Integrative Analysis of Single-cell and Bulk RNA Sequencing Reveals Macrophage Heterogeneity in Lyme Arthritis

Yan Dong[1,2], Yanshuang Luo[1], Meng Liu[1], Yantong Chen[1], Chao Song[1], Xiaorong Liu[1], Fusong Yang[1], Zhi Liang[1], Liye Zi[1], Hongbing Gao[1], Zhichao Cao[1], Fucun Gao[1], Xuesong Chen[1*], Guozhong Zhou[1,2*]

Abstract

Introduction Lyme arthritis (LA) is a chronic inflammatory joint disease caused by *Borrelia burgdorferi* infection, with approximately 10-20% of patients developing antibiotic resistance. Macrophages in the synovial microenvironment play a crucial role in disease progression, but their precise regulatory mechanisms remain unclear.

Methods scRNA-seq data of synovial tissue from LA patients was obtained from the GEO database (GSE233850) and integrated with transcriptome data from our research team's LA mouse model and the GSE125503 dataset. Multiple analytical strategies, including cell clustering, differential gene analysis, GO and KEGG enrichment analysis, pseudotime trajectory analysis, and cell-cell communication analysis, were employed to systematically investigate macrophage characteristics in LA. In vitro experiments using RAW cell lines were conducted to validate the role of key genes in response to *Borrelia burgdorferi* infection.

Results Through single-cell transcriptome analysis of 78,015 cells in the GSE233850 dataset, we identified and annotated 27,843 macrophages, further classifying them into 4 functional subgroups. Differential expression analysis revealed 387 macrophage-specific genes. Cell-cell communication network analysis showed that macrophages interact with other immune cells, including T cells and B cells, through multiple signaling pathways. GO and KEGG enrichment analyses indicated that these differentially expressed genes were mainly enriched in biological processes such as inflammatory response regulation, pathogen recognition, and cytokine signaling. By integrating transcriptome data from our team's LA mouse model and the GSE125503 dataset, we screened and validated three potential diagnostic marker genes (SIRPB1C,

FABP5, and PEPD). In vitro experiments using the RAW264.7 macrophage cell line showed significant expression changes in these genes following *Borrelia burgdorferi* stimulation ($P < 0.05$).

Conclusion This study provides the first systematic characterization of macrophage transcriptome in the LA synovial microenvironment, revealing their functional heterogeneity and identifying potential diagnostic markers, offering new targets for targeted therapeutic strategies.

Key words: *Borrelia burgdorferi*, Lyme arthritis, Single-cell RNA sequencing, Macrophage heterogeneity, Synovial microenvironment.

1. Introduction

Lyme arthritis (LA) is a manifestation of Lyme disease caused by infection with *Borrelia burgdorferi*, a tick-borne spirochete. It is the most common and persistent late-stage symptom of Lyme disease, typically appearing after several months [1, 2]. The main clinical manifestations of Lyme arthritis include intermittent or persistent swelling and pain in large joints (typically the human knee joint), which significantly affects patients' quality of life. Statistics show that approximately 60 % of individuals with Lyme disease will develop arthritis if they do not receive antibiotic treatment [3, 4]. The complexity of Lyme arthritis stems from the interaction between pathogens and the host immune system, leading to inflammation and, in some cases (particularly in antibiotic-resistant cases), chronic arthritis [5, 6]. Various studies have explored different aspects of this disease, including immune responses, bacterial components, and genetic susceptibility, yet challenges remain in fully elucidating the mechanisms involved.

Extensive research has shown that macrophages play a central role in the pathogenesis of Lyme arthritis [7, 8]. Following *Borrelia burgdorferi* infection, macrophages secrete chemokines such as CCL4 and CCL2, which attract monocytes and T cells to the infection site, while interacting with other immune cells through phagocytosis and

antigen presentation processes to amplify the inflammatory response [9, 10]. Both animal models and human studies have confirmed that macrophages not only have direct arthritis-inducing effects, but their secreted pro-inflammatory cytokines, including IL-1β, TNF-α, and IFN-γ, also play crucial roles in maintaining and exacerbating joint inflammation [11-13]. Although macrophages are central to the inflammatory response in Lyme arthritis, their role is not purely destructive. They also participate in the clearance of Borrelia bacteria and the resolution of inflammation [14]. In summary, macrophage infiltration and activity are subject to complex regulation by cellular polarization states, cytokine environment, and immune cell networks. Therefore, developing a thorough understanding of macrophages' role in the pathogenesis of Lyme arthritis is crucial.

With the rapid development of single-cell RNA sequencing (scRNA-seq) technology, researchers have been able to analyze tissue heterogeneity at the single-cell level [15-17], facilitating precise studies on the roles of individual cells such as macrophages in the progression of Lyme arthritis. However, studies integrating scRNA-seq with bulk RNA sequencing data to explore the role of macrophages in Lyme arthritis remain limited. To this end, this study established a Borrelia burgdorferi-infected C57BL/6 mouse model and conducted transcriptome sequencing of ankle joint tissues from infected mice to systematically analyze gene expression changes during the arthritic process. Furthermore, we adopted a multi-omics integrative analysis strategy, associating the transcriptome data of mouse joint tissues with bone marrow-derived macrophage transcriptome and scRNA-seq data from public databases, constructing a comprehensive gene expression network to elucidate the host immune response mechanisms to Bb infection. This multidimensional research approach not only deepens our understanding of the central role of macrophages in the pathogenesis of Lyme arthritis but also provides an important theoretical basis for the development of targeted therapeutic strategies.

2. Methods

2.1 Single-cell RNA Sequencing Data Quality Control and Processing

The scRNA-seq data of Lyme Arthritis was accessed from the GSE23385 dataset,

which is established in the GEO database corresponding to a published article "Single-cell RNA sequencing of murine ankle joints over time reveals distinct transcriptional changes following *Borrelia burgdorferi* infection" [18]. The single-cell RNA sequencing data processing workflow is as follows: Raw data was obtained from ankle joint samples of 2 normal and 4 *Borrelia burgdorferi*-infected C57 mice from the GSE233850 dataset. Quality control criteria include: more than 200 expressed genes per cell, each gene expressed in at least 3 cells, and mitochondrial RNA content less than 20% [19]. Data processing using the Seurat package: Normalization was performed using the NormalizeData function (natural logarithm of expression values multiplied by total gene expression ×10000), followed by identification of 2000 highly variable genes using FindVariableFeatures, data centering using ScaleData, and finally data integration using FindIntegrationAnchors (RunHarmony) with 2000 anchors to eliminate batch effects. A total of 78,015 cells were obtained for subsequent analysis [20].

2.2 Dimensionality Reduction and Data Clustering

Given that each gene in the samples exists as a dimension, high dimensional data becomes challenging to visualize. Therefore, we need to employ dimensionality reduction techniques to represent the true structure of the data using a reduced number of dimensions [21, 22]. For high-dimensional data visualization, the following dimensionality reduction and clustering steps were implemented: First, the RunPCA function was used to perform principal component analysis on highly variable genes, followed by data integration using RunHarmony with 2000 anchor points. Then, FindNeighbors and FindClusters functions were applied for clustering analysis. Finally, the RunUMAP function (dims=1:17, resolution=0.5) was used for dimensionality reduction visualization, effectively displaying the cell clustering results.

2.3 Cell Annotation, DEG and Marker Gene Identification

To characterize each cluster, a comprehensive manual annotation was performed. The FindAllMarkers function in the Seurat package was used to identify differentially expressed genes (DEGs), with selection criteria of |log2FoldChange|> 0.25 and

adjusted P-value< 0 .05. The FindAllMarkers function was further utilized to identify cell type-specific marker genes.

2.4 GO Enrichment Analysis and KEGG Pathway Enrichment Analysis

The clusterProfiler package (v3.14.3) was used to perform GO functional enrichment and KEGG pathway enrichment analysis on differentially expressed genes, and the analysis results were visualized using the ggplot2 package [23-25].

2.5 Cellular Trajectory Analysis (Pseudotime Analysis)

Cells undergo continuous dynamic changes, transforming from one cell type to another, which leads to changes in gene expression and functional states [26]. Using the monocle package to perform pseudotime analysis on macrophage subtypes, cell differentiation trajectories can be constructed through the plot_cell_trajectory function to demonstrate the dynamic changes of cells under different states. The BEAM function can optionally be used to identify key branch-determining genes, and results can be visualized using the plot_genes_branched_heatmap function.

2.6 Cell-Cell Communication Analysis

Using the CellChat package (v1.0.0) to analyze intercellular communication networks, cell communication relationships are constructed based on expression data of known ligand-receptor pairs to infer interaction patterns and their abundance between different cell types [27].

2.7 C57BL/6 Mouse Lyme Arthritis Modeling and Transcriptome Sequencing

Mice

Female C57BL/6J mice (22-24g) aged 4-6 weeks were obtained from Beijing Vital River Laboratory Animal Technology Co., Ltd. The mice were housed separately in SPF-grade facilities at the School of Basic Medicine, Kunming Medical University, under conditions of 12-hour light/dark cycles at 21.5°C, with free access to food and water and daily bedding changes. The experiment was approved by the Medical Ethics Committee of Anning First People's Hospital (Approval No.2024-046-01).

Bacteria propagation and infection

The *Borrelia burgdorferi* strain B31, purchased from ATCC (ATCC - 35210), was cultured in complete BSK-II medium at 37°C with 5 % CO2 and 100 % humidity. For

the experimental group, mice received bilateral footpad injections of spirochete suspension (1×10^6/50μl per side), while the control group received an equal volume of PBS. Ankle joint swelling was measured twice weekly using calipers, and X-ray examinations were performed on randomly selected mice.

Sample Collection and Transcriptome Analysis

On the 16th day after surgery, 6 mice from each group were sampled. Half of the samples were used for histopathological examination, and 4μm thick paraffin sections were prepared for HE staining. The other half of the samples (approximately 80mg of ankle joint tissue) were sent to LC Bio Technology CO., Ltd. (Hangzhou, China) for transcriptome sequencing. The sequencing was performed using the Illumina Novaseq 6000 platform, following standard operating procedures for paired-end sequencing with a sequencing mode of PE150. After obtaining the standardized clean data, bioinformatics analysis was conducted.

2.8 Transcriptome Data Integration

Based on our research focus on the role of macrophages in LA synovial tissue, we obtained transcriptome sequencing dataset (GSE125503, platform GPL17021) of *Borrelia burgdorferi*-infected macrophages from the Gene Expression Omnibus (GEO) database. This dataset includes 14 samples derived from bone marrow-derived macrophages (BMDMs) of C57BL/6 mice, comprising 4 control samples and 10 B. burgdorferi-infected samples. Subsequently, we integrated our team's Bulk-transcriptome data with the GSE125503 Bulk-transcriptome data. We merged different datasets by columns using the cbind function and obtained common genes through Reduce(intersect, geneList). Batch effects were then corrected using the ComBat function. Following this, we established criteria for identifying differentially expressed genes (DEGs) using the R package DESeq2. Specifically, we applied a threshold of |log2 FoldChange|>1 and adjusted P-value <0.05 to select DEGs for subsequent analysis [28, 29].

