scientific reports

Check for updates

OPEN

Exploring mitochondrial and ferroptotic mechanisms for systemic lupus erythematosus biomarker identification and therapy

Yunfeng Dai[1,2,4], Jianwen Liu[1,2,4], Yongxing Lai[1,2], Fei Gao[1,2], He Lin[1,2], Li Zhang[1,3✉] & Zhihan Chen[1,2✉]

Systemic lupus erythematosus (SLE) is a complex autoimmune disease with heterogeneous clinical manifestations. Understanding the molecular mechanisms of SLE is crucial for developing effective therapeutic strategies. This study downloaded microarray datasets from the Gene Expression Omnibus (GEO) database. Single-cell RNA sequencing (scRNA-seq) data was processed to identify 19 clusters and annotated five major cell types. Then we calculated mitochondrial-related genes (MRGs) and ferroptosis-related genes (FRGs) scores. FRGs scored the highest in Megakaryocytes, while MRGs scored the highest in B cells. By employing pseudotime analysis, cell-cell communication analysis, and Single-Cell Regulatory Network Inference and Clustering (SCENIC) analysis, we explored the heterogeneity of cells in SLE. Hub genes were identified using high-dimensional weighted correlation network analysis (hdWGNCA) and machine learning algorithms, leading to the development of a predictive diagnostic model with high predictive accuracy. Immune infiltration analysis revealed significant correlations between diagnostic biomarkers and various immune cells. Lastly, molecular docking studies suggested Doxorubicin may exert therapeutic effects by affecting these diagnostic biomarkers. This study offers new insights into the pathogenesis of SLE and provide valuable directions for future therapeutic research.

Keywords Systemic lupus erythematosus, Single-cell RNA sequencing, Ferroptosis, Mitochondria, Immune infiltration, Molecular Docking

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune connective tissue disease that has a wide spectrum of clinical manifestations[1]. According to a 2023 systematic review and modeling study that included 112 studies, the global SLE incidence and newly diagnosed population were estimated to be 5.14 (1.4–15.13) per 100 000 person-years and 0.40 million people annually, respectively[2]. The pathophysiology of SLE is intricate and influenced by multiple factors. Recent studies have implicated ferroptosis in the induction of immunocyte apoptosis, leading to the release of mitochondrial DNA (mtDNA) and damage-associated molecular patterns (DAMPs) in SLE, which activate inflammatory pathways, including the NLRP3 inflammasome and cGAS-STING pathway, thereby enhancing the production of type I interferons and pro-inflammatory cytokines, which in turn perpetuates autoimmunity and organ damage[3–5]. Moreover, mitochondrial dysfunction, marked by excessive mitochondrial reactive oxygen species (mROS), fuels immune cell activation, oxidative stress, and mtDNA oxidation[6,7]. This leads to impaired mitophagy and the accumulation of damaged mitochondria, creating a vicious cycle of inflammation and oxidative stress that is central to SLE progression. Our previous study indicated that increased mitochondrial fission may exacerbate podocyte injury and proteinuria, suggesting a potential role of mitochondrial dynamics in SLE pathogenesis[8]. Gaining a deeper understanding of the

[1]Fujian Medical University Shengli Clinical Medical College, Fuzhou 350000, China. [2]Department of Rheumatology, Fuzhou University Affiliated Provincial Hospital, No.134 Dongjie, Fuzhou 350000, China. [3]Department of Nephrology, Fuzhou University Affiliated Provincial Hospital, No.134 Dongjie, Fuzhou 350000, China. [4]Yunfeng Dai and Jianwen Liu contributed equally to this work. ✉email: lilydoctor9902@163.com; zhihanchen213@163.com

potential connections among immune function abnormalities, iron homeostasis disruptions, and mitochondrial dysfunction in SLE patients will enhance our knowledge of the disease and aid in prognostication.

Ferroptosis, characterized by glutathione depletion and lipid peroxidation (LPO), is a type of cell death dependent on iron[9]. This process is increasingly recognized for its role in a variety of pathological conditions. The mitochondrion, essential for ATP production and cellular metabolism, is particularly vulnerable to the damaging effects of LPO, which can initiate ferroptosis[10]. Given the mitochondria's pivotal role in maintaining cellular health, their susceptibility to LPO underscores a key mechanism that can lead to cellular dysfunction and demise[11]. LPO contributes to the pathological process of SLE by increasing oxidative stress, causing mitochondrial dysfunction, activating inflammatory pathways, promoting the production of autoantigens, and impairing mitophagy[12]. These observations align with our previous research, which explored the correlation between ferroptosis-related molecular subtypes and immunological characteristics in lupus nephritis, underscoring the importance of ferroptosis in the pathogenesis of SLE[13]. Building on these findings, our current study delves into the roles of ferroptosis-related genes (FRGs) and mitochondrial-related genes (MRGs). By exploring the expression patterns and regulatory networks of FRGs and MRGs, our research aims to uncover novel biomarkers and therapeutic targets for SLE. This could lead to more personalized treatment strategies and improved clinical outcomes for patients.

Single-cell RNA sequencing (scRNA-seq) provides valuable insights into the diversity of the immune system by identifying novel immune cell subsets, delineating intrinsic heterogeneity, and mapping developmental pathways[14]. The integration of scRNA-seq with bioinformatics analysis has become a powerful tool in studying autoimmune diseases, including the identification of key cellular and molecular mechanisms[15–17]. Despite these advances, the interplay between immunological dysregulation, ferroptosis, and mitochondrial dysfunction in SLE remains an area that is not fully explored. Further investigation is needed, particularly employing scRNA-seq in conjunction with advanced bioinformatics techniques, to uncover the complex interactions at play in SLE.

In the present study, we innovatively integrated the aforementioned methods to identify potential diagnostic genes for SLE and successfully constructed a diagnostic model. Moreover, we evaluated the interactions between these diagnostic markers and infiltrating immune cells. These findings provide clues for further exploring the pathogenesis of SLE and developing therapeutic strategies.

Methods

Data sources

In this study, we aim to investigate the differentially expressed genes (DEGs) between SLE and normal controls. Our methodology commenced with a search using the keywords "Systemic Lupus Erythematosus" and "Homo sapiens". Subsequently, we narrowed down the study types to "Expression profiling by array" or "Expression profiling by high throughput sequencing". We meticulously reviewed all resulting datasets and applied the following exclusion criteria: (i) datasets with a markedly insufficient sample size or those that failed to adequately represent the heterogeneity of SLE; (ii) datasets with low or incomplete data quality; (iii) datasets with restricted access; or (iv) datasets that did not comply with ethical standards.

