retrained with the remaining features, and the process is iteratively repeated until the desired number of features is selected.

We applied the SVM-RFE method using the `e1071` and `caret` packages in R to obtain both the importance and ranking of DE-MPRGs (differentially expressed marker gene profiles) and to record the error rate and accuracy of each feature subset. The optimal feature combination was determined by selecting the subset with the lowest error rate. As shown in Figure 4-D, when the number of genes is reduced to four, the error rate reaches its minimum at 0.237. These four genes—CD9, PTCHD1, NEURL1, and ST8SIA5—are selected as candidate genes.

3.4.4 Candidate Gene Screening

We used the "ggvenn" package in R to identify the intersection of genes selected by the three machine learning algorithms. As shown in Figure 4-E, all three algorithms consistently identified three common genes: PTCHD1, NEURL1, and ST8SIA5. These three genes were thus selected as candidate genes.

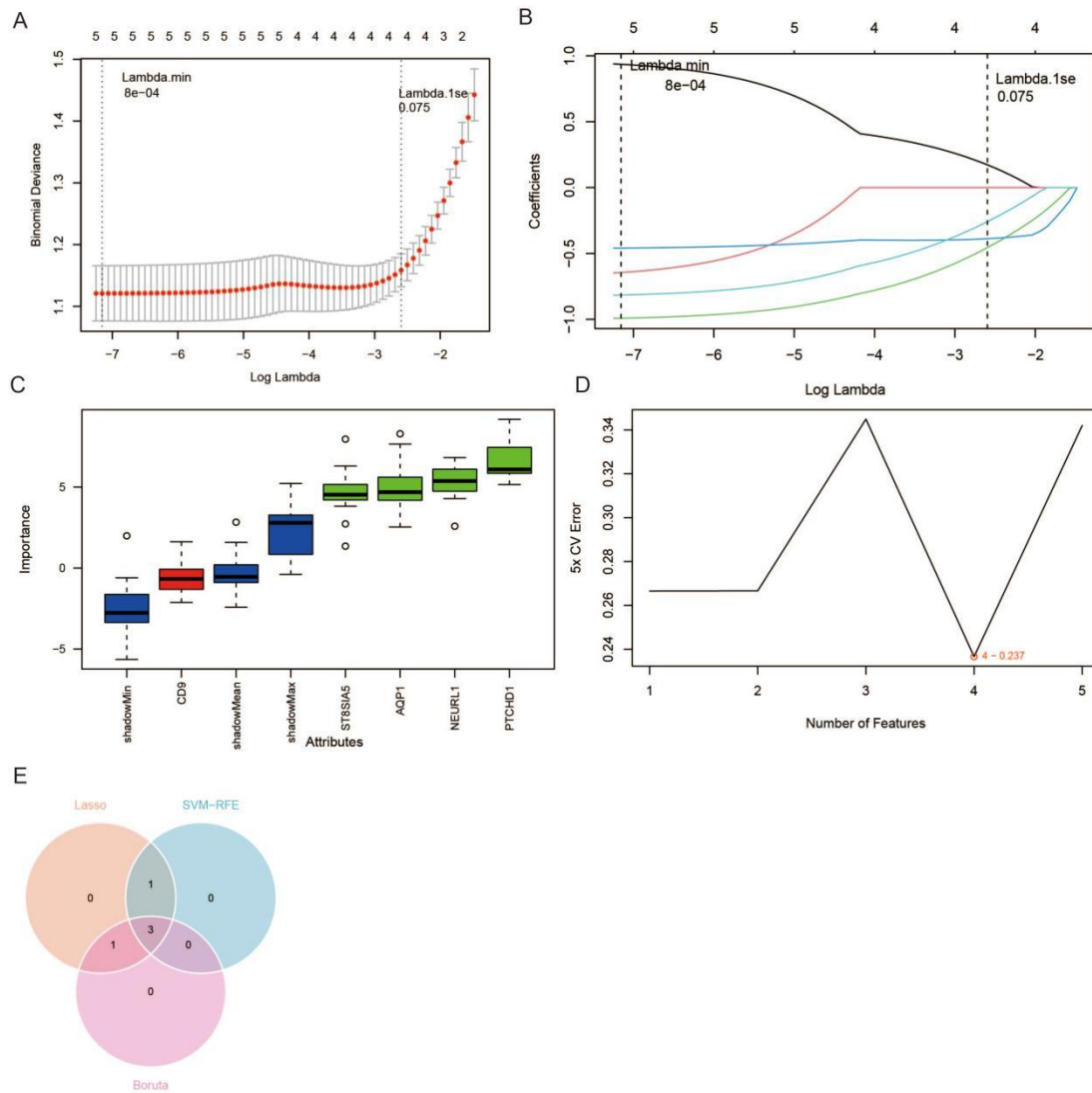


Figure 4. Machine learning-based selection of candidate genes from DE-MPRGs. (A-B) Lasso regression analysis. A: The horizontal axis represents $\log(\text{Lambda})$, and the vertical axis represents the cross-validation error. B: The horizontal axis represents $\log(\text{Lambda})$, and the vertical axis represents the coefficients of the genes. (C) Boruta algorithm. Green boxes: confirmed important genes; yellow boxes: tentative features; blue boxes: minimum, average, and maximum Z-scores of shadow attributes. Boruta selected 4 genes (AQP1, PTCHD1, NEURL1, ST8SIA5). (D) SVM-RFE (Support Vector Machine – Recursive Feature Elimination). (E) Venn diagram of genes selected by the three algorithms.

3.5 Biomarker Screening

To further screen for biomarkers, box plots were used to visualize the expression differences of candidate genes between disease and control samples, followed by intergroup T-tests. As shown in Figure 5-(A-B), all three candidate genes were significantly downregulated in both the training

and validation sets.

Based on gene expression results, PTCHD1, NEURL1, and ST8SIA5 exhibited consistent and significantly differential expression in both the training and validation sets. Therefore, PTCHD1, NEURL1, and ST8SIA5 were retained as biomarkers.

3.6 Immune Infiltration Analysis of Biomarkers

To further investigate immune cell characteristics in Control and Lesional samples, this study applied the ssGSEA algorithm from the GSVA package in R to calculate immune cell scores for each sample. Scores for 28 types of immune cells were computed across all samples, and the “ggplot2” package in R was used to create a bar plot showing the infiltration abundance percentages of these 28 immune cell types in Control and Lesional samples (Figure 5-C).