2.9 Least Absolute Shrinkage and Selection Operator (LASSO) Regression Analysis

We performed intersection analysis between differentially expressed genes (DEGs)

from macrophage-related single-cell sequencing and DEGs from the merged transcriptome dataset. Using the common differential genes obtained, we constructed and validated a model through the Least Absolute Shrinkage and Selection Operator (LASSO) method [30].

2.10 Quantitative Reverse Transcription PCR (qRT-PCR)

Cell Culture and Treatment: RAW264.7 cells were purchased from Shanghai Zhongqiao Xinzhou Biotechnology Co., Ltd. The cells were infected with *Borrelia burgdorferi* (MOI=1) for 6 hours, or treated with vehicle control (PBS). The samples were collected and stored at -80°C until needed. RNA Extraction and qRT-PCR Analysis: Total RNA was extracted using MolPure® Cell/Tissue Total RNA Kit (#19221ES50, ShangHai Yeasen Biotechnology) and reverse transcribed into complementary DNA using Hifair®III 1st Strand cDNA Synthesis SuperMix for qPCR (#11141ES60, ShangHai Yeasen Biotechnology). Quantitative reverse-transcription PCR was performed on the Gentier 96R Real-Time PCR System (Tian Long) using Hieff unicon universal blue Qpcr SYBR Master mix (#11184ES08, ShangHai Yeasen Biotechnology). The reaction conditions were as follows: 95°C for 2 minutes, followed by 40 cycles of 95°C for 1 second and 60°C for 30 seconds. Expression of target genes was normalized to Gapdh using the $2^{(-\Delta\Delta CT)}$ method. The primers were synthesized by Tsingke Biotechnology (Beijing). The primer sequences (5’-3’) used to amplify Sirpb1c were forward, TCAGCAGACAGTGGCCTTTA and reverse, CTGAGAGCGGACATCCCTAG. The primers used to amplify FABP5 were forward, CGGGAAGGAGAGCACGATAA and reverse, TTGTTGCATTTGACCGCTCA. The primers used to amplify Pepd were forward, AATCAACCCTGTTTGTGCCC and reverse, CACTGGCAATCTCATCCGTG. The primers used to amplify GAPDH were forward, TGT TGC CAT CAA TGA CCC CT and reverse TCG CCC CAC TTG ATT TTG GA.

2.11 Construction and Testing of Lyme Arthritis Diagnostic Nomogram Model

We employed the "pROC" R package for ROC curve analysis and the "rms" R package to create a nomogram for Lyme arthritis diagnosis [31, 32]. Risk scores were calculated based on the expression levels of each core gene, with the total risk

score defined as the sum of all individual gene risk scores. The diagnostic value of this nomogram for Lyme arthritis was evaluated through decision trees, calibration curves, and ROC curves.

2.12 Statistical Analysis

All analyses and visualizations were performed using R (version 4.2.1) and GraphPad Prism (version 9.0.2). Data are presented as mean ± SEM. Statistical comparisons between two groups were conducted using two-tailed Student's t-test or Mann-Whitney U test. Statistical significance was defined as $P < 0.05$ (*), $P < 0.01$ (**), and ns for not significant.

3. Results

3.1 Single-cell Transcriptome of Lyme Arthritis

In this study, we analyzed a total of 6 samples from the GSE233850 dataset, including ankle joint samples from 2 normal C57 mice uninfected with *Borrelia burgdorferi* and 4 paired C57 mice infected with Borrelia burgdorferi. Each sample underwent initial screening to remove low-quality data, retaining 78,015 cells for subsequent analysis. After normalization and PCA, the first 17 dimensions of cells were retained, followed by clustering visualization using the UMAP method. Cluster analysis divided the 78,015 cells into 24 clusters, as shown in Figure 2A. Cell annotation using the SingleR package revealed 8 cell types (Figure 2B): macrophages (27,843 cells, 35.7%) with high expression of Cd14 and CD68; T/NK cells with high expression of Nkg7, CD3e, and CD3g; B cells with high expression of Cd79a and Cd79b; DC cells with high expression of H2-Eb1, CD80, and CD86; endothelial cells with high expression of Pecam1, cldn5, and cdh5; fibroblasts with high expression of Col3a1 and Col1a1; synovial cells with high expression of Prg4; and muscle cells with high expression of Acta1 (Figures 2E-2G). Figure 2C shows the single-cell transcriptome atlas of normal and Lyme Arthritis samples. Figure 2D displays the single-cell transcriptome atlas of different samples. Figure 2H represents their proportions in different samples, with the Lyme Arthritis model showing higher proportions of macrophages, B cells, T/NK cells, and dendritic cells (DC), while fibroblasts, muscle cells, and synovial cells showed relatively lower proportions.

3.2 Single-cell Transcriptome Landscape

Given the crucial immunoregulatory role of macrophages in Lyme arthritis, this study conducted subsequent clustering and annotation analysis on 27,843 macrophages extracted from all samples. As shown in Figure 3, scRNA-seq analysis clustered macrophages into 9 distinct clusters (Figure 3A) and ultimately identified 4 macrophage subtypes (Figure 3B), comprising 8,114 Mac1 macrophages, 8,199 Mac2 macrophages, 6,533 Mac3 macrophages, and 4,997 Mac4 macrophages. The characteristics of these subtypes were demonstrated through multiple dimensions: comparison between normal and disease model groups (Figure 3C), distribution across different samples (Figure 3D), marker gene expression features of each subtype, including Pf4, Aqp1, Tppp3, Fxyd2, Mrc1, Hdc, G0s2, Retnlg, S100a8, S100a9, Thbs1, Ifitm6, Ifitm3, Plac8, Ly6c2, Timp2, Sparc, Vsig4, Cbr2 (Figures 3E-3G), and the compositional proportions of each subtype in normal and disease model groups displayed through a river plot (Figure 3H).

3.3 Macrophage-Related Enrichment Analysis

Figure 4A showed a volcano plot of differential analysis between overall cell subgroups. To elucidate the biological functions of marker genes in macrophages, GO and KEGG enrichment analyses were performed. KEGG pathway enrichment analysis of differentially expressed genes among macrophage populations (Figure 4B) revealed that signature genes were predominantly enriched in key pathways orchestrating macrophage activation, including Lysosome pathway, Chemokine signaling pathway, Pathogen infection response-related pathways, Endocytosis pathway, Apoptosis pathway, Osteoclast differentiation pathway, Lipid and atherosclerosis pathway, Epstein-Barr virus infection pathway, Efferocytosis pathway, and Fc gamma R-mediated phagocytosis pathway. GO functional enrichment analysis of differentially expressed genes between macrophage cells (Figure 4C-4D) revealed that: Biological Processes (BP) were mainly associated with immune response-regulating signaling pathway, immune response-activating signaling pathway, leukocyte migration, regulation of innate immune response, phagocytosis, and cytokine-mediated signaling pathway. Cellular Components (CC) were primarily

related to vacuolar membrane, endocytic vesicle, membrane raft, membrane microdomain, endosome membrane, and lysosomal membrane. Molecular Functions (MF) were associated with nucleoside-triphosphatase regulator activity, GTPase regulator activity, GTPase binding, GTPase activator activity, phospholipid binding, and small GTPase binding. Figure 4E showed a scatter plot of differential analysis between macrophage groups. Figure 4F showed a volcano plot of differential analysis between macrophage groups. The bar plots depicting the GO enrichment results for both upregulated and downregulated differentially expressed genes among macrophage subpopulations (Figure 4G) revealed distinct patterns: upregulated genes exhibited significant enrichment in pathways governing leukocyte intercellular adhesion, modulation of leukocyte adhesive interactions, T cell activation regulation, enhancement of leukocyte activation cascades, and positive regulation of cellular adhesion mechanisms. Conversely, downregulated genes demonstrated predominant enrichment in pathways controlling actin cytoskeletal architecture, peptidase activity modulation, actin filament dynamics, leukocyte migratory processes, and regulation of actin filament-dependent cellular mechanisms.

3.4 Pseudotime Analysis

To determine the relationships between macrophage subtypes and states, we performed differentiation trajectory analysis using the Monocle R package. In this study, Figure 5A shows four differentiation states of macrophages during development, while Figure 5B depicts the temporal sequence of cell subtype differentiation, with darker-colored cells gradually transitioning to lighter-colored cells. Figure 5C displays three branches of macrophage subtypes. According to pseudotime ordering: MAC1 macrophages progress from stage 1 to stage 3, and finally to stage 5, MAC2 macrophages mainly remain in stage 1, Similarly, MAC3 macrophages progress from stage 1 to stage 2, and finally to stage 3. MAC4 macrophages progress from stage 1 to stage 2, stage 3, and finally to stage 5. During the differentiation process from early to late stages, MAC1, MAC2, and MAC4 macrophages gradually accumulate in the late stage (stage 5), while MAC2 mainly accumulates in the early stage (Figure 5A-C).

Through GO biological process (BPs) enrichment analysis, we found that differentially expressed genes were mainly enriched in four functional modules. The first module relates to immune cell recruitment and inflammation regulation, including leukocyte and monocyte migration and recruitment, circulating immunoglobulin-mediated humoral immune response, classical complement activation pathway, and fine regulation of inflammatory responses, reflecting the dynamic recruitment of inflammatory cells and regulation of immune responses. The second module involves inflammatory effects and tissue remodeling, mainly including the production and regulation of pro-inflammatory factors (such as TNF), regulation of angiogenesis and vascular development, collagen metabolism and tissue remodeling, and changes in cellular energy metabolism, revealing tissue pathological changes and metabolic adaptation during disease progression. The third module is associated with antigen presentation and immune activation, including MHC class II molecule-mediated antigen processing and presentation, recognition and presentation of self-antigens, and $CD4^+$ T cell activation, demonstrating the initiation and maintenance mechanisms of autoimmune responses. The fourth module mainly relates to protein homeostasis regulation, covering ubiquitin-proteasome system regulation, regulation of protein post-translational modifications, and p53 signaling pathway-mediated cell fate regulation, reflecting precise regulation of cellular protein homeostasis. These four functional modules are interconnected, forming a complex regulatory network involved in disease development (Figure 5D). During the development of Lyme arthritis-related macrophages, the pseudotime dynamics of the top eight genes (C1qa, C1qb, C1qc, Ecm1, Emp1, Lmna, Pitp, Tppp3) in the four macrophage subgroups show (from left to right) the developmental process. Each point represents a cell, different colors represent different clusters, and the y-axis represents the expression level of each gene (Figure 5E).