Finally, we utilized the "GEOquery" R package to download the SLE related datasets GSE72326[18], GSE61635[19] and GSE50772[20] from the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) database. The GSE72326 dataset, which is based on the GPL10558 platform, includes a total of 177 samples, comprising 20 control samples and 157 SLE samples. The GSE61635 dataset, utilizing the GPL570 platform, consists of 129 samples, with 30 control samples and 99 SLE samples. Meanwhile, the GSE50772 dataset, utilizing the GPL570 platform, consists of 81 samples, with 20 control samples and 61 SLE samples. In this study, we designated the GSE72326 dataset as the training set, the GSE61635 dataset as the validation set and the GSE50772 dataset as the testing set. Expression data were preprocessed and normalized using quantile normalization with the "Limma" R package. Benjamini and Hochberg's approach was used to adjust P-values for controlling the false discovery rate (FDR).

We obtained the single cell sequencing (scRNA-seq) dataset GSE135779[21] from the GEO database, which corresponds to Homo sapiens and was assayed on the GPL20301 platform. For this study, we selected scRNA-seq data from seven adult SLE patients, specifically GSM4029943, GSM4029944, GSM4029945, GSM4029946, GSM4029950, GSM4029951, and GSM4029952. For the processing of single-cell data, we utilized the R package "Seurat" (version 4.4.0; https://github.com/satijalab/seurat)) to construct Seurat objects for seven samples. We filtered the gene expression data by retaining genes that were expressed in at least three cells and cells that expressed a minimum of 200 genes, using the parameters min.cells = 3 and min.features = 200. Subsequently, we refined our dataset by excluding cells with mitochondrial gene content exceeding 10%, ensuring that the number of detected features was greater than 300, and that the total RNA molecule count was between 1000 and 20,000, using the parameters percent.mt < 10, nFeature_RNA > 300, nCount_RNA > 1000, and nCount_RNA < 20000. After these steps, we obtained a dataset of 34,362 cells for subsequent analysis.

The mitochondrial protein database, MitoCarta3.0 (http://www.broadinstitute.org/mitocarta)[22], was visited to obtain 1,136 mitochondrial related genes (MRGs). The FerrDb (http://www.zhounan.org/ferrdb/current/)[23] is a manually curated resource for ferroptosis regulators and markers, as well as associations with ferroptosis-related diseases. And 111 ferroptosis-related genes (FRGs) were downloaded from the database. The gene lists are presented in Supplementary Table 1.

Cell type annotation

For the Seurat object of single-cell data, we initially applied the "NormalizeData" function of the R package "Seurat" (version 4.4.0) to standardize the gene expression data. We then proceeded to identify the 2000 most variable genes using the "FindVariableFeatures" function. Subsequently, we conducted Principal Component Analysis (PCA) with the "RunPCA" function and preserved the top 30 principal components for subsequent

analysis. We also employed the "RunHarmony" function for batch effect correction. The data were visualized using the Uniform Manifold Approximation and Projection (UMAP) algorithm, which delineated 19 distinct clusters. By manually annotating based on cell type-specific marker genes, we revealed 5 different cell types. The marker genes for CD4 memory T cells include CD3D and CD3E; for B cells, the markers are CD79A and MS4A1; for CD8 effector T cells, the markers are GNLY and GZMK; for Dendritic cells, the markers include CD14, MS4A7, FCER1A and CLEC4C. As for Megakaryocytes, the marker is PPBP. We used the function "FindAllMarkers" to calculate the differentially expressed genes (DEGs) between different cell populations. The threshold was set to logFC > 0.5 and adjusted p value < 0.05. Then we presented the analysis results through a heatmap.

Phenotype-related gene score analysis in single-cell data

The "UCell" R package is utilized to calculate gene signature enrichment scores in single-cell data. Leveraging the Mann-Whitney U test, it assesses single-cell feature scores for specified gene sets from scRNA-seq count matrices, known as UCell scores[24]. For the five unique cell types in SLE scRNA-seq, we determined UCell scores for both MRGs and FRGs. High-scoring cell populations were selected for further analysis, with results visualized through UMAP plots and violin plots for enhanced clarity.

Pseudotime analysis

Pseudotime analysis[25] enables the ordering of cells along a trajectory based on their temporal gene expression profiles. This method categorizes samples into cell populations representing various differentiation states and generates an intuitive, phylogenetic tree-like diagram that can predict the differentiation and developmental trajectories of cells. The outcomes of pseudotime analysis are validated by examining the distribution of cell type trajectories and the expression dynamics of signature genes to ascertain the starting and ending points of differentiation. We utilized the "Monocle" package (version 2.30.1; https://bioconductor.org/packages/release/bioc/html/monocle.html) to perform pseudotime sequence analysis on the RNA sequencing data of cell populations with high UCell scores, aiming to explore the dynamic changes in cellular development. In the selection of genes for trajectory analysis, we first calculated the dispersion of genes and filtered for genes with a mean expression level greater than 0.1 (mean_expression >= 0.1). Subsequently, we conducted a differential gene test (differentialGeneTest), selecting genes with a q-value less than 0.01 for ordering. Ultimately, we employed the "reduceDimension" function to dimensionally reduce the data to two dimensions, using the DDRTree method with max_components set to 2. We then ordered the cells using the "orderCells" function, generating a cellular developmental trajectory.

Cellular communication analysis

Cell–cell interaction analysis was performed using "CellChat" R package. CellChat[26] is a comprehensive knowledgebase of ligands, receptors, cofactors, and their interactions, accompanied by pathway annotations. Utilizing social network analysis tools, pattern recognition methods, and manifold learning techniques, CellChat identifies differentially expressed ligands and receptors in each cell type and clusters various communication patterns across different cell populations and pathways. Through these analyses, the specific signaling roles played by each cell population can be identified, and novel functional intercellular communication mechanisms for certain cell types can be discovered.

Single-cell regulatory network inference and clustering (SCENIC) analysis

Single-Cell Regulatory Network Inference and Clustering (SCENIC)[27], is a computational approach used in genomics research to analyze single-cell gene expression data and infer gene regulatory networks. It combines two main techniques: gene expression factorization and cis-regulatory motif analysis. The former assesses the activity of transcription factors (TFs) in individual cells, while the latter identifies TF binding sites within gene promoter regions. By integrating these techniques, SCENIC can deduce the regulatory networks that govern gene expression patterns.

High-dimensional weighted correlation network analysis

The high-dimensional weighted correlation network analysis (hdWGCNA) is a novel algorithm that provides a highly modular approach and can construct co-expression networks across multiscale cellular and spatial hierarchies, which is more suitable for scRNA-seq data[28]. This study utilizes the R package "hdWGCNA" and follows the official workflow (https://smorabit.github.io/hdWGCNA/articles/basic_tutorial.html) to analyze gene modules and their functional roles in cells with high UCell scores. Initially, we identified cell populations with high UCell scores from the Seurat object. We then normalized gene expression data using the "sctransform" function. Pearson correlation coefficients were calculated to detect co-expression relationships between gene pairs. The optimal soft threshold for constructing a scale-free network was determined using the "pickSoftThreshold" function. Based on this threshold, we generated an adjacency matrix and a topological overlap matrix (TOM). Co-expression modules were identified from the TOM using the "dynamicTreeCut" algorithm, with each module assigned a distinct color. The connectivity of hub genes within each module was assessed using the kME (gene-based connectivity) values calculated by the "ModuleConnectivity" function. Genes with higher kME values were considered more influential within their respective modules. Furthermore, the "GetMEs" function was used to compute the module eigengenes (ME), summarizing the gene expression profiles within each co-expression module. Finally, we extracted the top 500 module-associated genes from each module based on kME values using the "GetHubGenes" function. These genes were respectively intersected with MRGs and FRGs to identify key genes for this study.