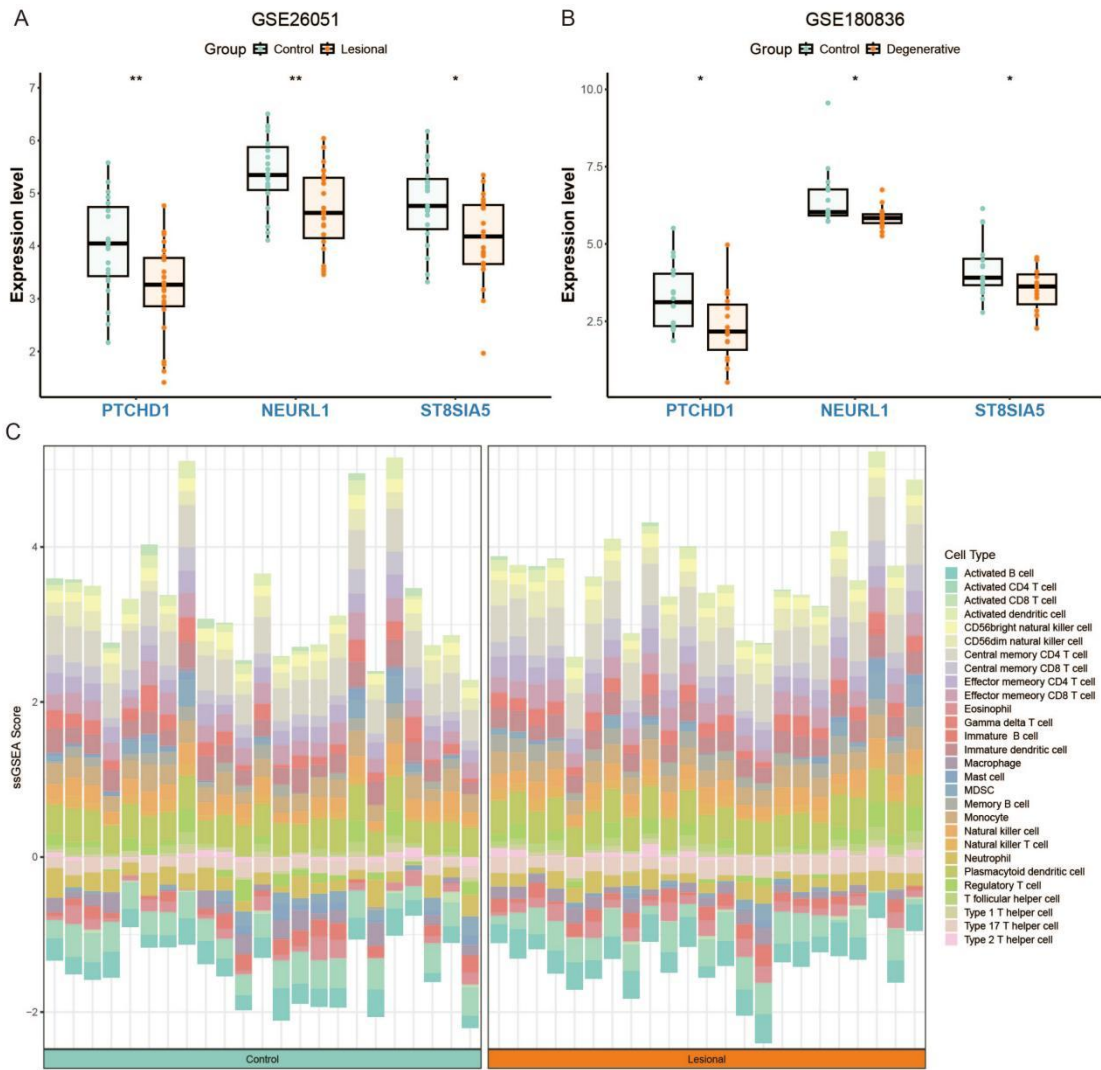


Figure 5. Biomarker expression validation and immune infiltration analysis. (A-B) Boxplots of candidate gene expression (training and validation sets). Upper panel: training set; lower panel:

validation set. Horizontal axis: genes (PTCHD1, NEURL1, ST8SIA5); vertical axis: expression level. Significance levels: ns (not significant), * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. All three genes are significantly downregulated in disease samples compared to controls in both datasets, and are retained as biomarkers. (C) Immune cell infiltration abundance (ssGSEA). Stacked bar plot showing the percentage of 28 immune cell types in Control vs. Lesional samples. Each bar represents the mean proportion of a specific immune cell type across samples.

To explore the relationship between biomarkers and differentially expressed immune cells, Pearson correlation analyses were performed between biomarkers and immune cells in both disease and control group samples. Correlation plots were generated using the "ggpubr" package in R. As shown in Figure 6-A, in the disease group, the highest positive correlation was observed between the NEURL1 gene and Activated B cells, with a correlation coefficient of 0.71 ($P < 0.001$), while the highest negative correlation was between the ST8SIA5 gene and Type 2 T helper cells, with a correlation coefficient of -0.61 ($P = 0.01$). In Figure 6-B, in the control group, the strongest positive correlation was observed between the ST8SIA5 gene and Type 17 T helper cells, with a correlation coefficient of 0.53 ($P = 0.01$), and the strongest negative correlation was between the NEURL1 gene and Activated dendritic cells, with a correlation coefficient of -0.59 ($P < 0.01$).

3.7 Construction of Diagnostic Nomogram

Using the "rms" package in R, we constructed a Nomogram model based on the expression levels of the three biomarkers. In the model, each gene's expression level corresponds to a specific score, and the sum of these individual scores yields a total score (Total Point). This total score is then used to predict the probability of disease diagnosis; higher scores indicate a greater likelihood of a disease sample. The prediction results are presented as a nomogram plot, as shown in Figure 7-C. To evaluate the accuracy of the nomogram model, a calibration curve corresponding to the nomogram was plotted. As shown in Figure 7-D, the slope of the curve approaches 1, indicating that the nomogram model is accurate and demonstrates high diagnostic efficacy in identifying disease samples.

To further validate the effectiveness of the predictive model, we plotted the ROC curve for the nomogram model and its associated traits. As shown in Figure 7-E, the nomogram model achieved

an AUC value greater than 0.7, indicating a good predictive capability.

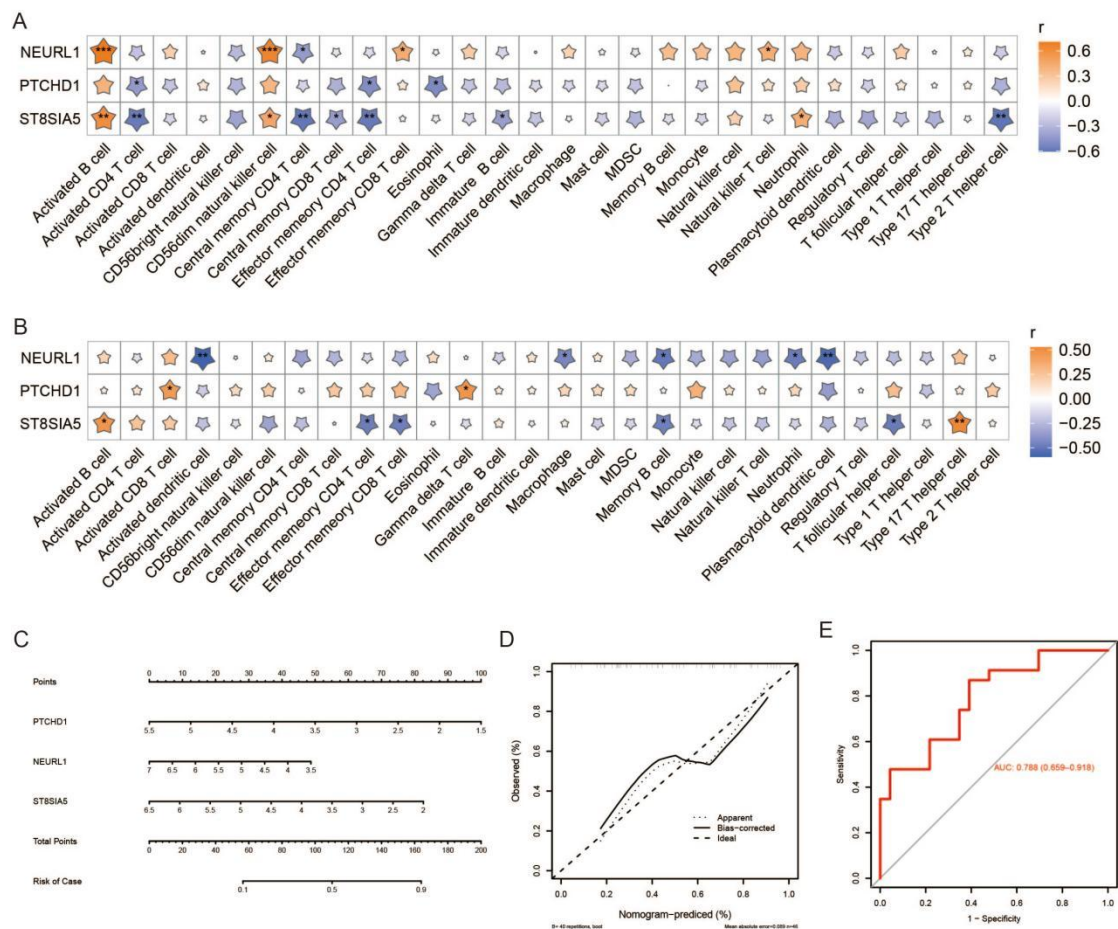


Figure 6. Correlation of biomarkers with immune cells and construction of a diagnostic nomogram.