3.5 Cell-to-Cell Communication Between Macrophages and Other Cells in Lyme Arthritis Tissue

To further investigate the influence of macrophages on other cells during Lyme arthritis pathogenesis, we conducted cell-to-cell communication analysis by tracking

ligand-receptor interactions using our scRNA-seq data. We found that ligand-receptor interactions between macrophages and other major cells were stronger in Lyme arthritis tissue compared to normal tissue (Figure 4A-4B). Total pathway analysis of both groups showed enhanced binding of LAMININ, GALECTIN, and THBS with their receptors in the Lyme arthritis group compared to normal tissue (Figure 4C). Subsequently, analysis of total receptor-ligand interactions and single-purpose pathway differences between the two groups revealed frequent occurrence of GALECTIN signaling pathways in Lyme arthritis joint tissue. This was evidenced by the expansion from normal tissue, where only macrophages and DC cells interacted, to Lyme arthritis tissue where Galectin-9 signaling and interactions occurred among six cell types: macrophages, B cells, DC cells, endothelial cells, fibroblasts, and synoviocytes. This suggests activation of GALECTIN signaling in these immune cells under disease conditions (Figure 4D-4E). Further analysis revealed that the LGALS9 (GALECTIN-9)/CD45 axis contributed most significantly to GALECTIN signaling (Figure 4F). The expression of Galectin signaling pathway-related genes between seven major cell types in both Lyme arthritis and normal tissues is shown in Figure 4G. Specifically, Galectin signaling pathways were widely present in ligand-receptor interactions across various cell types, with Ptprc, lghm, and CD44 showing high expression levels in multiple cell types. Finally, the crosstalk of ligand-receptor pairs LGALS9/CD45, LGALS9/CD44, and LGALS9/lghm between seven major cell types in both Lyme arthritis and normal tissues is shown in Figures 4H, 4I, and 4J, respectively.

3.6 Bulk-transcriptome sequencing of C57BL/6 Mouse Ankle Joints

In the Lyme arthritis mouse model, C3H mice began showing ankle joint swelling in their hind legs on day 2 after Bb injection, reaching peak swelling at around 2 weeks, with swelling gradually subsiding after 4-5 weeks (line graph). On day 16 post-injection, compared to the PBS control group, the experimental group mice exhibited significant joint swelling and mild stiffness, but no local skin ulceration or blackening was observed (mobile phone photo). X-ray examination revealed soft tissue edema (X-ray), and H&E staining further confirmed inflammatory cell

infiltration and significant edema in joint synovial and other soft tissues, though no obvious bone or cartilage destruction was observed (H&E staining) (see Supplementary Figure 1-3).

Based on this study's focus on macrophage specificity in Lyme arthritis, we merged transcriptome sequencing data from C57BL/6 mouse ankle joints with transcriptome data from C57BL/6 mouse bone marrow-derived macrophages (GSE125503, platform GPL17021). Subsequently, differential gene analysis was performed on the merged transcriptome dataset. Figures 7A-7B show the heatmap and volcano plot of differential genes. The 387 DEGs identified were intersected with 159 macrophage-related DEGs from scRNA-seq, yielding 10 overlapping genes (Figure 7C): SIRPB1C, CD14, FABP5, MEFV, PEPD, IL2RG, MMP14, CLEC4D, H2-AA, GYPC. First, the expression matrix of these genes was normalized, and LASSO regression analysis identified three core genes: SIRPB1C, FABP5, and PEPD (FIG.7D-E). The expression of these three core genes was observed in the merged transcriptome data (Figures 7F-G-H). RT-qPCR experiments were conducted in RAW264.7 cells to verify gene expression differences. Results showed high expression of TPP1, SPP1, SIRPB1C, FABP5, and PEPD in the *Borrelia burgdorferi* infection group (Figures 7I-J-K).

3.7 Construction and testing of a signature gene-based nomogram for predicting Lyme arthritis

Using the "RMS" R package, we constructed a diagnostic nomogram model for Lyme arthritis based on characteristic genes (SIRPB1C, FABP5, and PEPD) (Figure 8A), and evaluated its predictive capability using calibration curves. The calibration curve showed minimal difference between actual Lyme arthritis risk and predicted risk, indicating high accuracy of the nomogram Lyme arthritis model (Figure 8B). Decision curve analysis (DCA) demonstrated that patients could benefit from this nomogram (Figure 8C). ROC curve analysis also confirmed the accuracy of this model (Figure 8D).

4. Discussion

Lyme arthritis (LA) is one of the most common late-stage manifestations of Lyme

disease. Without timely diagnosis and treatment, it can lead to persistent arthritis and long-term functional impairment [33]. Although antibiotic therapy is effective in most cases, approximately 10-20% of patients develop antibiotic-refractory Lyme arthritis (ARLA) [34, 35]. The emergence of single-cell RNA sequencing (scRNA-seq) technology has provided us with unprecedented opportunities to understand disease pathogenesis at the single-cell level [36, 37]. This technology enables us to reveal cellular composition and transcriptomic changes under pathological conditions at single-cell resolution, thereby helping to identify features associated with LA pathogenesis and prognosis. Macrophages, as key effector cells of the innate immune system, play a central role in the pathogenesis of LA [38]. These cells not only participate in the regulation of inflammatory responses but also influence disease progression through the secretion of cytokines and chemokines.

C57BL/6 mice, being experimental animals with a uniform genetic background, provide excellent experimental conditions for investigating transcriptional regulation and cellular changes during infection. Notably, in the C57BL/6 mouse LA model, arthritis symptoms spontaneously resolve over time even when bacteria remain present, potentially providing a good simulation of how the host immune system gradually adapts to this non-life-threatening chronic infection [39-42]. Specifically, as the disease progresses, the number of *Borrelia burgdorferi* in mice gradually decreases while the host immune response undergoes changes, achieving long-term control of this bacterial infection [18]. This self-healing process provides important clues for understanding the self-limiting nature of LA.

The study utilized the GSE233850 scRNA-seq dataset to analyze the differentially expressed gene (DEG) profiles of macrophages in synovial tissue from C57BL/6 animal LA models. By integrating the research team's Bulk-transcriptome data and GSE125503 Bulk-transcriptome data through joint analysis and machine learning methods, the study explored core differential genes related to LA macrophages and conducted in vitro experimental validation. The synovial microenvironment (SME) plays a decisive role in LA development. It is a complex microenvironment formed by interactions between various immune cells, synovial cells, and stromal cells [43].

Studies have shown that SME can significantly influence disease progression and prognosis [44]. Compared to other types of arthritis, ARLA patients often have poorer prognosis and unsatisfactory treatment outcomes [5]. Previous studies have shown that in some LA patients, abnormal macrophage activation and inflammatory responses can persist even after clearing *Borrelia burgdorferi* infection [45]. Since infiltrating macrophages in synovial tissue may promote or inhibit inflammatory progression through different polarization states (M1/M2), understanding the related molecular mechanisms is crucial for improving treatment outcomes [46]. Currently, scRNA-seq technology has become a key tool for characterizing immune cells within synovial tissue, particularly in the context of inflammatory and autoimmune diseases such as psoriatic arthritis (PsA) and rheumatoid arthritis (RA) [47, 48]. Therefore, systematically studying the functional states of macrophages in the LA synovial microenvironment and their regulatory networks, and establishing a prognosis prediction model based on macrophage characteristics has important clinical significance.

Based on single-cell transcriptome sequencing analysis of ankle joint tissues from 2 uninfected control and 4 *Borrelia burgdorferi*-infected C57BL/6 mice, 8 cell types were identified after annotation, including macrophages, T/NK cells, B cells, DC cells, endothelial cells, fibroblasts, synovial cells, and muscle cells. Our comparative analysis revealed significant changes in cellular composition from normal to Lyme arthritis state. The proportion of macrophages increased substantially, becoming the predominant cell type. Conversely, the proportion of fibroblasts decreased significantly. B cells showed some increase, while muscle cells almost disappeared in the Lyme arthritis state. These changes indicated massive immune cell infiltration in the synovial microenvironment, particularly the notable increase in macrophages.

According to the reference [18], macrophages were categorized into four distinct subgroups (Mac1-Mac4) via unsupervised clustering analysis. In the context of the Lyme arthritis model, single-cell sequencing elucidated unique functional characteristics among these macrophage subgroups, notably highlighting MAC1's pivotal roles in tissue remodeling and immune regulation through essential genes such

as Pf4, Aqp1and Mrc1 [49-51]. The MAC2 subgroup is distinguished by elevated levels of pro-inflammatory genes such as Hdc, G0s2, Retnlg, and S100a8/S100a9, with the S100a8/S100a9 complex acting as a pivotal mediator in acute inflammation [52]. while HDC, as the key enzyme in histamine production, exhibits heightened expression correlated with intensified inflammatory responses [53]. The MAC3 subgroup is characterized by the expression of genes such as Thbs1, Ifitm6, Ifitm3, Plac8, and Ly6c2, which are primarily involved in immune modulation and the persistence of chronic inflammation, with Ifitm genes indicating possible antiviral activity [54], Thbs1 linked to immune cell infiltration in sepsis [55], and Ly6c2 serving as a marker for inflammatory monocytes [56]. The MAC4 subgroup, prominently expresses genes like Timp2, Sparc, Vsig4, and Cbr2, which are vital for tissue repair and calming inflammation, with Timp2 and Sparc aiding in extracellular matrix remodeling [57, 58], and Vsig4 exerting immunosuppressive effects linked to heightened NLRP3 inflammasome activity when deficient [59].

These findings unveiled a sophisticated immune regulatory network among the four macrophage subpopulations: MAC1, MAC2, and MAC3 orchestrate pro-inflammatory responses while executing distinct functional specializations - tissue remodeling, acute inflammatory responses, and chronic inflammation maintenance, respectively. In contrast, MAC4 functions as a regulatory sentinel, attenuating excessive inflammatory cascades through tissue regeneration and anti-inflammatory mechanisms. The river plot analysis of macrophage subpopulations in the Lyme arthritis model demonstrated an augmented presence of pro-inflammatory macrophage subsets (Mac1, Mac2, and Mac3), coupled with a diminished representation of tissue-homeostatic macrophages (Mac4), corroborating our observations and highlighting the dynamic phenotypic plasticity of macrophages under pathological conditions. Moreover, examination of macrophage pseudo-temporal trajectories revealed four distinct gene clusters that evolve along the developmental continuum, encompassing immune cell recruitment and inflammatory modulation, inflammatory response and tissue remodeling dynamics, antigen presentation and immune activation processes, and protein homeostasis regulation.