Functional enrichment and PPI analysis

The Gene Ontology (GO) analysis, a widely used method for large-scale functional enrichment analysis, encompasses three main categories: biological processes (BP), molecular functions (MF), and cellular components (CC)[29]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database resource for understanding high-level functions and biological systems from large-scale molecular datasets generated using high-throughput experimental technologies[30]. To identify the biological functions and signaling pathways of genes, we used the "clusterProfier" package for GO and KEGG pathway enrichment analysis. FDR < 0.05 were considered to be statistically significant. A bar chart is utilized to display the top 8 results for GO analysis. Meanwhile, the top 10 results from the KEGG analysis are represented in a bubble plot. A PPI network was established using the Search Tool for the Retrieval of Interacting Genes (STRING) (version 10.0) (http://string-db.org) online search tool[31], with interactions scoring > 0.4 considered statistically significant. The PPI network was visualized using Cytoscape (version 3.7.2) software. Additionally, the CytoHubba plugin within Cytoscape was used to analyze the hub genes of the PPI network.

Selection of potential diagnostic genes

We employed two established machine learning algorithms, Least Absolute Shrinkage and Selection Operator (LASSO) regression and Random Forest (RF), to identify diagnostic genes for SLE from hub genes. Initially, we established a LASSO regression model using the "glmnet" package, setting the "family" parameter to "binomial" and selecting the optimal lambda value through "lambda.1min". We then visualized the LASSO coefficient profiles and determined the best lambda value for the minimum standard error (1-SE). Subsequently, we applied the Random Forest algorithm with the "randomForest" package to classify significant genes, using an ensemble of decision trees to identify the most important variables. This process allowed us to filter and identify shared disease signature genes. We constructed a random forest model with 100 trees on the discovery cohort and optimized the number of trees through cross-validation. Genes were ranked by importance, and the top 10 significant genes were plotted. The intersection of genes identified by both algorithms was visualized using a Venn diagram.

Construction and validation of the diagnostic model

We utilized logistic regression to construct a diagnostic model based on the diagnostic biomarkers previously identified for SLE patients. Nomogram was drawn by the "rms" package to elevate the operability and practicability of the diagnostic model. The discriminative performance of the model was visually demonstrated using clinical impact curves, calibration curves, and decision curve analysis (DCA). Finally, the model's generalization efficacy was meticulously evaluated using a suite of performance metrics, including AUC, precision, recall, F1 score, specificity and accuracy.

Immune infiltration analysis

CIBERSORT, an universal deconvolution algorithm, was employed to examine the immune cell subset proportions using RNA expression profiles[32]. Herein, we obtained a matrix of 22 immune cell subsets in GSE72326 via the R package "CIBERSORT". A bar plot displayed the percent of each immune cell and compare infiltrating levels between SLE patients and controls using the "ggplot2 (v3.3.0)" package. We also created correlation heatmaps to reflect the relationships between immune cells and diagnostic biomarkers. Subsequently, we compared the scores of different immune cells between SLE and control samples to identify those with varying infiltration levels across the samples.

Verification of the diagnostic genes using qPCR

Peripheral blood of 5 SLE patients and 5 healthy people was collected from Fujian Provincial Hospital. The diagnosis of this disease followed the 1997 revised American College of Rheumatology (ACR) classification criteria[33]. All protocols were approved by the ethics committee of the Fujian Provincial Hospital (No.K2024-03-016) and all procedures followed the tenets of the Declaration of Helsinki, and written informed consents were obtained from all participants. All methods were performed in accordance with the relevant guidelines and regulations.

The potential diagnostic genes were subjected to qPCR analysis to verify the reliability of the bioinformatics findings. Total RNA was extracted from the cell lines using TRI REAGENT BD(MRCGENE), following the manufacturer's instructions. And then total RNA was reverse transcribed to complementary DNA (cDNA) by using the Gene Amp PCR System 9700 (Applied Biosystems). qPCR was performed using QuantStudio5 Real-time PCR System (Applied Biosystems). β-actin was used as an internal control. The relative levels of hub genes were estimated using the $2^{-\triangle\triangle Ct}$ method. The primer sequences of hub genes were listed in Supplementary Table 2.

Drug screening and Docking

We queried the CTD database[34] (http://ctdbase.org/) for small molecule drugs associated with SLE and identified those interacting with each diagnostic biomarker. By intersecting all query results, we further refined the list of small molecule drugs that could potentially influence the aforementioned diagnostic biomarkers and serve as therapeutic strategies for SLE. To elucidate the binding mechanisms of these drugs with the diagnostic biomarkers, we obtained the three-dimensional crystal structures of the biomarkers from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) database (https://rcsb.org/) and the crystal structures of the small molecule drugs from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/). Molecular docking was performed using the Molecular Operating Environment (MOE) software. Then we selected the most promising small molecules for each biomarker based on their docking affinity scores.