(A) Correlation heatmap – disease group. Pearson correlations between the three biomarkers (PTCHD1, NEURL1, ST8SIA5) and 28 immune cell types in disease samples. Strongest positive correlation: NEURL1 with Activated B cells ($r = 0.71$, $p < 0.001$). Strongest negative correlation: ST8SIA5 with Type 2 T helper cells ($r = -0.61$, $p = 0.01$). (B) Correlation heatmap – control group. Pearson correlations in control samples. Strongest positive correlation: ST8SIA5 with Type 17 T helper cells ($r = 0.53$, $p = 0.01$). Strongest negative correlation: NEURL1 with Activated dendritic cells ($r = -0.59$, $p < 0.01$). (C) Nomogram model. Based on expression levels of PTCHD1, NEURL1, and ST8SIA5. Points are assigned to each gene; the total point sum predicts the probability of disease diagnosis (higher total point $\rightarrow$ higher likelihood of disease). (D) Calibration curve of the nomogram. Horizontal axis: predicted probability; vertical axis: actual

probability. A slope close to 1 indicates high predictive accuracy. (E) ROC curve of the nomogram. The area under the curve (AUC) is >0.7 , demonstrating good diagnostic performance.

3.8 GSEA and GSVA Analysis

3.8.1 Gene Set Enrichment Analysis (GSEA)

To gain further insight into pathway differences associated with varying expression levels of different biomarkers in disease progression, we used the GSVA package in R, applying the "KEGG" background gene set from the MSigDB database (c2.cp.kegg.v7.0.symbols.gmt). In the training set, correlations between each gene and the biomarkers were calculated and ranked by correlation strength, followed by GSVA enrichment analysis. Pathways with $|NES| > 1$ and $P_{\text{adjust}} < 0.05$ were considered significantly activated or suppressed. Details are provided in Figures 7-A, 7-B, and 7-C.

3.8.2 Gene Set Variation Analysis (GSVA)

Gene Set Variation Analysis (GSVA) is a non-parametric, unsupervised method primarily used to assess gene set enrichment in microarray and transcriptomic data. GSVA converts the gene expression matrix across different samples into a gene set expression matrix, allowing for the evaluation of whether specific metabolic pathways are enriched across samples. Unlike individual gene expression differences, GSVA identifies pathways with significant inter-sample variation, offering greater biological relevance.

To gain further insights into pathway variations throughout disease progression, we employed the GSVA and limma packages in R. Based on the "KEGG" background gene sets from the MSigDB database (c2.cp.kegg.v7.0.symbols.gmt), we conducted GSVA enrichment analyses on disease and control groups within the training set. Differential analysis was then performed on each pathway entry (disease group vs. control group) using the limma package, with thresholds set at $|t| > 2$ and $P\text{-value} < 0.05$. A total of 39 pathways showed significant differences. Pathways with the most notable differences were ranked and displayed according to GSVA differential analysis t-values, as shown in Figure 7-D. Additionally, a heatmap of scores for significantly different pathways is presented in Figure 7-E.