Cell-cell communication analysis revealed significant alterations in galectin-receptor interactions during pathological states, highlighting these molecules as compelling therapeutic targets for Lyme arthritis intervention. Heat map analyses demonstrated that under disease conditions, GALECTIN signaling pathway gene expression predominantly manifests in specific cellular interactions: between macrophages and endothelial cells, endothelial cells and B cells, endothelial cells and dendritic cells, and B cells and fibroblasts. Furthermore, our investigation unveiled critical ligand-receptor pairs (GALS9/CD45, LGALS9/CD44, LGALS9/lghm) within GALECTIN signaling networks, exhibiting enhanced intercellular communication across diverse cell types in Lyme arthritis compared to control conditions. Prior research has established Galectin-9's pivotal role in macrophage polarization dynamics, whereby its elevated expression promotes M2-phenotype development, characterized by amplified CD206 expression and attenuated pro-inflammatory cytokine production, including tumor necrosis factor α and IL-6. Conversely, Gal-9 suppression induces M1-phenotype polarization, marked by heightened NOS2 expression and pro-inflammatory cytokine production - a regulatory mechanism crucial in inflammatory conditions, neoplastic processes, and gestational states [60-63]. Consequently, therapeutic strategies targeting galectin pathways may represent a pioneering approach in Lyme arthritis treatment, potentially yielding improved clinical outcomes.

KEGG enrichment analysis of differentially expressed genes in macrophages revealed that these genes were primarily enriched in pathways related to lysosomes, bacterial infections (such as Salmonella and Yersinia infections), chemokine signaling pathways, endocytosis, and apoptosis. This indicates that in the Lyme arthritis environment, macrophages may be more actively involved in pathogen defense, cell signaling, and cell death processes. GO functional enrichment analysis found that in Biological Process (BP), these genes mainly participate in immune response regulation, leukocyte migration, and phagocytosis, highlighting the core immunoregulatory function of macrophages in the inflammatory microenvironment. Regarding Cellular Component (CC), the differentially expressed genes were

primarily enriched in membrane structures such as vacuolar membranes, endosomes, and membrane rafts, which are closely related to macrophage phagocytosis and antigen presentation functions. At the Molecular Function (MF) level, these genes mainly involve GTPase activity and phospholipid binding, which play crucial roles in cell signal transduction and inflammation regulation. These results collectively indicate that macrophages undergo significant functional remodeling in the synovial microenvironment of Lyme arthritis. They may be more actively engaged in immune regulation, pathogen defense, and cell-cell interactions, while their migration capabilities may be altered.

In this research, we developed a Lyme arthritis model in C57BL/6 mice and conducted transcriptome sequencing on infected ankle joint tissues. We retrieved a public dataset (GSE125503, GPL17021) that contains transcriptomic data from C57BL/6 bone marrow-derived macrophages. Focusing on macrophage-specific roles in Lyme arthritis, we combined the two transcriptomic datasets and executed differential gene expression analysis. Utilizing LASSO regression on the differentially expressed genes, we identified three macrophage marker genes (SIRPB1C, FABP5, and PEPD).

SIRPB1C (Signal Regulatory Protein Beta 1C) is a key member of the signal regulatory protein β family. Despite the lack of direct reports on SIRPB1C, research indicates that the SIRP family is intrinsically linked to inflammation and immunity. Investigations into SIRPB1 knockout in macrophages revealed diminished levels of inflammatory cytokines like IL1RA, CCL2, and IL-8, underscoring its function in inflammation regulation [64]. Another constituent of the SIRP family, SIRPα, functions as an immune inhibitory receptor that attenuates phagocytosis and the migration of immune cells; agonistic antibodies directed against SIRPα have proven effective in alleviating inflammatory disorders such as pulmonary inflammation and colitis by limiting the infiltration of neutrophils and monocytes [65, 66]. Fatty acid-binding protein 5 (FABP5) is a lipid chaperone implicated in inflammatory pain modulation. Recent investigations indicate that FABP5 is present in human synovium [67, 68], and its inhibition markedly decreases cytokine and chemokine secretion,

highlighting its potential influence on osteoarthritis-related inflammatory pain [69]. Additionally, FABP5 is notably upregulated in intestinal macrophages during inflammatory bowel diseases, and its inhibition can thwart macrophage differentiation into the inflammatory M1 phenotype, thereby alleviating DSS-induced colitis in murine models [70]. The PEPD (Peptidase D) gene encodes prolidase, which is vital for collagen metabolism and tissue integrity [71, 72]. Additionally, PEPD is implicated in inflammatory processes and obesity-related metabolic disorders, as its reduced expression in adipose tissue contributes to fibrosis and insulin resistance [73]. Furthermore, PEPD has functions that inhibit p53, which are linked to cancer development [74]. These association studies reinforce our findings regarding macrophage marker genes in Lyme arthritis, suggesting potential targets for investigating the molecular mechanisms of its progression.

Although our research provides valuable insights, it has several limitations that need to be addressed in future studies. First, the sample size of the integrated bulk RNA sequencing data is relatively limited. Second, regarding the exploration of molecular mechanisms: while we validated the transcriptional changes of core differential genes through in vitro experiments, we have not fully understood the functional roles of these genes in LA pathogenesis. Future research should aim to validate these findings in larger cohorts and employ functional experiments to confirm the roles of these cells, genes, and pathways in disease progression.

5. Conclusion

In conclusion, through innovative single-cell and transcriptome integrated analysis, we systematically investigated the functional characteristics of macrophages in LA synovial tissue. Furthermore, through cell clustering analysis, pseudotime trajectory analysis, and cell-cell communication network analysis, we revealed the dynamic changes of macrophages in the synovial microenvironment and their interaction patterns with other immune cells. Finally, we identified and validated three characteristic genes (SIRPB1C, FABP5, and PEPD) that can serve as diagnostic markers for Lyme arthritis through macrophage marker gene identification, and conducted in vitro validation. These findings provide new insights into understanding

the pathogenesis of LA and potential molecular targets for optimizing disease treatment strategies.

Author Affiliations

[1]Department of Pain Medicine, The Affiliated Anning First People's Hospital of Kunming University of Science and Technology, Kunming, Yunnan Province, 650302, China.

[2]School of Basic Medical Sciences, Kunming University of Science and Technology, Kunming, Yunnan Province, 650500, China.

Acknowledgement We extend our thanks to Professor Yi-Qun Kuang for providing the Biosafety Laboratory of the Research Center for Clinical Medicine, First Affiliated Hospital of Kunming Medical University, Kunming Medical University; and the publicly available datasets from the GEO databases.

Contributors Yan Dong conducted comprehensive bioinformatics analyses; Yanshuang Luo, Meng Liu, Yantong Chen, Xuesong Chen, Xiaorong Liu, Fusong Yang, Chao Song, Zhi Liang, Liye Zi, Hongbing Gao, Zhichao Cao, and Fucun Gao executed experimental procedures, performed data interpretation, and drafted the manuscript; Guozhong Zhou conceptualized the study design and provided overall project supervision. All authors reviewed and endorsed the final manuscript for publication.

Funding The study was supported by the Yunnan Fundamental Research Projects (grant NO. 202301BE070001-036) and the internal projects of the First People's Hospital of Anning, affiliated with Kunming University of Science and Technology (grant NO. 2024AYY001 and 2024AYY005).

Competing interests The author declares no competing interests.

Ethical approval This study was approved by the Scientific Research Ethics Committee of the Jiaxing University Affiliated Maternity and Child Hospital (No. 2021-65), and informed consent was obtained from each participant.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Publicly available datasets that support the findings of this study are openly available in the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/info/datasets.html). This data can be found here: GSE233850 and GSE125503.

Code availability any additional information required to re-analyze the data reported in this work paper is available from the lead contact upon request.

Publisher's note All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

[1] Lochhead R.B., Strle K., Arvikar S.L., Weis J.J., Steere A.C. Lyme arthritis: linking infection, inflammation and autoimmunity. Nat Rev Rheumatol 17 (8) (2021) 449-461. doi:10.1038/s41584-021-00648-5.

[2] Kanjana K., Strle K., Lochhead R.B., Pianta A., Mateyka L.M., Wang Q., et al. Autoimmunity to synovial extracellular matrix proteins in patients with postinfectious lyme arthritis. J Clin Invest 133 (17) (2023). doi:10.1172/JCI161170.

[3] Nadelman R.B., Hanincová K., Mukherjee P., Liveris D., Nowakowski J., Mckenna D., et al. Differentiation of reinfection from relapse in recurrent lyme disease. N Engl J Med 367 (20) (2012) 1883-1890. doi:10.1056/NEJMoa1114362.

[4] Coburn J., Garcia B., Hu L.T., Jewett M.W., Kraiczy P., Norris S.J., et al. Lyme disease pathogenesis. Curr Issues Mol Biol 42 (2021) 473-518. doi:10.21775/cimb.042.473.

[5] Brouwer M., van de Schoor F.R., Vrijmoeth H.D., Netea M.G., Joosten L. A joint effort: the interplay between the innate and the adaptive immune system in lyme arthritis. Immunol Rev 294 (1) (2020) 63-79. doi:10.1111/imr.12837.

[6] Steere A.C. Posttreatment lyme disease syndromes: distinct pathogenesis caused by maladaptive