Statistical analysis

R software 4.1.2 was performed for statistical analyses and visualization. To assess the differences between two continuous variables, we utilized the independent Student's t-test for those with a normal distribution and the Mann-Whitney U test for non-parametric comparisons. Forest plots for both univariate and multivariate analyses were generated using generalized linear models (glm). Unless otherwise specified, the correlation coefficients among different molecules in this study were calculated using Spearman's rank correlation analysis. All statistical P-values were two-tailed. And p-value < 0.05 was considered a significant difference.
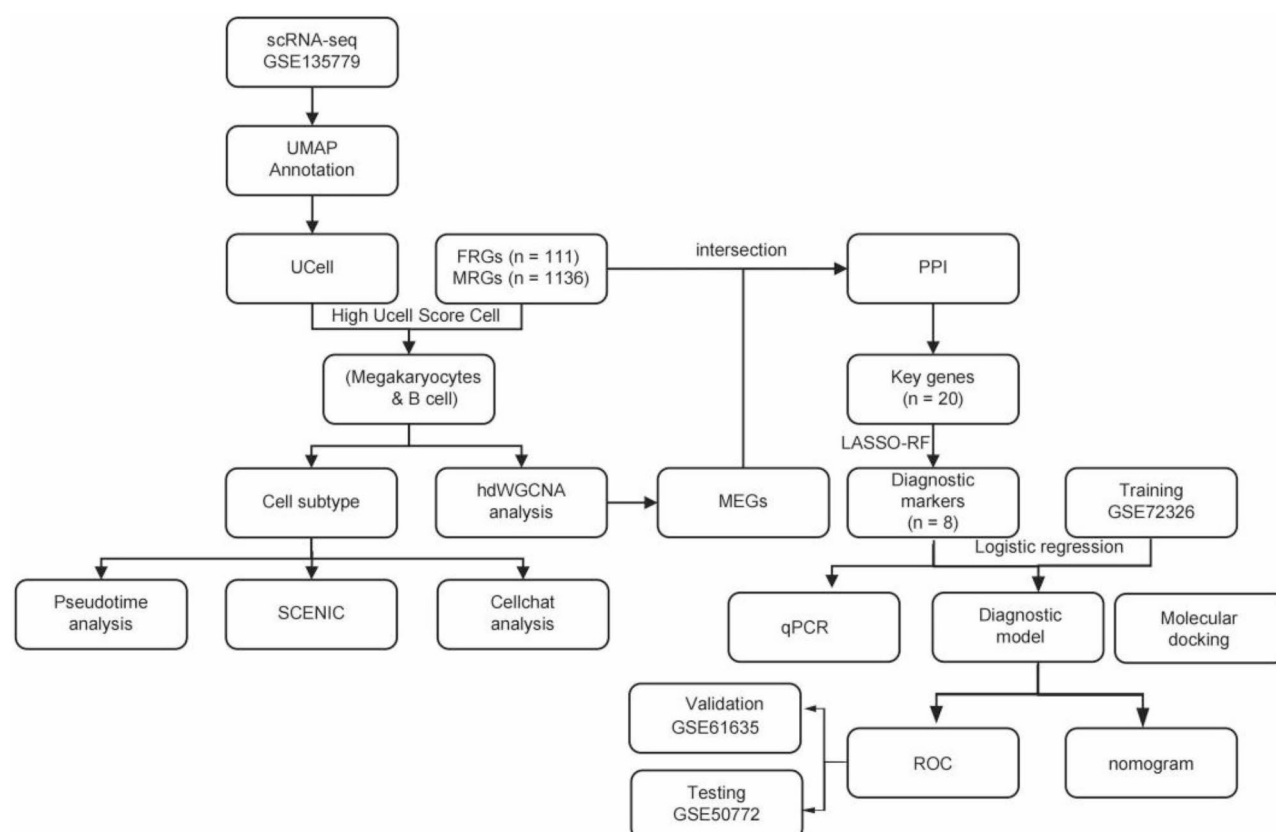
Results

Workflow

The entire working processes are shown in Fig. 1.

Single-cell heterogeneity

As described in the method, we obtained 19 cell clusters in the SLE single-cell dataset using UMAP clustering (Fig. 2A). These clusters were significant as they provide a detailed map of the cellular landscape within SLE, revealing the heterogeneity of immune cell populations that contribute to the disease's complexity. Through manual annotation, we identified 5 cell types among the 19 clusters (CD4 memory T cells, B cells, CD8 effector T cells, Dendritic cells, and Megakaryocytes) (Fig. 2C). The identification of these specific cell types was crucial for understanding their distinct roles and contributions to SLE pathology. We analyzed the expression of single-cell subpopulation marker genes across different cell clusters and presented the results in a bubble plot (Fig. 2B). Additionally, through differential expression analysis, we identified DEGs between different cell types ($n = 1409$, see Supplementary Table 3), and displayed the expression of the TOP20 DEGs across various cell types in a heatmap (Fig. 2D). In addition, we performed the GO/KEGG analysis of all DEGs, which revealed that several pathways play pivotal roles in the pathogenesis of SLE (Fig. 2E, F). These pathways encompass the activation of immune cells, the initiation of inflammatory responses, the disruption of self-tolerance, the escalation of oxidative stress, and the modulation of apoptosis. For further details, refer to Supplementary Table 4.

Phenotype-related gene score analysis in single-cell data

Utilizing the "UCell" R package, we computed scores for organelle function-related genes across different cell types and visualized the results using UMAP plots. Notably, FRGs scores were higher in Megakaryocytes and CD4 memory T cells, while lower in B cells and CD8 effector T cells (Supplementary Fig. 1A). While MRGs scores were elevated in CD4 memory T cells and lower in CD8 effector T cells and Dendritic cells (Supplementary

Fig. 1. Entire working processes of the study.
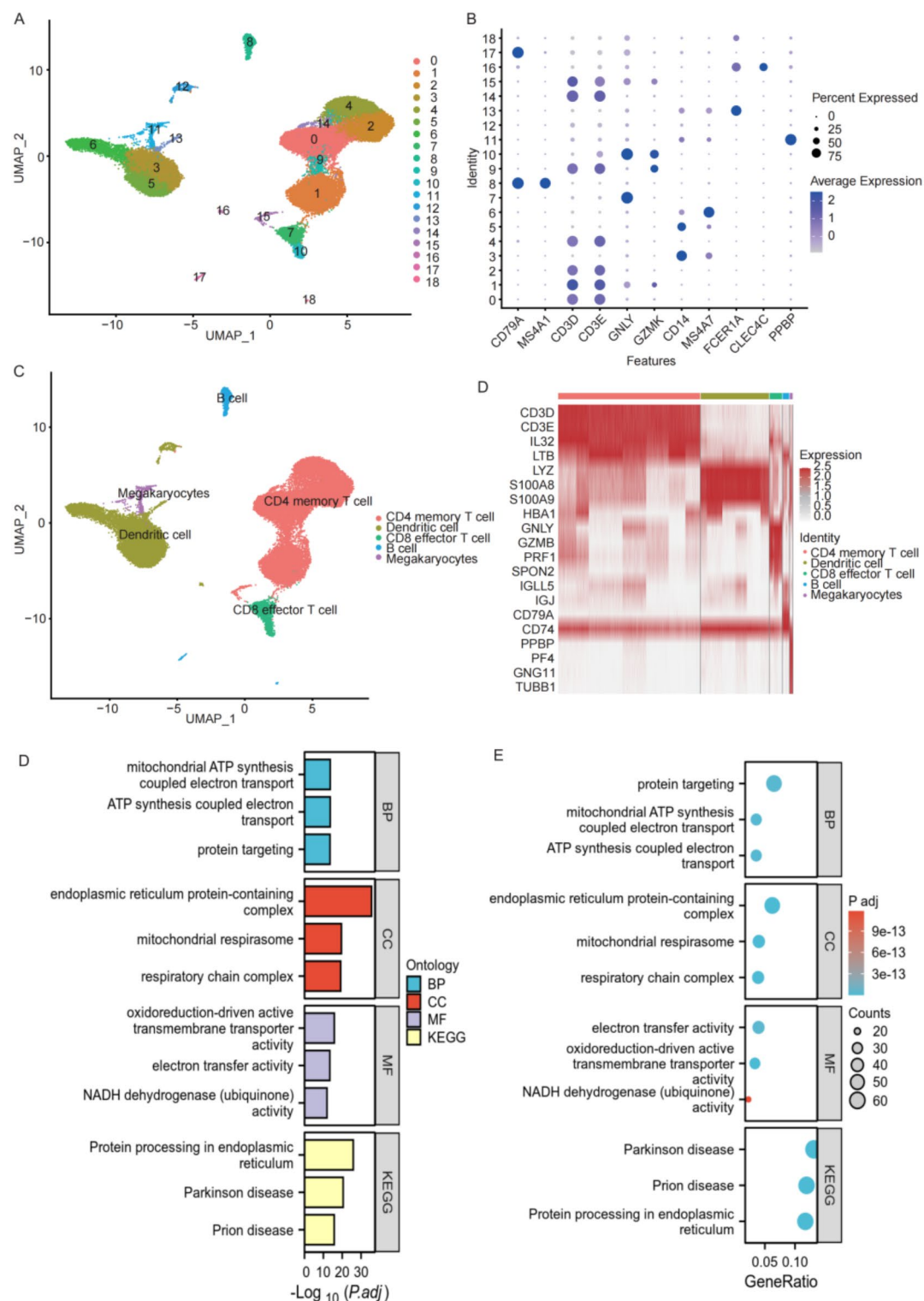
Fig. 2. Single-cell heterogeneity. (**A**) UMAP plot visualized the single-cell data clustering outcomes. (**B**) Bubble plot of marker gene expression across different cell clusters. (**C**) Annotation results of cell populations. (**D**) Heatmap of differential gene expression in single-cell transcriptomes (displaying the top 20, with red indicating upregulation and blue indicating downregulation). (**E, F**) GO and KEGG enrichment analysis results for all DEGs.