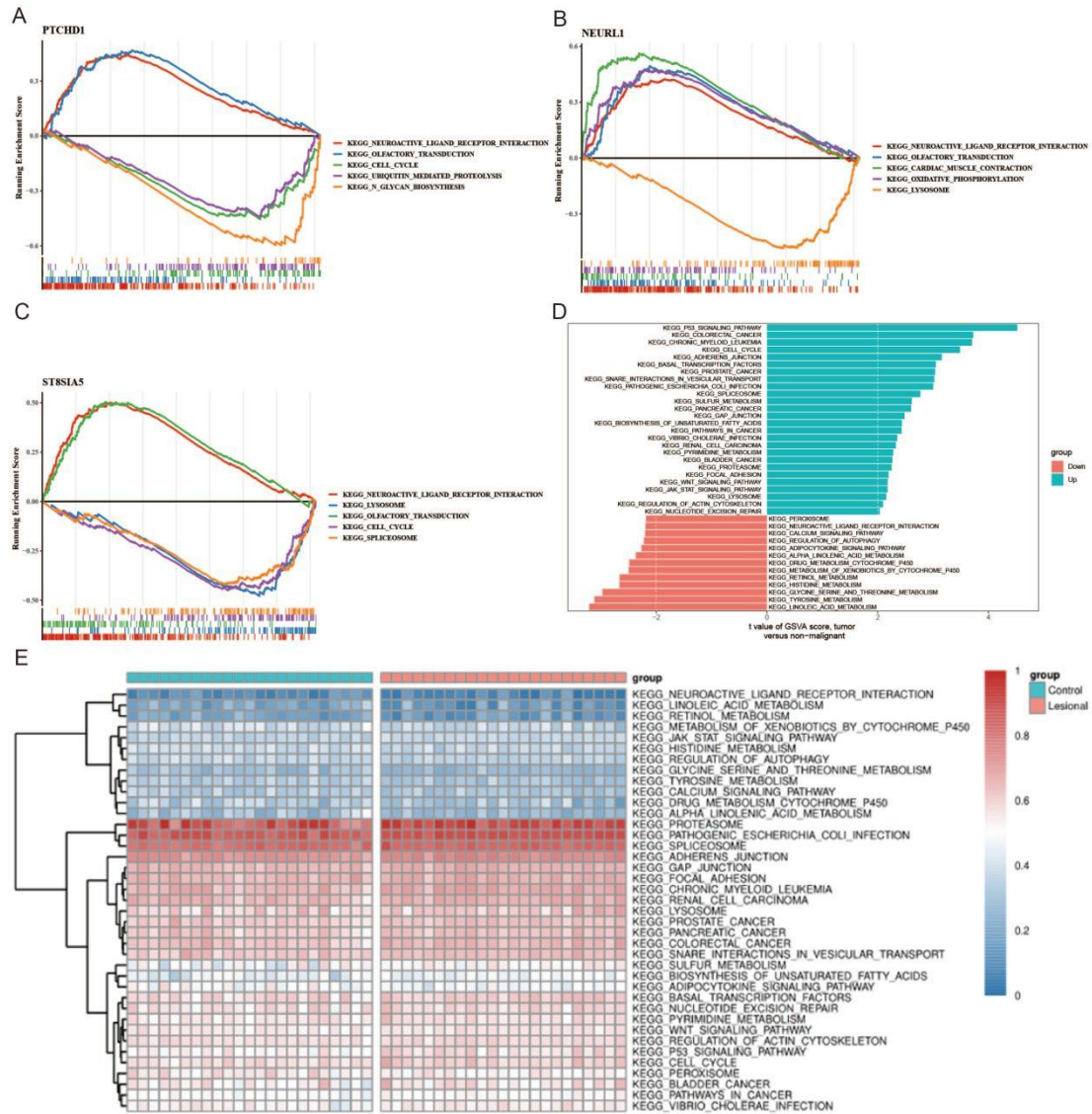


Figure 7. GSEA and GSVA analyses of the identified biomarkers. (A) GSEA for PTCHD1. Enrichment plots of significantly activated or suppressed KEGG pathways ($|NES| > 1$, $P_{\text{adjust}} < 0.05$) based on correlation with PTCHD1 expression. (B) GSEA for ST8SIA5. Significantly enriched KEGG pathways ($|NES| > 1$, $P_{\text{adjust}} < 0.05$) associated with ST8SIA5 expression levels. (C) GSEA for NEURL1. Significantly enriched KEGG pathways ($|NES| > 1$, $P_{\text{adjust}} < 0.05$) associated with NEURL1 expression. (D) GSVA differential analysis (disease vs. control). (E) GSVA heatmap of significantly different pathways. (Data source: KEGG (Kyoto Encyclopedia of Genes and Genomes) database. Pathway images were obtained from the KEGG database (<http://www.kegg.jp>). Citation: Kanehisa, M., et al. (2016); Kanehisa, M. and Goto, S. (2000).)

3.9 TF/miRNA-mRNA Regulatory Network Construction

Transcription factors (TFs) are key participants in gene regulation, impacting cellular functions by binding to DNA and modulating gene transcription levels. Predicting TFs for mRNAs provides a more comprehensive understanding of the structure and dynamics of gene regulatory networks, revealing interactions between genes.

To gain deeper insights into interactions within the gene regulatory network, this study utilized the ChEA3 database (Version 3) to predict TFs for selected biomarkers, identifying 36 TFs associated with 3 biomarkers. Network visualization was performed using Cytoscape software (Version 3.10.2), as shown in Figure 8-A. The resulting network comprises 39 nodes and 37 edges.

The miRDB database (miRDB Version 6.0) was used to predict miRNAs potentially regulating the biomarker SPARC, filtering for miRNAs with an interaction score >80. Network visualization was conducted using Cytoscape software (Version 3.10.2), as shown in Figure 8-B. The resulting network consists of 25 nodes and 24 edges.

3.10 Drug-Gene Interaction Analysis

Potential therapeutic drugs targeting biomarkers were identified through the DSIgDB database. A total of 16 drugs were predicted to interact with the three biomarkers, resulting in 18 interaction pairs. The drug-biomarker interaction network was constructed and visualized in Cytoscape (Figure 8-C).

3.11 RT-qPCR validation of biomarker expression

Within the training cohort, significant differences in gene expression were observed for both biomarkers between ST and control samples. Specifically, PTCHD1, NEURL1 and ST8IA5 expression were reduced in ST samples relative to controls (Figures 8-D). To further validate these findings, RT-qPCR was performed to assess their expression in clinical samples. Consistent with the transcriptomic analysis, RT-qPCR confirmed that PTCHD1, NEURL1 and ST8IA5 expression were significantly decreased in ST samples compared with controls ($p < 0.05$). These concordant results across independent datasets and experimental validation further supported the robustness and reproducibility of the identified expression patterns.

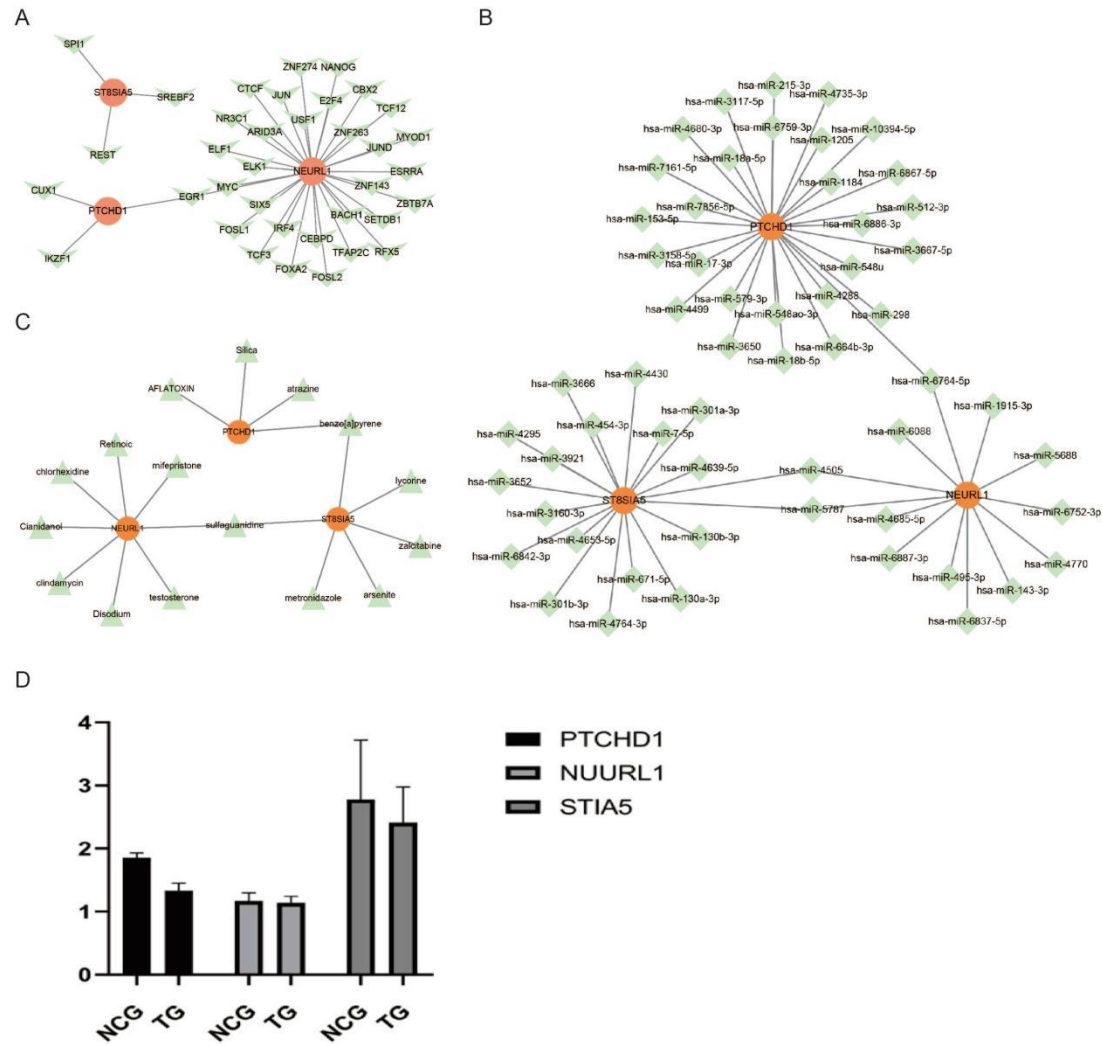


Figure 8. TF/miRNA-mRNA regulatory network, drug-gene interactions, and RT-qPCR validation. (A) mRNA-TF network. Circles: mRNAs (PTCHD1, NEURL1, ST8SIA5). Arrows: transcription factors (TFs) predicted by ChEA3 database. The network contains 39 nodes and 37 edges. (B) miRNA-mRNA network. Circles: mRNAs; diamonds: miRNAs predicted by miRDB database (interaction score >80). The network contains 25 nodes and 24 edges. (C) Drug-gene interaction network. Circles: biomarkers; triangles: potential therapeutic drugs predicted by DSigDB database. A total of 16 drugs with 18 interaction pairs are shown. (D) RT-qPCR validation of biomarker expression in clinical samples. Bar plot (or corresponding graph) showing relative expression levels of PTCHD1, NEURL1, and ST8SIA5 in ST (lesional) vs. control samples. All three biomarkers were significantly downregulated in ST samples compared to controls (* $p < 0.05$), consistent with transcriptomic findings from the training cohort (Figure 5A).