host responses. J Clin Invest 130 (5) (2020) 2148-2151. doi:10.1172/JCI138062.
[7] Lasky C.E., Olson R.M., Brown C.R. Macrophage polarization during murine lyme borreliosis. Infect Immun 83 (7) (2015) 2627-2635. doi:10.1128/IAI.00369-15.
[8] Shin J.J., Strle K., Glickstein L.J., Luster A.D., Steere A.C. Borrelia burgdorferi stimulation of chemokine secretion by cells of monocyte lineage in patients with lyme arthritis. Arthritis Res Ther 12 (5) (2010) R168. doi:10.1186/ar3128.
[9] Codolo G., Bossi F., Durigutto P., Bella C.D., Fischetti F., Amedei A., et al. Orchestration of inflammation and adaptive immunity in borrelia burgdorferi-induced arthritis by neutrophil-activating protein a. Arthritis Rheum 65 (5) (2013) 1232-1242. doi:10.1002/art.37875.
[10] Codolo G., Bossi F., Durigutto P., Bella C.D., Fischetti F., Amedei A., et al. Orchestration of inflammation and adaptive immunity in borrelia burgdorferi-induced arthritis by neutrophil-activating protein a. Arthritis Rheum 65 (5) (2013) 1232-1242. doi:10.1002/art.37875.
[11] Duchateau B.K., Munson E.L., England D.M., Lovrich S.D., Callister S.M., Jensen J.R., et al. Macrophages interact with enriched populations of distinct t lymphocyte subsets for the induction of severe destructive lyme arthritis. J Leukoc Biol 65 (2) (1999) 162-170. doi:10.1002/jlb.65.2.162.
[12] Benjamin S.J., Hawley K.L., Vera-Licona P., La Vake C.J., Cervantes J.L., Ruan Y., et al. Macrophage mediated recognition and clearance of borrelia burgdorferi elicits myd88-dependent and -independent phagosomal signals that contribute to phagocytosis and inflammation. BMC Immunol 22 (1) (2021) 32. doi:10.1186/s12865-021-00418-8.
[13] Harjacek M., Diaz-Cano S., Alman B.A., Coburn J., Ruthazer R., Wolfe H., et al. Prominent expression of mrna for proinflammatory cytokines in synovium in patients with juvenile rheumatoid arthritis or chronic lyme arthritis. J Rheumatol 27 (2) (2000) 497-503.
[14] Sonderegger F.L., Ma Y., Maylor-Hagan H., Brewster J., Huang X., Spangrude G.J., et al. Localized production of il-10 suppresses early inflammatory cell infiltration and subsequent development of ifn- γ -mediated lyme arthritis. J Immunol 188 (3) (2012) 1381-1393. doi:10.4049/jimmunol.1102359.
[15] Liu W., Jia J., Dai Y., Chen W., Pei G., Yan Q., et al. Delineating covid-19 immunological features using single-cell rna sequencing. Innovation (Camb) 3 (5) (2022) 100289. doi:10.1016/j.xinn.2022.100289.
[16] Onuora S. Single-cell rna sequencing sheds light on cell-type specific gene expression in immune cells. Nat Rev Rheumatol 18 (7) (2022) 363. doi:10.1038/s41584-022-00802-7.
[17] Su C., Xu Z., Shan X., Cai B., Zhao H., Zhang J. Cell-type-specific co-expression inference from single cell rna-sequencing data. Nat Commun 14 (1) (2023) 4846. doi:10.1038/s41467-023-40503-7.
[18] Helble J.D., Walsh M.J., Mccarthy J.E., Smith N.P., Tirard A.J., Arnold B.Y., et al. Single-cell rna sequencing of murine ankle joints over time reveals distinct transcriptional changes following borrelia burgdorferi infection. iScience 26 (11) (2023) 108217. doi:10.1016/j.isci.2023.108217.
[19] Butler A., Hoffman P., Smibert P., Papalexi E., Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36 (5) (2018) 411-420. doi:10.1038/nbt.4096.
[20] Stuart T., Butler A., Hoffman P., Hafemeister C., Papalexi E., Mauck W.R., et al. Comprehensive integration of single-cell data. Cell 177 (7) (2019) 1888-1902. doi:10.1016/j.cell.2019.05.031.
[21] Gong B., Zhou Y., Purdom E. Cobolt: integrative analysis of multimodal single-cell sequencing data.

Genome Biol 22 (1) (2021) 351. doi:10.1186/s13059-021-02556-z.
[22] Stuart T., Satija R. Integrative single-cell analysis. Nat Rev Genet 20 (5) (2019) 257-272. doi:10.1038/s41576-019-0093-7.
[23] Wu T., Hu E., Xu S., Chen M., Guo P., Dai Z., et al. Clusterprofiler 4.0: a universal enrichment tool for interpreting omics data. Innovation (Camb) 2 (3) (2021) 100141. doi:10.1016/j.xinn.2021.100141.
[24] Sherman B.T., Hao M., Qiu J., Jiao X., Baseler M.W., Lane H.C., et al. David: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res 50 (W1) (2022) W216-W221. doi:10.1093/nar/gkac194.
[25] Zhou Y., Zhou B., Pache L., Chang M., Khodabakhshi A.H., Tanaseichuk O., et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 10 (1) (2019) 1523. doi:10.1038/s41467-019-09234-6.
[26] Haghverdi L., Büttner M., Wolf F.A., Buettner F., Theis F.J. Diffusion pseudotime robustly reconstructs lineage branching. Nat Methods 13 (10) (2016) 845-848. doi:10.1038/nmeth.3971.
[27] Jin S., Plikus M.V., Nie Q. Cellchat for systematic analysis of cell-cell communication from single-cell transcriptomics. Nat Protoc (2024). doi:10.1038/s41596-024-01045-4.
[28] Love M.I., Huber W., Anders S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. Genome Biol 15 (12) (2014) 550. doi:10.1186/s13059-014-0550-8.
[29] Liu S., Wang Z., Zhu R., Wang F., Cheng Y., Liu Y. Three differential expression analysis methods for rna sequencing: limma, edger, deseq2. J Vis Exp (175) (2021). doi:10.3791/62528.
[30] Tibshirani R. The lasso method for variable selection in the cox model. Stat Med 16 (4) (1997) 385-395. doi:10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3.
[31] Jaume S.P., Skaltsa K., Pallarès N., Carrasco J.L. Thresholdroc: optimum threshold estimation tools for continuous diagnostic tests in r. J Stat Softw 82 (1548-7660) (2017) 0-0. doi:10.18637/jss.v082.i04.
[32] Yee J. Rlsm: r package for least squares monte carlo. arXiv (Cornell University) 0 (2018) 0-0. doi:null.
[33] Keyßer G. Lyme disease and lyme arthritis. Aktuelle Rheumatol 0 (0341-051X) (2024) 0-0. doi:10.1055/a-2215-0830.
[34] Steere A.C., Strle F., Wormser G.P., Hu L.T., Branda J.A., Hovius J.W., et al. Lyme borreliosis. Nat Rev Dis Primers 2 (2056-676X) (2016) 0-0. doi:10.1038/nrdp.2016.90.
[35] Onuora S. Lyme arthritis: spirochaete remnants could explain antibiotic-refractory lyme arthritis. Nat Rev Rheumatol 8 (8) (2012) 440. doi:10.1038/nrrheum.2012.115.
[36] Van de Sande B., Lee J.S., Mutasa-Gottgens E., Naughton B., Bacon W., Manning J., et al. Applications of single-cell rna sequencing in drug discovery and development. Nat Rev Drug Discov 22 (6) (2023) 496-520. doi:10.1038/s41573-023-00688-4.
[37] Papalexi E., Satija R. Single-cell rna sequencing to explore immune cell heterogeneity. Nat Rev Immunol 18 (1) (2018) 35-45. doi:10.1038/nri.2017.76.
[38] Arvikar S.L., Crowley J.T., Sulka K.B., Steere A.C. Autoimmune arthritides, rheumatoid arthritis, psoriatic arthritis, or peripheral spondyloarthritis following lyme disease. Arthritis Rheumatol 69 (1) (2017) 194-202. doi:10.1002/art.39866.
[39] Koloski C.W., Hurry G., Foley-Eby A., Adam H., Goldstein S., Zvionow P., et al. Male c57bl/6j mice have higher presence and abundance of borrelia burgdorferi in their ventral skin compared to female mice. Ticks Tick Borne Dis 15 (2) (2024) 102308. doi:10.1016/j.ttbdis.2024.102308.

[40] Hammond E.M., Baumgarth N. Cd4 t cell responses in persistent borrelia burgdorferi infection. Curr Opin Immunol 77 (2022) 102187. doi:10.1016/j.coi.2022.102187.

[41] Bernard Q., Wang Z., Di Nardo A., Boulanger N. Interaction of primary mast cells with borrelia burgdorferi (sensu stricto): role in transmission and dissemination in c57bl/6 mice. Parasit Vectors 10 (1) (2017) 313. doi:10.1186/s13071-017-2243-0.

[42] Malovrh T., Skoberne M., Gruntar I., Kotnik V. The cell-mediated immune response to borrelia afzelii, garinii and burgdorferi in c57bl/6 mice is dependent on antigen specificity. FEMS Immunol Med Microbiol 28 (3) (2000) 233-240. doi:10.1111/j.1574-695X.2000.tb01482.x.

[43] Akin E., Aversa J., Steere A.C. Expression of adhesion molecules in synovia of patients with treatment-resistant lyme arthritis. Infect Immun 69 (3) (2001) 1774-1780. doi:10.1128/IAI.69.3.1774-1780.2001.

[44] Steere A.C., Duray P.H., Butcher E.C. Spirochetal antigens and lymphoid cell surface markers in lyme synovitis. Comparison with rheumatoid synovium and tonsillar lymphoid tissue. Arthritis Rheum 31 (4) (1988) 487-495. doi:10.1002/art.1780310405.

[45] Shin J.J., Glickstein L.J., Steere A.C. High levels of inflammatory chemokines and cytokines in joint fluid and synovial tissue throughout the course of antibiotic-refractory lyme arthritis. Arthritis Rheum 56 (4) (2007) 1325-1335. doi:10.1002/art.22441.

[46] Zerrouk N., Alcraft R., Hall B.A., Augé F., Niarakis A. Large-scale computational modelling of the m1 and m2 synovial macrophages in rheumatoid arthritis. NPJ Syst Biol Appl 10 (1) (2024) 10. doi:10.1038/s41540-024-00337-5.

[47] Bruno D., Somma D., Tolusso B., Di Mario C., Coletto L.A., Elmesmari A., et al. Pos0425 single-cell rna sequencing of synovial tissue-derived myeloid cells in psoriatic arthritis patients across disease phases. null 0 (2023) 0-0. doi:10.1136/annrheumdis-2023-eular.3047.

[48] Edalat S.G., Gerber R., Houtman M., Kuret T., Izanc N., Micheroli R., et al. A reference single-cell map of freshly dissociated human synovium in inflammatory arthritis with an optimized dissociation protocol for prospective synovial biopsy collection. bioRxiv (Cold Spring Harbor Laboratory) 0 (2022) 0-0. doi:10.1101/2022.06.01.493823.

[49] павлов О.В., Chepanov S.V., Selutin A.V., Selkov S.A. Platelet-leukocyte interactions: immunoregulatory role and pathophysiological relevance. Медицинская иммунология 24 (1563-0625) (2022) 871-888. doi:10.15789/1563-0625-pli-2511.

[50] Wagner K., Unger L., Salman M.M., Kitchen P., Bill R.M., Yool A.J. Signaling mechanisms and pharmacological modulators governing diverse aquaporin functions in human health and disease. Int J Mol Sci 23 (3) (2022). doi:10.3390/ijms23031388.

[51] Nour J., Moregola A., Svecla M., Da D.L., Bellini R., Neyrolles O., et al. Mannose receptor deficiency impacts bone marrow and circulating immune cells during high fat diet induced obesity. Metabolites 12 (12) (2022). doi:10.3390/metabo12121205.

[52] Möller A., Jauch-Speer S.L., Gandhi S., Vogl T., Roth J., Fehler O. The roles of toll-like receptor 4, cd33, cd68, cd69, or cd147/emmprin for monocyte activation by the damp s100a8/s100a9. Front Immunol 14 (2023) 1110185. doi:10.3389/fimmu.2023.1110185.