Fig. 1C). Further statistical analysis using the Wilcoxon test revealed that FRGs scores were significantly highest in Megakaryocytes (p-value < 0.001) (Supplementary Fig. 1B), and MRGs scores were most prominent in B cells (p-value < 0.001) (Supplementary Fig. 1D).

Analysis of novel subtypes with high UCell scores

To uncover novel cell subtypes with high UCell scores, we performed cluster analyses on Megakaryoctes and B cell clusters. Within the Megakaryocytes cluster, we identified four distinct Megakaryoctes subtypes (Fig. 3A). For these subtypes, we employed the "FindAllMarkers" function to determine marker genes among them (n = 1327, Supplementary Table 3), and annotated them based on the genes with significant differential expression (Fig. 3B), ultimately characterizing S100A8 Megakaryocytes, PF4 Megakaryocytes, IL32 Megakaryocytes, and FCGR3A Megakaryocytes (Fig. 3C). The expression patterns of these marker genes across the subtypes were visualized using a heatmap (Fig. 3D). In the B cell cluster, we identified five B cell subtypes (Fig. 3E). Similarly, we calculated marker genes for these subtypes (n = 1051, Supplementary Table 3) and annotated them based on differentially expressed genes (Fig. 3F), resulting in the identification of ANXA1 B cells, TCL1A B cells, CRIP1 B cells, IGJ B cells, and KLRB1 B cells (Fig. 3G). The heatmap was also used to illustrate the expression of marker genes across these subtypes (Fig. 3H).

Functional enrichment analysis

Based on the results of the aforementioned analysis, we conducted GO and KEGG enrichment analyses for DEGs between Megakaryocytes subtypes and B cell subtypes. More details are presented in Supplementary Tables 5 and 6, respectively. The GO analysis revealed that DEGs among Megakaryocytes subtypes are are linked to processes like cytoplasmic translation and cell adhesion regulation, and cellular components including ribosomes and MHC complexes, along with molecular functions involving MHC binding (Supplementary Fig. 2A). For B cell subpopulations, DEGs are associated with antigen presentation and T cell activation regulation, with connections to components like the endoplasmic reticulum and MHC complexes, and molecular functions such as peptide antigen binding (Supplementary Fig. 2B). The KEGG analysis indicated that DEGs among Megakaryocytes subtypes are related to pathways such as Coronavirus disease - COVID-19, Ribosome, and Leishmaniasis (Supplementary Fig. 2C), while those among B cell subpopulations are associated with pathways like Protein processing in endoplasmic reticulum, Parkinson's disease, and Prion disease (Supplementary Fig. 2D).

Monocle pseudotime analysis of high UCell score subtypes

Utilizing Monocle's pseudotime analysis, we delineated the developmental trajectories of four Megakaryocytes subtypes and visualized through a differentiation trajectory plot (Fig. 4A) and a timeline plot (Fig. 4B), suggesting a sequential development from FCGR3A to S100A8, IL32, and PF4 Megakaryocytes. To investigate the expression changes of marker genes among the four subtypes during their developmental trajectories, we displayed the expression patterns of these markers using a heatmap (Fig. 4C). Our findings revealed that genes such as FCGR3A, MS4A7, AIF1, IL8, RP11-290F20.3, LST1, S100A12, LYZ, S100A9, and S100A8 were downregulated during the developmental stages of Megakaryocytes subtypes. Conversely, genes like PF4, TUBB1, RGS18, PPBP, and TSC22D1 were upregulated. Additionally, the expression of genes CD3E, IL32, PTPRCAP, CD69, and LTB initially increased and then decreased.

Likewise, we traced the developmental paths of five B cell subtypes, as depicted in the differentiation trajectory plot (Fig. 4D) and timeline plot (Fig. 4E), which revealed a trajectory from TCL1A, CRIP1, KLRB1, ANXA1, to IGJ B cells. The heatmap (Fig. 4F) revealed that genes such as CXCR4, TCL1A, and BANK1 exhibited downregulated expression during the developmental phases of B cell subtypes. In contrast, genes including IGJ, IGLL5, MZB1, and HSP90B1 showed upregulated expression. The expression of genes like PRSS57, CRIP1, and IL32 initially increased and subsequently decreased.

Cellular communication analysis

Cell communication analysis revealed the extent of interactions between different subtypes of Megakaryocytes and B cells, respectively. By analyzing and comparing the interactions, we found tight interconnectivity within Megakaryocyte subtypes (Supplementary Fig. 3A, B) and among B cell subtypes (Supplementary Fig. 3D, E). Notably, we identified a markedly enhanced APP_CD74 receptor-ligand interaction in PF4 and S100A8 Megakaryocytes (Supplementary Fig. 3C). This interaction could be pivotal in the context of SLE, potentially influencing the pro-inflammatory cytokine profile and contributing to the aberrant immune responses characteristic of the disease. The APP_CD74 interaction might also be implicated in the activation and survival of Megakaryocytes, which are known to be dysregulated in SLE.