Discussion

By integrating WGCNA, differential-expression analysis and machine learning, we identified PTCHD1, NEURL1 and ST8SIA5 as biomarkers that are closely associated with macrophage polarization in human tendinopathy. Across two independent cohorts we found that all three genes were reproducibly down-regulated and displayed robust diagnostic accuracy ($AUC > 0.76$), supporting their potential as minimally invasive molecular signatures.

We first examined the published literature to understand why these genes might matter. Previous studies have established that PTCHD1 is an X-linked synaptic-membrane protein implicated in neurodevelopmental disorders through regulation of neuronal activity and distal enhancers . Although PTCHD1 has not previously been associated with macrophages or tendinopathy, we noted that recent work shows mouse bone-marrow-derived macrophages express PTCHD1 mRNA under both resting and LPS-stimulated conditions, and that this expression is dependent on RasGEF1b—a factor known to participate in inflammatory responses and the production of inflammatory mediators—suggesting that PTCHD1 may be embedded within the macrophage inflammatory gene-regulatory network(Bugda Gwilt et al., 2019; Burkett et al., 2023; Chang et al., 2023).

We next considered NEURL1 published work indicates that NEURL1 is an E3 ubiquitin ligase that promotes endocytosis and positively activates Notch signalling via mono-ubiquitination of Jagged-1; inhibition of the NEURL1–Notch axis lowers Notch activity, thereby suppressing M1 macrophage activation and fostering M2-mediated tissue repair(J. Zhu et al., 2023a, 2023b, 2023c).

For ST8SIA5, we found that it is the rate-limiting enzyme in the synthesis of polysialylated neural cell-adhesion molecule (PSA-NCAM). PSA chains weaken cell–cell and cell–matrix adhesion through negative-charge repulsion, promoting cell migration and neurite outgrowth(El Maarouf et al., 2012). In tumour-microenvironment studies ST8SIA5 is classified as a member of the sialyltransferase family(Xiao et al., 2013); high-level sialylation can inhibit macrophage pro-inflammatory activity and skew macrophages toward an M2-like phenotype by engaging Siglec receptors on their surface.

Although direct evidence demonstrating the roles of these three genes in macrophage polarization and tendinopathy is still lacking, we reasoned that the signalling pathways in which they participate have been shown—directly or indirectly—to exert critical regulatory effects on

macrophage activation(Chen et al., 2026; N. Zhang et al., 2025). We therefore propose that down-regulation of PTCHD1, NEURL1 and ST8SIA5 promotes a pro-inflammatory M1 phenotype while simultaneously suppressing the tissue-reparative M2 phenotype, thereby contributing to the chronic inflammation and tissue damage characteristic of tendinopathy(Blanco et al., 2025; Jeong et al., 2025).

To explore possible mechanisms, we performed GSEA and observed that PTCHD1, NEURL1 and ST8SIA5 may influence macrophage polarization—and thus tendinopathy progression and tissue repair—through distinct signalling cascades. The biomarkers were significantly enriched in several key pathways, most notably PI3K–Akt signalling and extracellular-matrix (ECM) receptor interactions(X. Wang et al., 2025; Z. Wang et al., 2025). ECM components such as collagen, fibronectin and laminin engage trans-membrane receptors including integrins, rapidly inducing autophosphorylation of focal-adhesion kinase (FAK). Activated FAK recruits and activates PI3K, generating PIP3 and promoting Akt membrane translocation and full phosphorylation. This sequential axis governs survival, proliferation, migration, differentiation and matrix remodelling, and has been validated in diverse pathophysiological contexts including osteogenesis, tendon/ligament repair, tumour drug resistance and fibrosis (Paneru and Hill, 2023).

We also noted that enrichment results linked the three genes to the metabolic states of aerobic glycolysis and oxidative phosphorylation. Macrophage polarization is accompanied by striking, opposing metabolic reprogramming: the M1 phenotype, triggered by LPS/IFN- γ , relies almost exclusively on aerobic glycolysis(M. de-Brito et al., 2020; H. Zhang et al., 2025). The TCA cycle is “broken” at succinate dehydrogenase and isocitrate dehydrogenase, leading to succinate accumulation that stabilises HIF-1 α and persistently up-regulates glycolytic enzymes, thereby driving expression of pro-inflammatory genes such as IL-1 β , IL-6 and iNOS. Inhibition of glycolysis is sufficient to attenuate ROS/NO production and cytokine secretion (Damen et al., 2022; Koenis et al., 2018). In contrast, the M2 phenotype, induced by IL-4/IL-13, depends predominantly on oxidative phosphorylation (OXPHOS) and fatty-acid oxidation (FAO)(Fang et al., 2025). The TCA cycle remains intact, the electron-transport chain is highly active, and the STAT6–PPAR- γ –PGC-1 β axis promotes mitochondrial biogenesis; blocking FAO or OXPHOS markedly suppresses M2 markers such as Arg-1 and CD206 (Koo and Garg, 2019; Y. Zhou et al., 2025). Although low-level glycolysis is retained for metabolic plasticity, ATP depletion disrupts

M2 differentiation only when both glycolysis and the Krebs cycle are simultaneously inhibited, demonstrating that OXPHOS/FAO constitutes the core of M2 polarization(Luque-Campos et al., 2021). We therefore propose that enrichment of the PI3K–Akt pathway may indirectly modulate macrophage energy metabolism by orchestrating cellular metabolic reprogramming, thereby promoting a pro-inflammatory phenotype.

We next asked whether these biomarkers reflect the local immune milieu. Immune-infiltration analysis revealed that expression levels of PTCHD1, NEURL1 and ST8SIA5 correlated with infiltration scores of multiple immune-cell subsets, including activated B cells and Th2 cells. Integrated pathway enrichment showed that the PI3K–Akt axis acts as a “positive gatekeeper” for immune-cell infiltration: BCR signalling via PI3K δ –Akt drives proliferation, class switching and intratumoral aggregation of activated B cells—effects rapidly reversed by the PI3K δ inhibitor idelalisib(Zhong et al., 2023). IL-4R–STAT6 and Akt form a positive-feedback loop in which Akt phosphorylation boosts GATA3 expression, driving Th2 differentiation and IL-4/5/13 secretion; the mTOR inhibitor INK128 simultaneously suppresses both Th2 cytokines and intratumoral Th2 numbers(Vebr et al., 2023). The infiltration of activated B cells and Th2 cells is significantly correlated with PI3K–Akt activity, underscoring this axis as a critical signalling node that maintains their function and population. This pattern suggests that down-regulation of PTCHD1, NEURL1 and ST8SIA5 is associated with immune-microenvironment dysbalance, and the phenotypic hallmark may be a shift in macrophage polarization toward a pro-inflammatory state. These biomarkers may therefore serve as gauges of overall immune status within tendon tissue, implying a coordination imbalance between tendon cells and immune cells during disease progression(Russo et al., 2022). Notably, recent single-cell RNA-seq studies have revealed an abnormal distribution of macrophage subsets in tendinopathic tissue, with a significant increase in the proportion of M1 macrophages.