[53] Takai J., Ohtsu H., Sato A., Uemura S., Fujimura T., Yamamoto M., et al. Lipopolysaccharide-induced expansion of histidine decarboxylase-expressing ly6g(+) myeloid cells identified by exploiting histidine decarboxylase bac-gfp transgenic mice. Sci Rep 9 (1) (2019)

15603. doi:10.1038/s41598-019-51716-6.
[54] Schelle L., Abrantes J., Baldauf H.M., Esteves P.J. Evolution of primate interferon-induced transmembrane proteins (ifitms): a story of gain and loss with a differentiation into a canonical cluster and ifitm retrogenes. Front Microbiol 14 (2023) 1213685. doi:10.3389/fmicb.2023.1213685.
[55] Liu Y., Yu K., Li C. Identification of platelets related thbs1 as a critical gene in sepsis related thrombocytopenia via an integrated bioinformatic analysis. Research Square (Research Square) 0 (2022) 0-0. doi:10.21203/rs.3.rs-2239983/v1.
[56] Hoenow S., Yan K., Noll J., Groneberg M., Casar C., Lory N.C., et al. The properties of proinflammatory ly6c(hi) monocytes are differentially shaped by parasitic and bacterial liver infections. Cells 11 (16) (2022). doi:10.3390/cells11162539.
[57] Chen J., Qin S., Liu S., Zhong K., Jing Y., Wu X., et al. Targeting matrix metalloproteases in diabetic wound healing. Front Immunol 14 (2023) 1089001. doi:10.3389/fimmu.2023.1089001.
[58] Ghanemi A., Yoshioka M., St-Amand J. Secreted protein acidic and rich in cysteine as a regeneration factor: beyond the tissue repair. Life (Basel) 11 (1) (2021). doi:10.3390/life11010038.
[59] Huang X., Feng Z., Jiang Y., Li J., Xiang Q., Guo S., et al. Vsig4 mediates transcriptional inhibition of nlrp3 and il-1β in macrophages. Sci Adv 5 (1) (2019) eaau7426. doi:10.1126/sciadv.aau7426.
[60] Lv R., Bao Q., Li Y. Regulation of m1-type and m2-type macrophage polarization in raw264.7 cells by galectin-9. Mol Med Rep 16 (6) (2017) 9111-9119. doi:10.3892/mmr.2017.7719.
[61] Arikawa T., Saita N., Oomizu S., Ueno M., Matsukawa A., Katoh S., et al. Galectin-9 expands immunosuppressive macrophages to ameliorate t-cell-mediated lung inflammation. Eur J Immunol 40 (2) (2010) 548-558. doi:10.1002/eji.200939886.
[62] Morrison H.M., Craft J., Rivera-Lugo R., Johnson J.R., Golovkine G.R., Bell S.L., et al. Deficiency in galectin-3, -8, and -9 impairs immunity to chronic mycobacterium tuberculosis infection but not acute infection with multiple intracellular pathogens. PLoS Pathog 19 (6) (2023) e1011088. doi:10.1371/journal.ppat.1011088.
[63] Sumbayev V.V., Yasinska I., Schlichtner S., Lall G.S., Fasler-Kan E., Bf G. P01.07 galectin 9 dependent immune evasion machinery as a potential target for personalised immunotherapy of cancer. Poster presentations 0 (2022) 0-0. doi:10.1136/jitc-2022-itoc9.19.
[64] Tang J.M., Wan X., Zhang J., Diao N., Zhang C., Gao X., et al. A frameshift variant in the sirpb1 gene confers susceptibility to crohn's disease in a chinese population. Front Genet 14 (1664-8021) (2023) 0-0. doi:10.3389/fgene.2023.1130529.
[65] Kanazawa Y., Saito Y., Supriatna Y., Tezuka H., Kotani T., Murata Y., et al. Role of sirpα in regulation of mucosal immunity in the intestine. Genes Cells 15 (12) (2010) 1189-1200. doi:10.1111/j.1365-2443.2010.01453.x.
[66] Kong X.N., Yan H.X., Chen L., Dong L.W., Yang W., Liu Q., et al. Lps-induced down-regulation of signal regulatory protein {alpha} contributes to innate immune activation in macrophages. J Exp Med 204 (11) (2007) 2719-2731. doi:10.1084/jem.20062611.

[67] Zhou Y., Elmes M.W., Sweeney J.M., Joseph O.M., Che J., Hsu H.C., et al. Identification of fatty acid binding protein 5 inhibitors through similarity-based screening. Biochemistry 58 (42) (2019) 4304-4316. doi:10.1021/acs.biochem.9b00625.

[68] Bando Y., Sakashita H., Nagasaka A., Sakiyama K., Tokuda N., Iseki S., et al. Septoclasts expressing epidermal fatty acid-binding protein (e-fabp, fabp5) in endochondral ossification. J Oral Biosci 64 (1) (2022) 18-25. doi:10.1016/j.job.2021.12.003.

[69] Gordon C., Trainor J., Shah R.J., Studholme K.M., Gelman A., Doswell F., et al. Fatty acid binding protein 5 inhibition attenuates pronociceptive cytokine/chemokine expression and suppresses osteoarthritis pain: a comparative human and rat study. Osteoarthritis Cartilage 0 (1063-4584) (2023) 0-0. doi:10.1016/j.joca.2023.11.010.

[70] Xu J., Zheng B., Xie C., Zhao Y., Wu H., Wang Y., et al. Inhibition of fabp5 attenuates inflammatory bowel disease by modulating macrophage alternative activation. Biochem Pharmacol 219 (0006-2952) (2024) 115974. doi:10.1016/j.bcp.2023.115974.

[71] Karna E., Szoka L., Huynh T., Palka J.A. Proline-dependent regulation of collagen metabolism. Cell Mol Life Sci 77 (10) (2020) 1911-1918. doi:10.1007/s00018-019-03363-3.

[72] Misiura M., Miltyk W. Current understanding of the emerging role of prolidase in cellular metabolism. Int J Mol Sci 21 (16) (2020). doi:10.3390/ijms21165906.

[73] Pellegrinelli V., Rodriguez-Cuenca S., Rouault C., Figueroa-Juarez E., Schilbert H., Virtue S., et al. Dysregulation of macrophage pepd in obesity determines adipose tissue fibro-inflammation and insulin resistance. Nat Metab 4 (4) (2022) 476-494. doi:10.1038/s42255-022-00561-5.

[74] Yang L., Li Y., Bhattacharya A., Zhang Y. Pepd is a pivotal regulator of p53 tumor suppressor. Nat Commun 8 (1) (2017) 2052. doi:10.1038/s41467-017-02097-9.

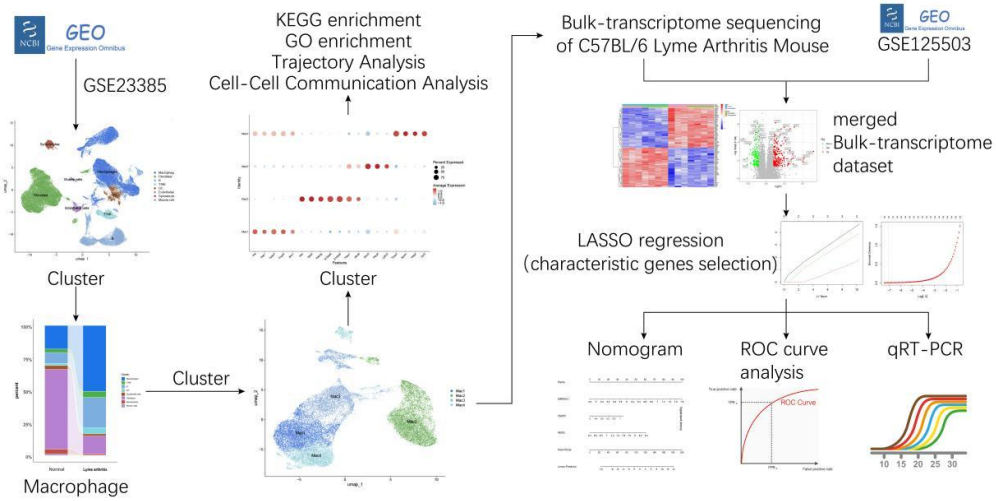


Figure 1. A detailed flow chart in this investigation.

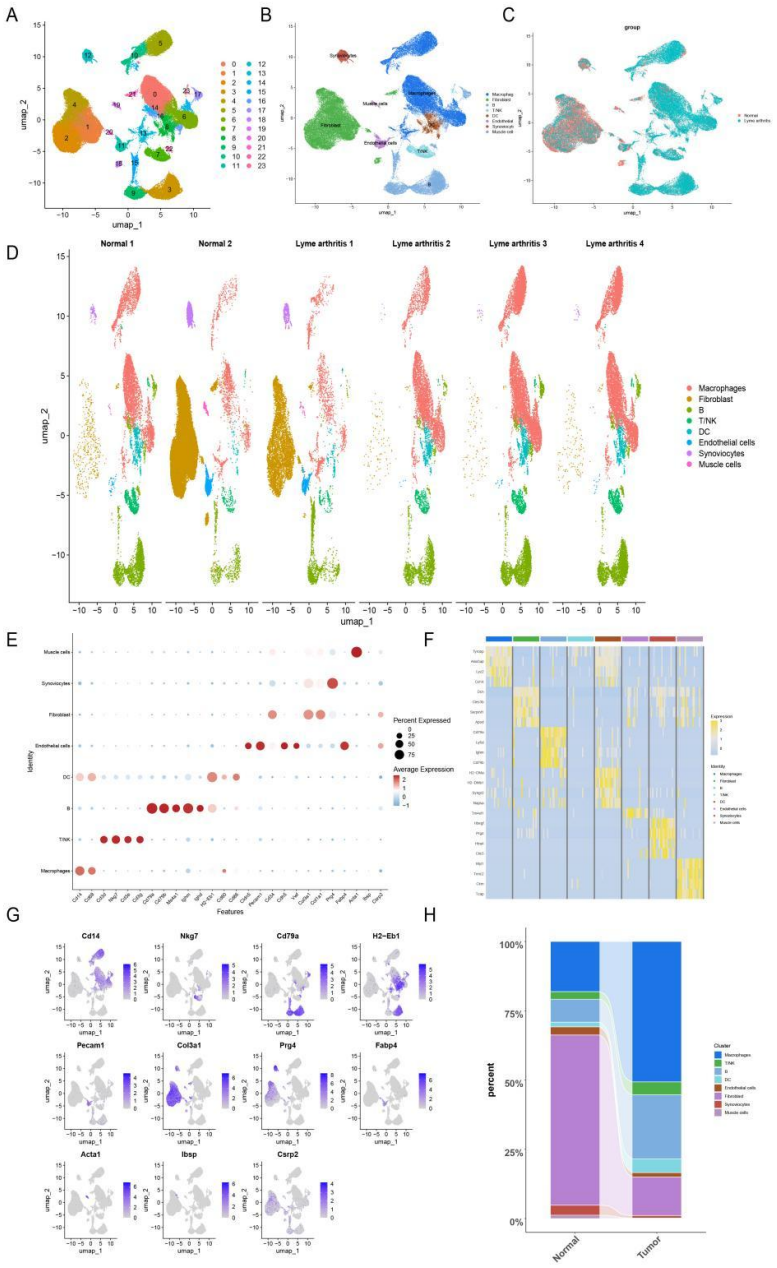


Figure 2. Overview of single-cell atlas of normal and Lyme arthritis C57BL/6 mouse samples. (A) UMAP plot showing clustering of single-cell samples into 24 clusters; (B) Eight cell types identified based on marker gene expression; (C) UMAP clustering plot comparing normal and Lyme arthritis model groups; (D) UMAP clustering plot of six samples; (E) Bubble plot showing marker gene expression across eight cell types; (F) Heatmap showing differential gene expression among eight cell types; (G) UMAP plot highlighting expression patterns of marker genes for eight cell types; (H) Cell type proportions in normal and Lyme arthritis model groups from scRNA-seq dataset.