Similarly, a robust MIF_CD74 + CXCR4 interaction was observed in IGJ and CRIP1 B cells (Supplementary Fig. 3F). This interaction is biologically significant as it could modulate the migration and activation of B cells, which are central to the production of autoantibodies in SLE. The MIF_CD74 + CXCR4 interaction might also be involved in the pathogenesis of SLE by influencing the trafficking and positioning of B cells within the inflammatory milieu. These findings underscore the importance of specific cell-cell interactions in the pathophysiology of SLE.

SCENIC analyses

We first conducted SCENIC analysis for four Megakaryocyte subtypes and five B cell subpopulations, visualizing the AUC values of transcription factors (TFs) that showed statistical differences across various cell subtypes in a heatmap (Supplementary Fig. 4A, B). We then visualized the UMAP for two TFs that showed statistical differences within the Megakaryocyte subpopulations (Supplementary Fig. 4C). Our findings indicated that FOSB_extended (62 g) has a stronger effect on S100A8 Megakaryocytes, but a weaker effect on PF4 Megakaryocytes. While
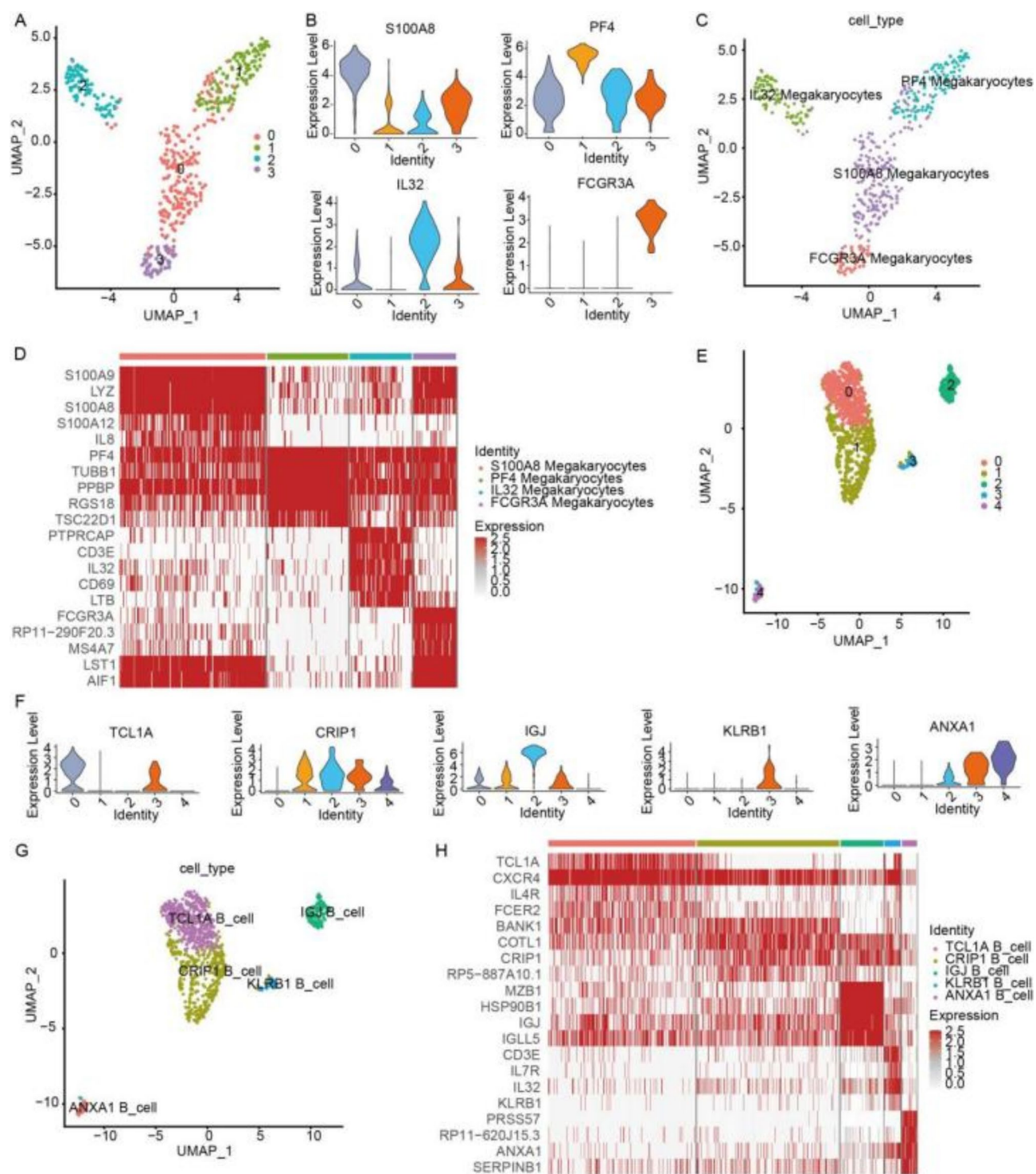
Fig. 3. Analysis of novel subtypes with high UCell scores. (**A**) UMAP plot of Megakaryocytes subtypes. (**B**) Violin plot of marker gene expression in Megakaryocytes subtypes. (**C**) UMAP plot of annotated Megakaryocytes subtypes. (**E**) UMAP plot of B cell subtypes. (**F**) Violin plot of marker gene expression in B cell subtypes. (**G**) UMAP plot of annotated B cell subtypes. (**D, H**) Heatmaps of marker gene expression in Megakaryocytes and B cell subtypes, respectively, with red indicating upregulation and blue indicating downregulation.
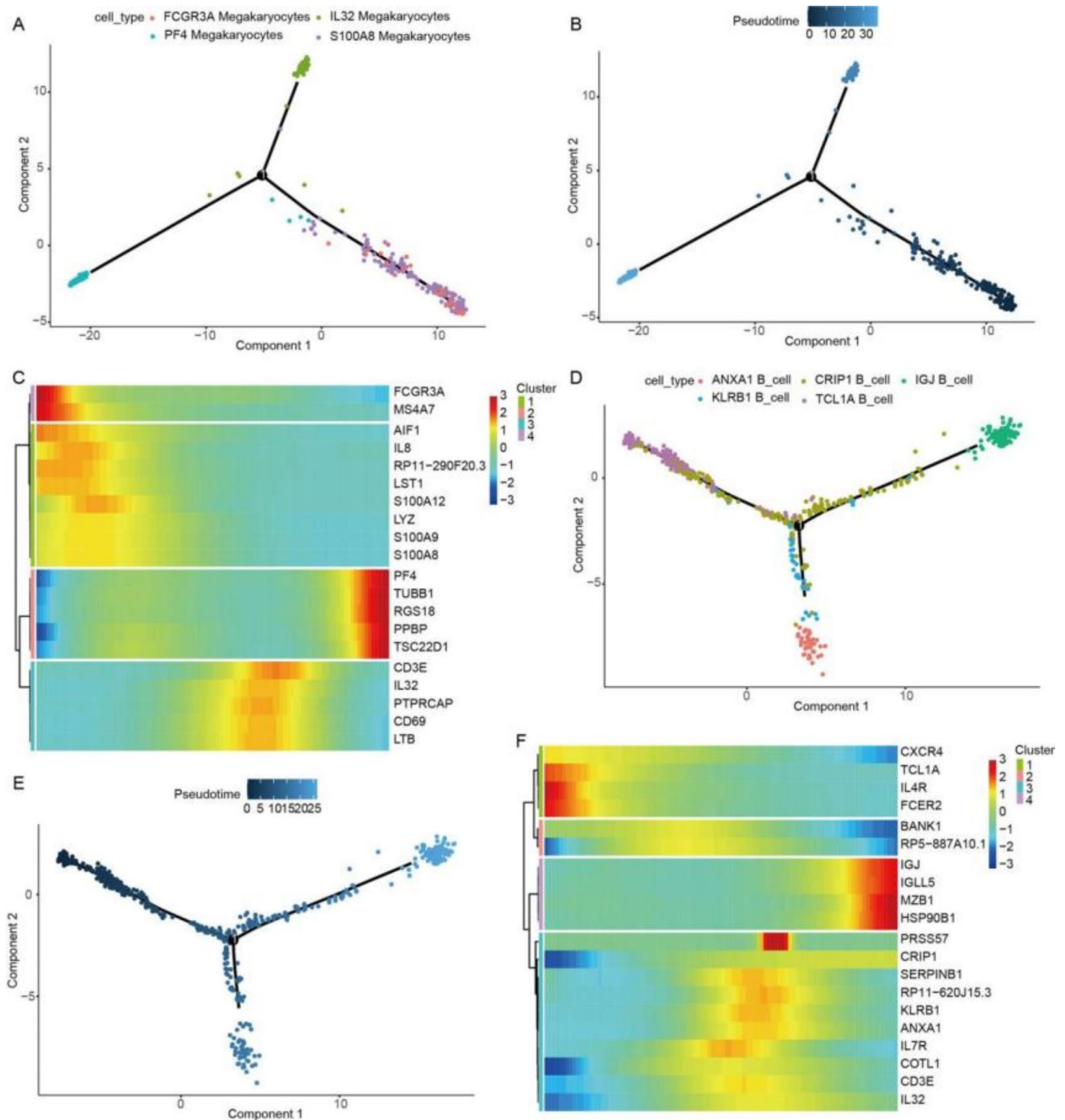
Fig. 4. Pseudotime analysis of high UCell score subtypes. (**A**, **B**) The differentiation trajectory plots of Megakaryocytes and B cell subtypes, respectively, with different colors indicating distinct cell subtypes. (**B**, **D**) The developmental timelines for Megakaryocytes and B cell subtypes, where color transitions from dark to light signify the order of pseudo-time progression. (**C**, **F**) The heatmaps depict the pseudotime expression dynamics of marker genes specific to either four Megakaryocytes subtypes (**C**) or five B cell subtypes (**F**), with upregulated expression indicated in red and downregulated expression in blue.