Building on these findings, we will leverage single-cell datasets to validate the cell-type-specific expression of the identified genes and clarify whether their signal is chiefly derived from macrophages, thereby sharpening their mechanistic interpretability. In the present study we linked the three-gene signature to the local immune landscape: across both cohorts, low expression of PTCHD1, NEURL1 or ST8SIA5 was consistently accompanied by a reduced M2-like score and an elevated pro-inflammatory cytokine module, indicating a shift toward an M1-dominant,

inflammation-prone microenvironment. This pattern implies that sustained M1 polarization may impede tendon healing.

To dissect the regulatory circuitry governing these genes we generated an atlas of transcription factors and miRNAs predicted to target PTCHD1, NEURL1 and ST8SIA5. Several interferon-regulatory factors and miR-146a-5p emerged recurrently, offering entry points for small-molecule or oligonucleotide modulation. Integrating these regulators into future CRISPR screens should pinpoint the nodes most capable of efficiently reinstating M1/M2 homeostasis. Using an integrated bioinformatics pipeline coupled with three machine-learning frameworks—LASSO, SVM-RFE and Boruta—we identified PTCHD1, NEURL1 and ST8SIA5 as robust candidates. Their specificity for MPRGs was confirmed in a clinically matched trauma-control cohort, and the resulting nomogram showed good calibration and discrimination for all three genes. Together they constitute a quantitative signature that could readily be translated into qPCR or other rapid-assay platforms for evaluation of biopsy or ultrasound-guided needle specimens. Although clinical-phenotype data were not included in the present work, future studies should determine whether expression levels correlate with pain intensity, imaging grade or treatment response in tendinopathy patients. Demonstration of a close relationship with disease severity or prognosis in clinical samples would provide stronger evidence for their use as minimally invasive molecular diagnostics and would accelerate their integration into individualized management algorithms. Drug–gene network analysis nominated several repurposable compounds (e.g., dasatinib, resveratrol, quercetin) predicted to restore M1/M2 homeostasis, offering a rational basis for precision clinical trials. Given their putative role in sculpting the immune microenvironment, PTCHD1, NEURL1 and ST8SIA5 may also serve as candidate targets for immunomodulatory therapy; feasibility and safety in immune-intervention strategies should be systematically evaluated to underpin future precision applications.

Limitations and Future Directions

This study is based on publicly available datasets that contain limited clinical metadata, necessitating validation in prospective clinical cohorts. Furthermore, functional experiments are essential to confirm the causal roles of PTCHD1, NEURL1, and ST8SIA5 in the pathogenesis of tendinopathy.

Conclusion

We identified PTCHD1, NEURL1, and ST8SIA5 as robust biomarkers related to macrophage polarization in tendinopathy. These genes hold promise as minimally invasive diagnostic tools and therapeutic targets, thereby supporting precision medicine approaches for the management of tendinopathy.

Data availability statement

The gene expression datasets analyzed in this study are publicly available in the Gene Expression Omnibus (GEO) database, under accession numbers GSE26051 and GSE180836. The macrophage polarization-related genes (MPRGs) were sourced from the Molecular Signatures Database (MSigDB), accessible at <https://www.gsea-msigdb.org/gsea/msigdb>.

Declarations

Ethical Approval and Informed Consent

This study involving humans were approved by the Sichuan Ba Yi Rehabilitation Center(Sichuan Rehabilitation Hospital),Ethical approval numberNo. CKLL-20231103.This study has been registered with the Chinese Clinical Trials RegistryTrial No. ChiCTR2400083469 (Registration Date: April 26, 2024) .All procedures were conducted in accordance with the 1964 Declaration of Helsinki and its subsequent amendments. Written informed consent was obtained from all participants or their legal guardians.

Author contributions

Ling Ren: Conceptualization, Writing – original draft, Software, Visualization, Validation. Yuxi Qin and Chengqi He: Supervision, Methodology, Project administration, Writing – review and editing.

Funding

This study was sponsored by the National Natural Science Foundation of China (82302884),Sichuan Provincial Administration of Traditional Chinese Medicine Science and Technology Research Special Program(25MSZX531).

Acknowledgments

We thank the provider for the GEO database for datasets GSE26051 and GSE180836, all study participants for validation samples, our laboratory technicians for their support in RT-qPCR, and the open-source bioinformatics community for providing R analysis tools.

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.

References

- Agborbesong, E., Bissler, J., Li, X., 2023. Liquid Biopsy at the Frontier of Kidney Diseases: Application of Exosomes in Diagnostics and Therapeutics. *Genes* 14, 1367. <https://doi.org/10.3390/genes14071367>
- Baldwin, M.J., Mimpfen, J.Y., Cribbs, A.P., Stace, E., Philpott, M., Dakin, S.G., Carr, A.J., Snelling, S.Jb., 2022. Electrospun Scaffold Micro-Architecture Induces an Activated Transcriptional Phenotype within Tendon Fibroblasts. *Front. Bioeng. Biotechnol.* 9, 795748. <https://doi.org/10.3389/fbioe.2021.795748>
- Blanco, S., Muñoz-Gallardo, M.D.M., Hernández, R., Peinado, M.Á., 2025. The Interplay Between Melatonin and Nitric Oxide: Mechanisms and Implications in Stroke Pathophysiology. *Antioxidants* 14, 724. <https://doi.org/10.3390/antiox14060724>
- Bugda Gwilt, K., Olliffe, N., Hoffing, R.A., Miller, G.M., 2019. Trace amine associated receptor 1 (TAAR1) expression and modulation of inflammatory cytokine production in mouse bone marrow-derived macrophages: a novel mechanism for inflammation in ulcerative colitis. *Immunopharmacol. Immunotoxicol.* 41, 577–585. <https://doi.org/10.1080/08923973.2019.1672178>
- Burkett, J.B., Doran, A.C., Gannon, M., 2023. Harnessing prostaglandin E2 signaling to ameliorate autoimmunity. *Trends Immunol.* 44, 162–171. <https://doi.org/10.1016/j.it.2023.01.004>
- Chang, J., Wu, H., Wu, J., Liu, M., Zhang, W., Hu, Y., Zhang, X., Xu, J., Li, L., Yu, P., Zhu, J., 2023. Constructing a novel mitochondrial-related gene signature for evaluating the tumor immune microenvironment and predicting survival in stomach adenocarcinoma. *J. Transl. Med.* 21, 191. <https://doi.org/10.1186/s12967-023-04033-6>
- Chen, H., Xu, Y., Qiu, J., Guan, Z., Yu, S., Ding, Y., 2026. Differential regulatory mechanisms of gut-derived SCFAs (acetate, propionate, butyrate) on the pulmonary TLR2 signaling pathway and their roles in the inflammatory balance of bacterial pneumonia. *Eur. J. Med. Res.* 31, 249. <https://doi.org/10.1186/s40001-025-03666-1>
- Damen, M.S.M.A., Alarcon, P.C., Shah, A.S., Divanovic, S., 2022. Greasing the inflammatory pathogenesis of viral pneumonias in diabetes. *Obes. Rev.* 23, e13415. <https://doi.org/10.1111/obr.13415>
- El Maarouf, A., Moyo-Lee Yaw, D., Lindhout, T., Pearse, D.D., Wakarchuk, W., Rutishauser, U., 2012. Enzymatic Engineering of Polysialic Acid on Cells in Vitro and in Vivo Using a Purified Bacterial Polysialyltransferase. *J. Biol. Chem.* 287, 32770–32779. <https://doi.org/10.1074/jbc.M112.377614>
- Fang, F., Wang, E., Yang, H., Zhao, T., Wang, Q., Zhang, Z., Song, Y., Liu, X., 2025. Reprogramming mitochondrial metabolism and epigenetics of macrophages via miR-10a liposomes for atherosclerosis therapy. *Nat. Commun.* 16, 9117. <https://doi.org/10.1038/s41467-025-64201-8>
- Jeong, S., Yang, K., Lee, Y., Park, J.-W., Park, E.-M., Kang, J.L., 2025. Gas6 induces AIM to suppress acute lung injury in mice by inhibiting NLRP3 inflammasome activation and inducing autophagy. *Front. Immunol.* 16, 1523166. <https://doi.org/10.3389/fimmu.2025.1523166>
- Jia, J., Zhang, Y., Lu, Q., Tian, S., Zhao, Y., Fan, H., 2025. Dual Roles of Hypoxia-Inducible Factor 1 in Acute Lung Injury: Tissue-Specific Mechanisms and Therapeutic Modulation. *Cells* 14, 1089.