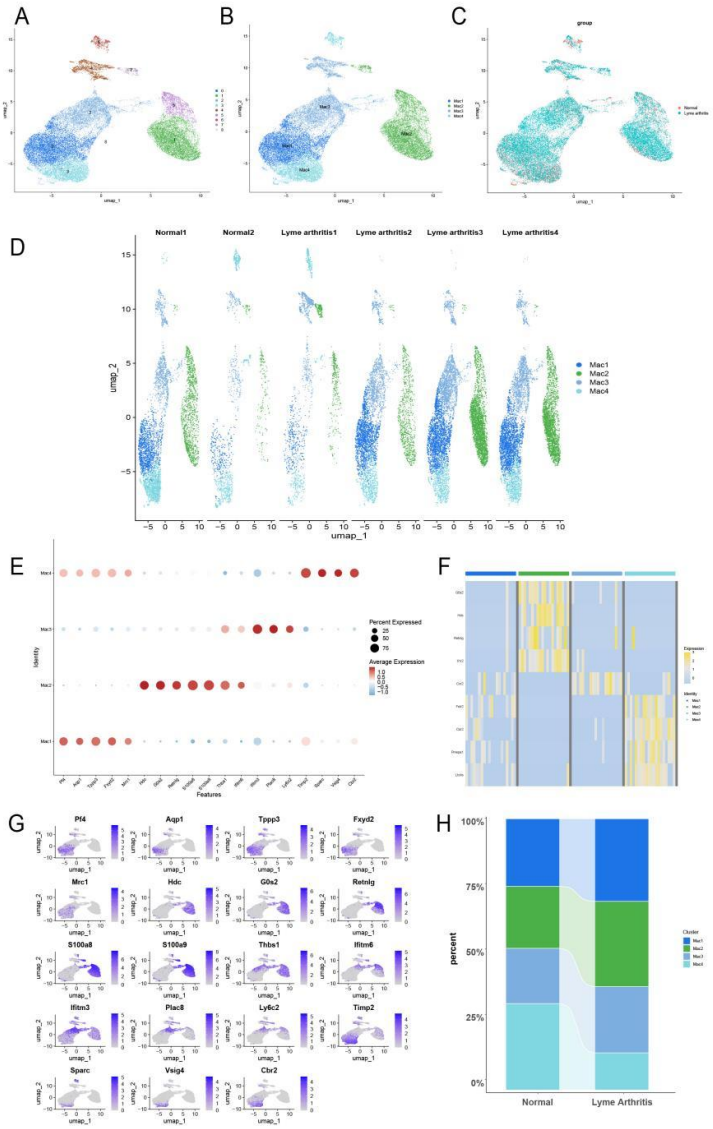


Figure 3. Single-cell transcriptome atlas of macrophages in scRNA-seq. (A) UMAP plot of macrophages in scRNA-seq, clustered into 9 distinct clusters; (B) UMAP plot highlighting the identification of four macrophages subtypes characterized by their signature molecular marker; (C) Comparative UMAP clustering plot between normal and Lyme arthritis model groups; (D) UMAP clustering plot of six samples; (E) Bubble plot visualization depicting the differential expression patterns of signature genes across four distinct macrophage subpopulations; (F) Heatmap showing differential gene expression for each of the four macrophages subtypes; (G) UMAP plot highlighting expression patterns of marker genes for the four cell types; (H) Cell type proportions between normal and Lyme arthritis model groups in scRNA-seq dataset.

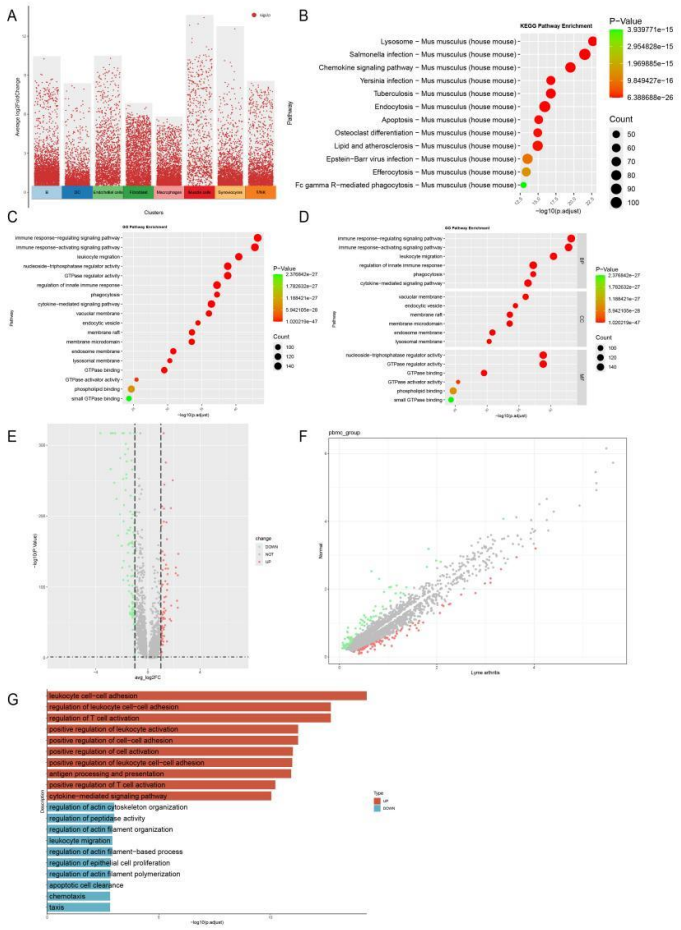


Figure 4. Enrichment Analysis Between Specific Subgroup Cells and Groups. (A) Volcano plot illustrating differential gene expression analysis across cellular subpopulations, wherein red dots denote cluster-specific genes exhibiting significant enrichment compared to other cellular clusters; (B) KEGG pathway enrichment analysis depicting the functional distribution of differentially expressed genes among macrophage populations; (C) GO functional enrichment plot of differentially expressed genes between macrophage cells; (D) GO functional enrichment analysis visualizing differentially expressed genes in macrophages, stratified by biological processes (BP), molecular functions (MF), and cellular components (CC) ontological categories; (E) Scatter plot of differential analysis between macrophage groups; (F) Diagonal volcano plots illustrating differential gene expression analysis between macrophage subpopulations; (G) Bar plots depicting GO enrichment analysis of upregulated and downregulated genes among macrophage subpopulations.

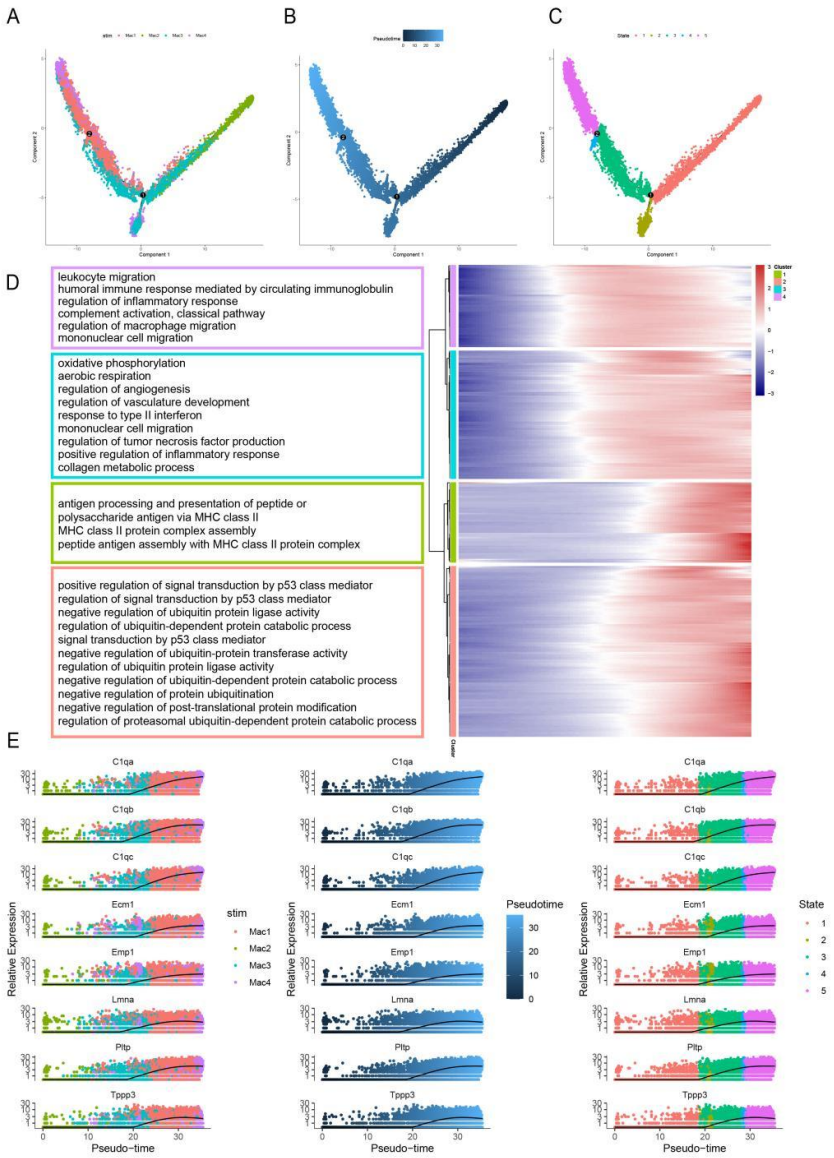


Figure 5. Pseudotime analysis of Macrophage in scRNA-seq. (A) Tajectory plot displaying the identified clusters of Macrophage extracted in scRNA-seq; (B) Trajectory plot depicting the developmental time course of Macrophage; (C) Trajectory plot showcasing the different stages of Macrophage differentiation along a pseudo-time axis; (D) Heatmap of pseudo-temporal differentially expressed genes and their enrichment patterns (displaying genes with temporal variations in the pseudo-timeline, clustered into four groups; the text in the left bar represents GO enrichment pathways); (E) Scatter plot of gene expression changes in pseudotime (for observing temporal changes of genes of interest).