JUND_extended (95 g) exerts stronger regulation on S100A8 Megakaryocytes and relatively weaker regulation on FCGR3A Megakaryocytes. Similarly, for B cell subpopulations, UMAP visualization revealed that JUN_extended (325 g) has a stronger regulatory effect on TCL1A B cells and a relatively weaker effect on ANXA1 B cells; SPIB_extended (62 g) exerts stronger regulation on CRIP1 B cells and relatively weaker regulation on ANXA1 B cells (Supplementary Fig. 4D).

High-dimensional weighted correlation network analysis

To investigate gene modules, we performed high-dimensional weighted gene co-expression network analysis (hdWGCNA). The analysis of Megakaryocytes identified five modules with an optimal soft threshold of 11 (Supplementary Fig. 5A). The WGCNA dendrogram visually depicted the hierarchical clustering of genes based on their expression patterns (Supplementary Fig. 5B). From each module, we selected the top 500 genes as signature genes and illustrated their expression patterns with a UMAP plot (Supplementary Fig. 5C), with details listed in Supplementary Table 6. Similarly, for B cells, an optimal soft threshold of 7 yielded seven modules (Supplementary Fig. 6A), with hierarchical clustering visualized in a dendrogram (Supplementary Fig. 6B). The top 500 signature genes from each modules were similarly plotted on a UMAP (Supplementary Fig. 6C), with gene lists provided in Supplementary Table 7.

Selection of potential diagnostic genes

Initially, we performed an intersection of the module genes derived from the hdWGCNA analysis of Megakaryocytes with the FRGs, yielding a set of 22 genes (Supplementary Table 8) (Supplementary Fig. 7A). Similarly, for B cells, the intersection of module genes with MRGs yielded 136 genes (Supplementary Table 8) (Supplementary Fig. 7B). We then constructed PPI networks for these 22 and 136 genes, respectively. The interactions were imported into Cytoscape software, and the CytoHubba plugin was employed to identify the top 10 genes as hub genes using the MCC algorithm. Among the 22 genes, we identified TLN1, MYL6, FLNA, PRDX1, ACTG1, ACTB, RAC1, IQGAP1, MYH9, and CAPZB as hub genes (Supplementary Fig. 7C), and among the 136 genes, we identified NDUFS5, NDUFS6, NDUFB3, NDUFA6, NDUFB7, NDUFA13, NDUFB11, NDUFA1, NDUFB6, and NDUFA8 (Supplementary Fig. 7D). By merging these two gene groups, we determined a total of 20 key genes for this study (Supplementary Fig. 7E).

To identify signature genes for SLE and assess their diagnostic potential, we employed LASSO regression method based on the 20 hub genes. As the value of λ increased, the number of feature decreased while the absolute values of the coefficients increased (Supplementary Fig. 7F). After feature selection, we obtained two models (the optimal model and the simplest model) (Supplementary Fig. 7G). We selected the optimal model and identified 11 genes—ACTB, ACTG1, CAPZB, FLNA, IQGAP1, MYH9, MYL6, NDUFA1, NDUFA6, NDUFB3, PRDX1—as potential signature genes for SLE.

To enhance the accuracy of signature genes, we employed the RF algorithm to filter feature genes, selecting the top 10 genes ranked by importance (ACTB, FLNA, NDUFS6, MYL6, CAPZB, NDUFB3, PRDX1, ACTG1, NDUFA1, NDUFA8) as potential diagnostic genes for SLE (Supplementary Fig. 7H, I). Finally, by intersecting the potential feature genes identified by both the LASSO and RF algorithms, we obtained eight genes—ACTB, ACTG1, CAPZB, FLNA, MYL6, NDUFA1, NDUFB3, PRDX1—as diagnostic biomarkers for SLE (Supplementary Fig. 7J).