<https://doi.org/10.3390/cells14141089>

Kan, T., Ran, Z., Sun, L., Jiang, X., Hou, L., Yang, Y., Jia, Z., Zhang, W., Wang, L., Yan, M., Xie, K., 2023. Cell-free fat extract-loaded microneedles attenuate inflammation-induced apoptosis and mitochondrial damage in tendinopathy. *Mater. Today Bio* 22, 100738. <https://doi.org/10.1016/j.mtbio.2023.100738>

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M.; KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457-D462 (2016).

Kanehisa, M. and Goto, S.,: KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27-30 (2000).

Koenis, D.S., Medzikovic, L., Van Loenen, P.B., Van Weeghel, M., Huveneers, S., Vos, M., Evers-van Gogh, I.J., Van Den Bossche, J., Speijer, D., Kim, Y., Wessels, L., Zelcer, N., Zwart, W., Kalkhoven, E., De Vries, C.J., 2018. Nuclear Receptor Nur77 Limits the Macrophage Inflammatory Response through Transcriptional Reprogramming of Mitochondrial Metabolism. *Cell Rep.* 24, 2127-2140.e7. <https://doi.org/10.1016/j.celrep.2018.07.065>

Koo, S., Garg, N.J., 2019. Metabolic programming of macrophage functions and pathogens control. *Redox Biol.* 24, 101198. <https://doi.org/10.1016/j.redox.2019.101198>

Li, S., Liu, H., Ruan, Z., Guo, R., Sun, C., Tang, Y., Huang, X., Gao, T., Hao, S., Li, H., Song, N., Su, Y., Ning, F., Li, Z., Chang, T., 2023. Landscape analysis of m6A modification regulators related biological functions and immune characteristics in myasthenia gravis. *J. Transl. Med.* 21, 166. <https://doi.org/10.1186/s12967-023-03947-5>

Liu, S., Wang, C., Yang, H., Zhu, T., Jiang, H., Chen, J., 2020. Weighted gene co-expression network analysis identifies FCER1G as a key gene associated with diabetic kidney disease. *Ann. Transl. Med.* 8, 1427–1427. <https://doi.org/10.21037/atm-20-1087>

Liu, X., Jin, L., Yang, Y., Lu, M., Xue, B., 2025. SNRPA upregulation promotes mitochondrial function and drives CRPC aggressiveness. *Cell Death Dis.* 17, 74. <https://doi.org/10.1038/s41419-025-08302-8>

Luque-Campos, N., Bustamante-Barrientos, F.A., Pradenas, C., García, C., Araya, M.J., Bohaud, C., Contreras-López, R., Elizondo-Vega, R., Djouad, F., Luz-Crawford, P., Vega-Letter, A.M., 2021. The Macrophage Response Is Driven by Mesenchymal Stem Cell-Mediated Metabolic Reprogramming. *Front. Immunol.* 12, 624746. <https://doi.org/10.3389/fimmu.2021.624746>

M. de-Brito, N., Duncan-Moretti, J., C. da-Costa, H., Saldanha-Gama, R., Paula-Neto, H.A., G. Dorighello, G., L. Simões, R., Barja-Fidalgo, C., 2020. Aerobic glycolysis is a metabolic requirement to maintain the M2-like polarization of tumor-associated macrophages. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1867, 118604. <https://doi.org/10.1016/j.bbamcr.2019.118604>

Millar, N.L., Silbernagel, K.G., Thorborg, K., Kirwan, P.D., Galatz, L.M., Abrams, G.D., Murrell, G.A.C., McInnes, I.B., Rodeo, S.A., 2021. Tendinopathy. *Nat. Rev. Dis. Primer* 7, 1. <https://doi.org/10.1038/s41572-020-00234-1>

Paneru, B.D., Hill, D.A., 2023. The role of extracellular vesicle-derived miRNAs in adipose tissue function and metabolic health. *Immunometabolism* 5, e00027. <https://doi.org/10.1097/IN9.0000000000000027>

Peng, J., Zhang, H., Wang, H., Meng, Q., Li, D., Gao, M., Li, Y., Ma, X., Xia, L., Xu, R., Zhu, J., 2025. Mechanistic insights into Guizhi Fuzi decoction for lumbar disc herniation: Integrating network pharmacology and bioinformatics approach. *Medicine (Baltimore)* 104, e41917. <https://doi.org/10.1097/MD.00000000000041917>

Qiao, C., Sun, Z., Muhammad, F., Gong, W., Lin, J., Lv, J., Cheng, X., Fei, Y., Xu, N., Xie, Y., Jiang, R., Lin, C., Yang, H., Lv, Z., Wang, Z., Liu, Y., Lu, Z., Gu, X., Zhang, Z., Hu, C., Zhang, H., Meng, J., Yang, X., Shi, D., Wei, H., Bao, N., 2025. Ruthenium Oxide Nanozyme for Tendinopathy Treatment via Countering Ferroptosis. *Adv. Healthc. Mater.* 14, 2501035. <https://doi.org/10.1002/adhm.202501035>

Quintero, D., Perucca Orfei, C., Kaplan, L.D., De Girolamo, L., Best, T.M., Kouroupis, D., 2023. The roles and therapeutic potential of mesenchymal stem/stromal cells and their extracellular vesicles in tendinopathies. *Front. Bioeng. Biotechnol.* 11, 1040762. <https://doi.org/10.3389/fbioe.2023.1040762>

Russo, V., El Khatib, M., Prencipe, G., Citeroni, M.R., Faydaver, M., Mauro, A., Berardinelli, P., Cerveró-Varona, A., Haidar-Montes, A.A., Turriani, M., Di Giacinto, O., Raspa, M., Scavizzi, F., Bonaventura, F., Stöckl, J., Barboni, B., 2022. Tendon Immune Regeneration: Insights on the Synergetic Role of Stem and Immune Cells during Tendon Regeneration. *Cells* 11, 434. <https://doi.org/10.3390/cells11030434>

Schulze-Tanzil, G.G., Delgado Cáceres, M., Stange, R., Wildemann, B., Docheva, D., 2022. Tendon healing: a concise review on cellular and molecular mechanisms with a particular focus on the Achilles tendon. *Bone Jt. Res.* 11, 561–574. <https://doi.org/10.1302/2046-3758.118.BJR-2021-0576.R1>

Shan, Z., Li, J., Mu, X., Zhang, J., Ren, S., 2026. Identification and validation of PANoptosis-related genes in ankylosing spondylitis. *BMC Med. Genomics* 19, 28. <https://doi.org/10.1186/s12920-025-02301-6>

Shao, Q., Li, N., Qiu, H., Zhao, M., Jiang, C., Wan, C., 2025. Identification Exploring the Mechanism and Clinical Validation of Mitochondrial Dynamics-Related Genes in Membranous Nephropathy Based on Mendelian Randomization Study and Bioinformatics Analysis. *Biomedicine* 13, 1489. <https://doi.org/10.3390/biomedicine13061489>