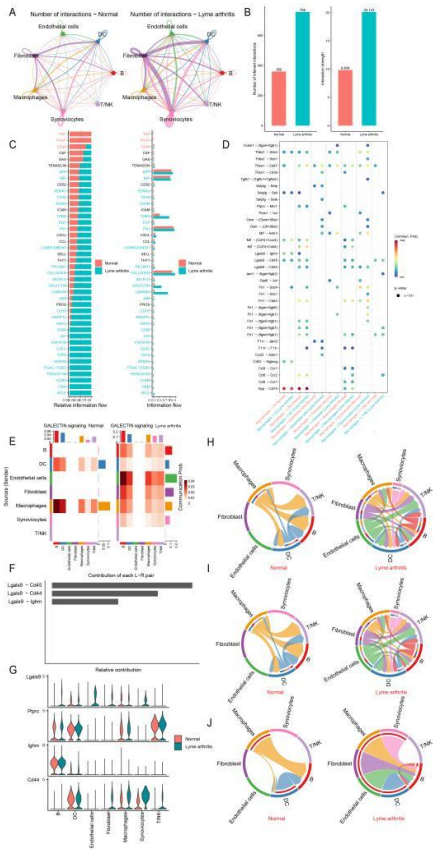


Figure 6. Cell-cell communication analysis. (A) Network diagram showing the number of cell-cell interactions separately in normal and Lyme arthritis model groups; (B) Comparison of the number and strength of cellular communications between normal and Lyme arthritis model groups; (C) Total pathway analysis between Lyme arthritis model group and normal group; (D) Ligand-receptor pair analysis between macrophages and other cell types in control and model groups; (E) Heatmap of ligand-receptor interaction patterns in Galectin signaling pathway among different cell types in control and model groups; (F) Relative contribution of receptor-ligand pairs in Galectin pathway in the model group; (G) Expression profiles of representative genes in the Galectin signaling pathway across different cell types between model and normal samples. (H) Cell-cell communication mediated by single LGALS9/CD45 ligand-receptor pair in normal and model groups; (I) Cell-cell communication mediated by single LGALS9/CD44 ligand-receptor pair in normal and model groups; (J) Cell-cell communication mediated by single LGALS9/Ighm ligand-receptor pair in normal and model groups.

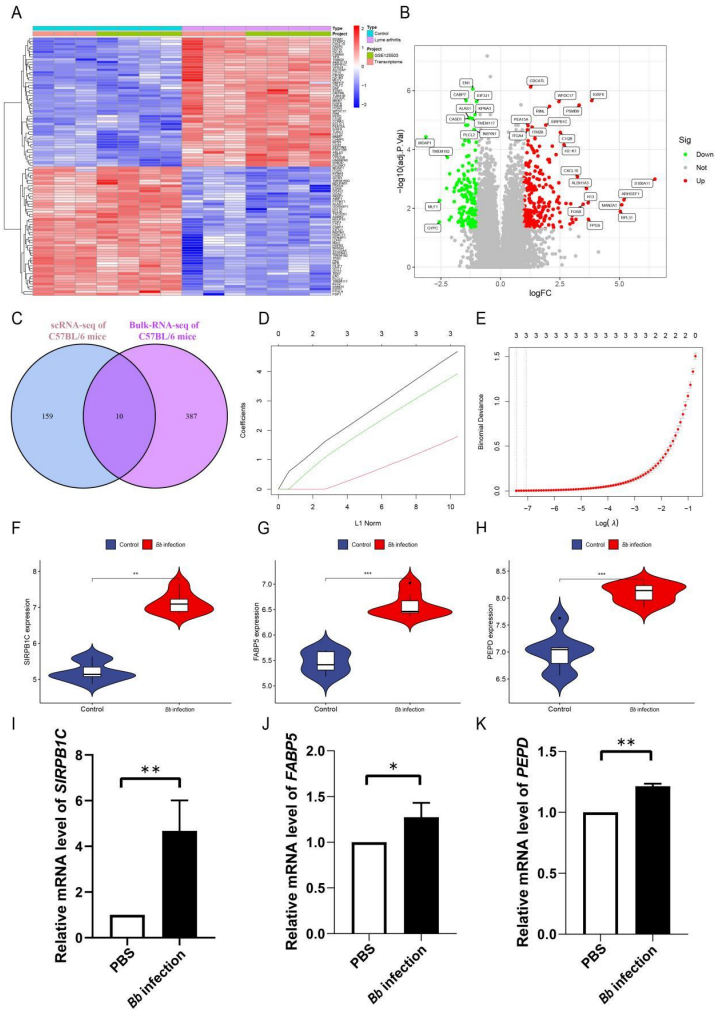


Figure 7. Prediction and Validation of Macrophage Predictive Gene Features. (A) Heatmap of differentially expressed genes in merged transcriptome; (B) Volcano plot of differentially expressed genes in merged transcriptome; (C) Venn diagram showing intersection of DEGs between merged transcriptome and macrophage-related genes in single-cell data: 10 differential genes; (D) Construction of LASSO regression model using merged transcriptome samples as training set; (E) Model parameter optimization cross-validation (left line indicates optimal λ value): resulting in 3 core genes (SIRPB1C, FABP5 and PEPD); (F-G-H) Expression of 3 core genes (SIRPB1C, FABP5 and PEPD) in merged transcriptome samples; (I-J-K) RT-qPCR in vitro experimental validation of expression of 3 core genes (SIRPB1C, FABP5 and PEPD). $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

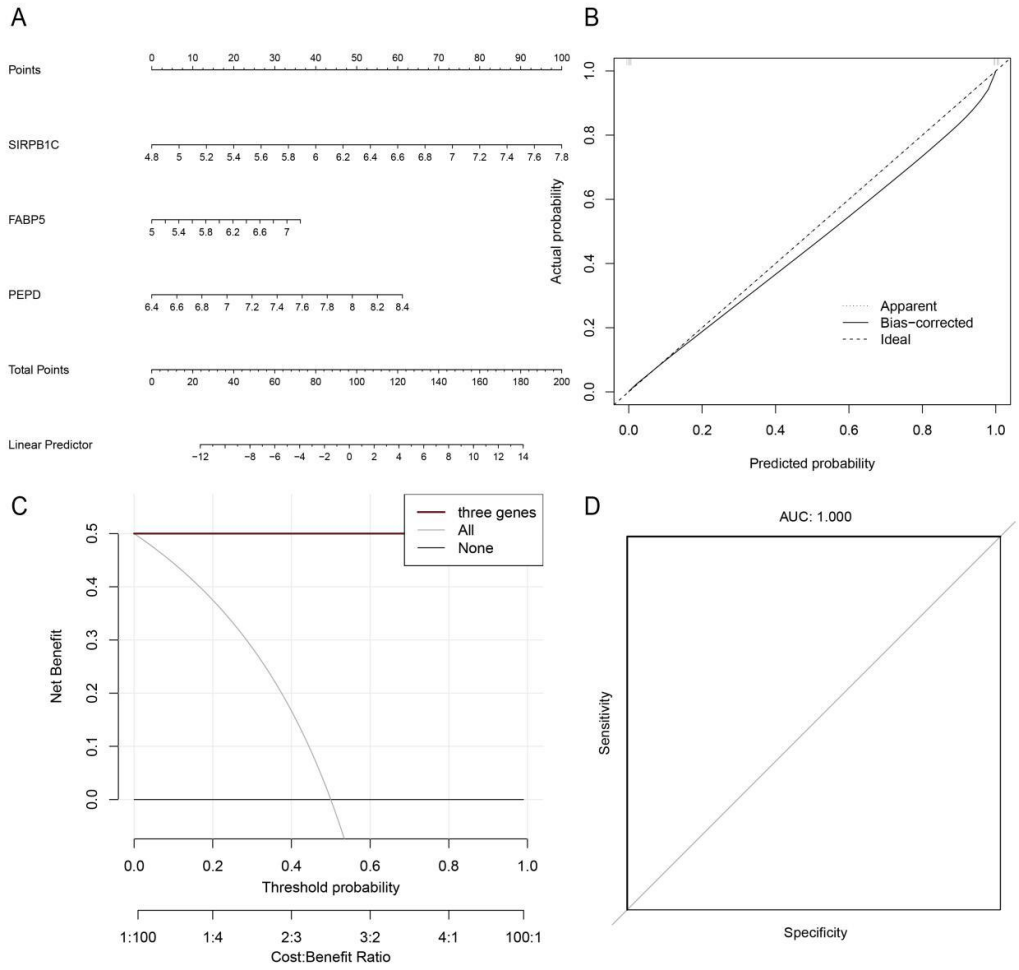
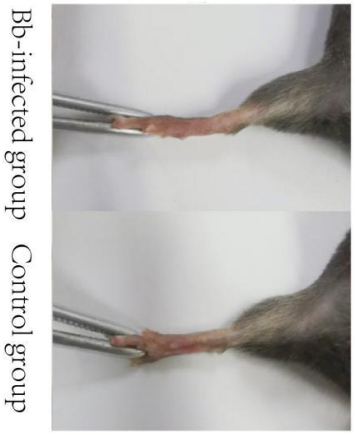
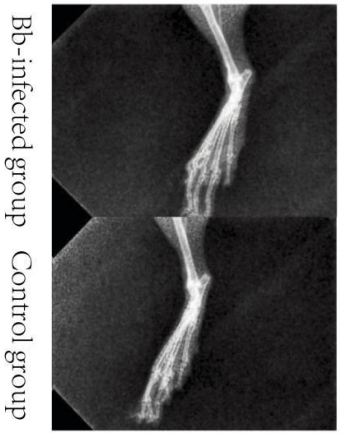


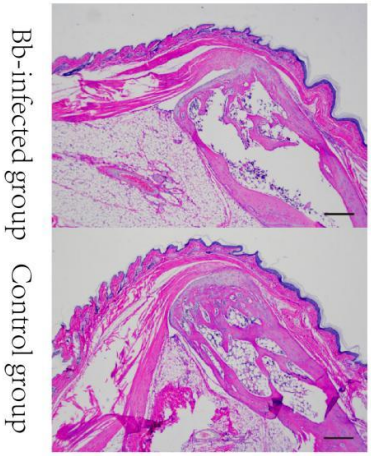
Figure 8. Construction and Validation of a Nomogram Model for Lyme Arthritis Diagnosis Based on Key Genes. (A) Using nomogram to predict the occurrence of Lyme arthritis; (B) Calibration curve to evaluate the predictive ability of the nomogram model; (C) DCA curve to assess the value of the nomogram model; (D) ROC curve to evaluate the nomogram model. Abbreviations: ROC, Receiver operating characteristic, DCA, Decision Curve Analysis.



Supplementary Figure 1. Swelling of hind limbs in C57BL/6 mice 16 days after *Borrelia burgdorferi* or PBS injection.



Supplementary Figure 2. X-ray of hind limbs in C57BL/6 mice 16 days after *Borrelia burgdorferi* or PBS injection.



Supplementary Figure 3. H&E staining of hind limbs in C57BL/6 mice 16 days after *Borrelia burgdorferi* or PBS injection.