Suresh, A., Martens, P., Tang, W.H.W., 2022. Biomarkers for Myocarditis and Inflammatory Cardiomyopathy. *Curr. Heart Fail. Rep.* 19, 346–355. <https://doi.org/10.1007/s11897-022-00569-8>

Tang, A., Shu, Q., Jia, S., Lai, Z., Tian, J., 2024. Adipose Mesenchymal Stem Cell-Derived Exosomes as Nanocarriers for Treating Musculoskeletal Disorders. *Int. J. Nanomedicine* Volume 19, 13547–13562. <https://doi.org/10.2147/IJN.S486622>

Tang, M.S., Miraldi, E.R., Girgis, N.M., Bonneau, R.A., Loke, P., 2020. Alternative Activation of Macrophages Is Accompanied by Chromatin Remodeling Associated with Lineage-Dependent DNA Shape Features Flanking PU.1 Motifs. *J. Immunol.* 205, 1070–1083. <https://doi.org/10.4049/jimmunol.2000258>

Vebr, M., Pomahačová, R., Sýkora, J., Schwarz, J., 2023. A Narrative Review of Cytokine Networks: Pathophysiological and Therapeutic Implications for Inflammatory Bowel Disease Pathogenesis. *Biomedicine* 11, 3229. <https://doi.org/10.3390/biomedicine11123229>

Wang, X., Li, M., Liu, X., Sun, G., Zhang, D., Sun, L., Yin, Y., Zhang, W., Hao, J., 2025. Isolation of Murine Pancreatic Stellate Cells and the Establishment of a New ex-vivo Activation Model. *Clin. Exp. Gastroenterol.* Volume 18, 79–89. <https://doi.org/10.2147/CEG.S507384>

Wang, Z., Mao, J., Zhang, Y., Yang, W., Sun, D., Lu, Z., Lu, X., Gao, B., 2025. Pan-cancer analysis reveals the potential role of DHCR24 in bladder cancer via interactions with HRAS to facilitate cholesterol synthesis. *Oncol. Lett.* 30, 1–18. <https://doi.org/10.3892/ol.2025.15131>

Wearing, S.C., Hooper, S.L., Langton, C.M., Keiner, M., Horstmann, T., Crevier-Denoix, N., Pourcelot, P., 2024. The Biomechanics of Musculoskeletal Tissues during Activities of Daily Living: Dynamic Assessment Using Quantitative Transmission-Mode Ultrasound Techniques. *Healthcare* 12, 1254. <https://doi.org/10.3390/healthcare12131254>

Xiao, S., Li, R., Diao, H., Zhao, F., Ye, X., 2013. Progesterone Receptor-Mediated Regulation of N-Acetylneuraminase Pyruvate Lyase (NPL) in Mouse Uterine Luminal Epithelium and Nonessential Role of NPL in Uterine Function. *PLoS ONE* 8, e65607. <https://doi.org/10.1371/journal.pone.0065607>

Yuan, T., Lai, C.-T., Yang, S.-Q., Meng, J., Qian, H., Yu, X., Jiang, H., Cao, Q.-G., Xu, J.-D., Bao, N.-R., 2024. The rat as a novel model for chronic rotator cuff injuries. *Sci. Rep.* 14, 5344. <https://doi.org/10.1038/s41598-024-55281-5>

Zhang, H., Huang, X., Chen, R., Lin, W., Li, L., Peng, X., Yang, X., Chen, X., Guo, T., Weng, J., Li, J., Yuan, Q., Tan, H., Wang, M., 2025. Interfacial Coordination-Engineered Emulgel Modulates the Osteo-Immune Microenvironment in Periodontitis Treatment Via HIF1 Signaling Pathway-Induced Metabolic Reprogramming. *Adv. Sci.* 12, e09998. <https://doi.org/10.1002/adv.202509998>

Zhang, N., Zhao, F., Chen, H., Wang, J., Li, H., 2025. UBD-mediated glycolytic reprogramming promotes M2 macrophage polarization in ovarian cancer immune evasion. *J. Cell Commun. Signal.* 19, e70034. <https://doi.org/10.1002/ccs3.70034>

Zhang, Y., Wang, Y., Gu, Y., Liu, Y., Liu, G., Wu, J., Bai, N., 2025. Regulation of FOXL2 gene in ovarian granulosa cell tumor by JNK inhibitor. *Sci. Rep.* 15, 12167. <https://doi.org/10.1038/s41598-025-85439-8>

Zheng, Q., Hou, W., 2021. Regulation of angiogenesis by microRNAs in cancer. *Mol. Med. Rep.* 24, 583. <https://doi.org/10.3892/mmr.2021.12222>

Zhong, F.-M., Yao, F.-Y., Yang, Y.-L., Liu, J., Li, M.-Y., Jiang, J.-Y., Zhang, N., Xu, Y.-M., Li, S.-Q., Cheng, Y., Xu, S., Huang, B., Wang, X.-Z., 2023. Molecular subtypes predict therapeutic responses and identifying and validating diagnostic signatures based on machine learning in chronic myeloid leukemia. *Cancer Cell Int.* 23, 61. <https://doi.org/10.1186/s12935-023-02905-x>

Zhou, J., Wang, H., Wang, J., Zhou, F., 2025. Discovering biomarkers for chronic sinusitis with nasal polyps: a study integrating bioinformatics analysis and experimental validation of macrophage polarization and metabolism-related genes. *Front. Bioinforma.* 5, 1613136. <https://doi.org/10.3389/fbinf.2025.1613136>

Zhou, Y., Teng, W., Wu, J., Luo, Y., Wang, Y., Li, Y., 2025. Piperlongumine Inhibits Lung Cancer Growth by Inducing Endoplasmic Reticulum Stress Leading to Suppression of M2 Macrophage Polarization. *Biol. Proced. Online* 27, 18. <https://doi.org/10.1186/s12575-025-00279-0>

Zhu, J., Li, Yu, Zhong, C., Zhu, M., Zheng, Y., Xiong, A., Meng, P., Shan, L., Li, Yang, Huang, J., 2023a. Neuritin affects the activity of neuralized-like 1 by promoting degradation and weakening its affinity for substrate. *Acta Biochim. Biophys. Sin.* 55, 1650–1658. <https://doi.org/10.3724/abbs.2023098>

Zhu, J., Li, Yu, Zhong, C., Zhu, M., Zheng, Y., Xiong, A., Meng, P., Shan, L., Li, Yang, Huang, J., 2023b. Neuritin affects the activity of neuralized-like 1 by promoting degradation and weakening its affinity for substrate. *Acta Biochim. Biophys. Sin.* 55, 1650–1658. <https://doi.org/10.3724/abbs.2023098>

Zhu, J., Li, Yu, Zhong, C., Zhu, M., Zheng, Y., Xiong, A., Meng, P., Shan, L., Li, Yang, Huang, J., 2023c. Neuritin affects the activity of neuralized-like 1 by promoting degradation and weakening its

affinity for substrate. *Acta Biochim. Biophys. Sin.* 55, 1650–1658.
<https://doi.org/10.3724/abbs.2023098>

Zhu, L., Wang, Y., Jin, S., Niu, Y., Yu, M., Li, Z., Chen, L., Wu, X., Ding, C., Wu, T., Shi, X., Zhang, Y., Luo, D., Liu, Y., 2023. Parishin A-loaded mesoporous silica nanoparticles modulate macrophage polarization to attenuate tendinopathy. *Npj Regen. Med.* 8, 14.
<https://doi.org/10.1038/s41536-023-00289-0>