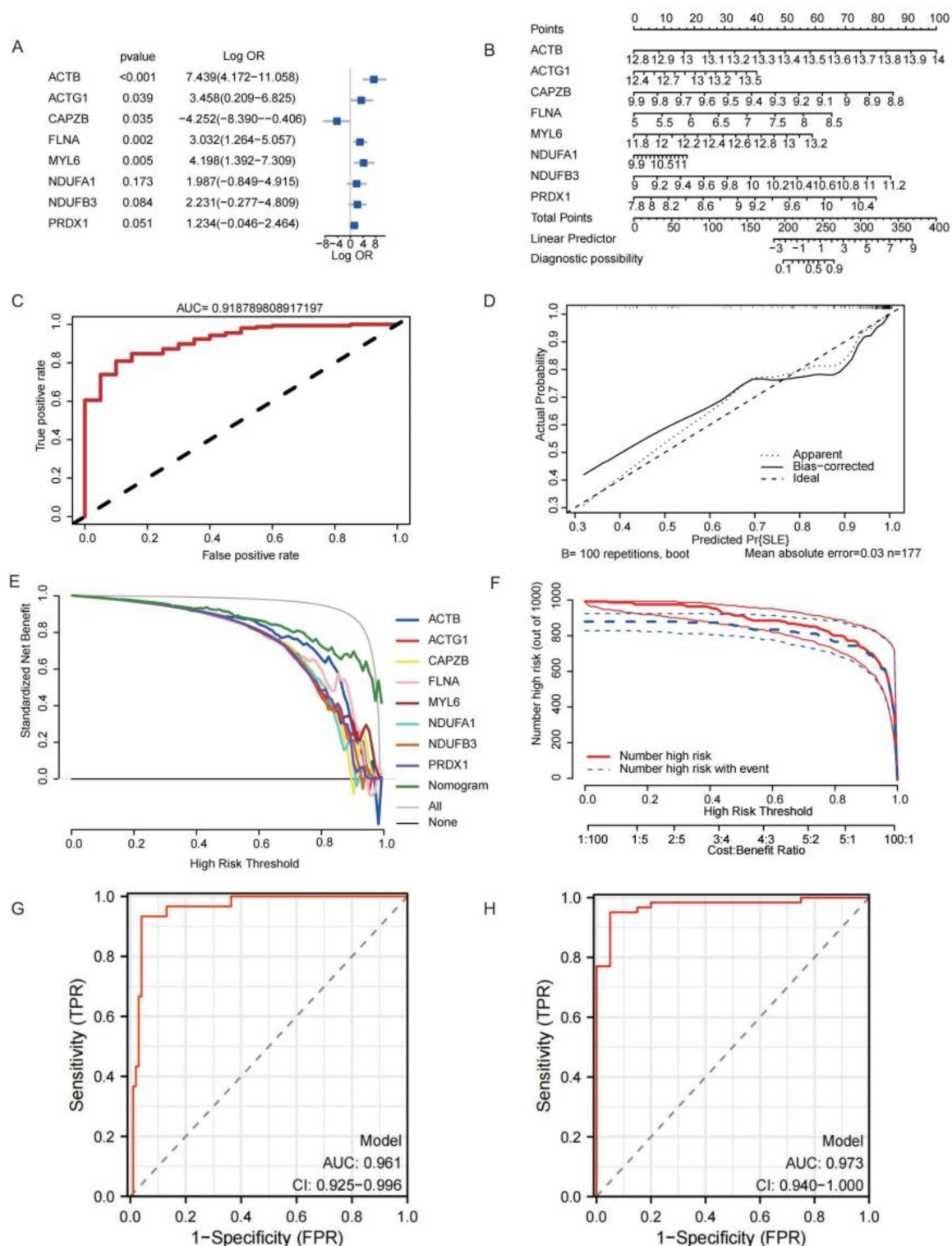### Construction and validation of the diagnostic model

We established a diagnostic model using logistic regression and presented it in a nomogram (Fig. 5A, B). This model demonstrated a high AUC value of 0.919 in training cohort, with the calibration plot showing a bias-corrected line nearly overlapping the ideal diagonal, indicating excellent predictive performance (Fig. 5C, D). Moreover, the DCA curve showed that the nomogram model had certain predictive performance for SLE compared to models based on single biomarkers (Fig. 5E). To more vividly illustrate the clinical effectiveness, we further plotted the clinical impact curve. The close proximity of the high-risk population curve to the high SLE risk event curve indicates robust predictive power (Fig. 5F). The AUC value for the validation set is 0.961 (Fig. 5G), while for the test set, it is 0.973 (Fig. 5H). In the training set, the model achieved an AUC of 0.919, a Precision of 0.985, a Recall of 0.809, an F1-score of 0.889, a Specificity of 0.9, and an Accuracy of 0.82. For the validation set (GSE61635), the metrics were an AUC of 0.961, a Precision of 0.875, a Recall of 0.933, an F1-score of 0.903, a Specificity of 0.96, and an Accuracy of 0.959. In the test set (GSE50772), the model demonstrated an AUC of 0.973, a Precision of 0.983, a Recall of 0.951, an F1-score of 0.967, a Specificity of 0.95, and an Accuracy of 0.951 (Supplementary Table 9). These results suggest that the model achieved excellent predictive performance.

Immune infiltration analysis

Using CIBERSORT, we calculated immune cell infiltration percentages in SLE cohorts, identifying 18 cell types after excluding zero-score cells (Supplementary Fig. 8A). We then observed significant differences in the infiltration scores of T cells CD4 memory resting, NK cells resting, monocytes, and activated dendritic cells between the SLE and control groups (Supplementary Fig. 8B). We also analyzed the correlations between the eight diagnostic biomarkers and immune cell infiltration. In the control group, the strongest positive correlation was observed between NDUFA1 and T cells CD8 ($r = 0.81$, $p = 1.38e-5$), and the strongest negative correlation was between NDUFA1 and Neutrophils ($r = -0.8$, $p = 2.59e-5$) (Supplementary Fig. 9C). In SLE patients, MYL6 showed the highest positive correlation with Neutrophils ($r = 0.52$, $p = 3.12e-12$), while PRDX1 had the highest negative correlation with Neutrophils ($r = -0.65$, $p = 1.62e-20$) (Supplementary Fig. 8D).

Verification of the diagnostic genes using qPCR

The bioinformatics analysis revealed significant differences in the expression of potential diagnostic markers between the SLE group and the control group. To validate these findings, we performed qPCR to assess the relative expression of the identified hub genes. As depicted in Supplementary Fig. 9, the relative expression of PRDX1 was significantly higher ($p < 0.05$) in SLE patients compared to controls. Conversely, the relative expression of ACTG1, CAPZB, FLNA, MYL6, and NDUFA1 was significantly lower ($p < 0.05$) in SLE patients. These results substantiate the accuracy and reliability of our initial bioinformatics analyses. However, the expression levels of ACTB and NDUFB3 did not show statistical significance, possibly due to the limited sample size of the qPCR

Fig. 5. Construction and Validation of the diagnostic model. (**A**) Forest plot illustrates the logistic regression analysis for the eight diagnostic biomarkers. (**B**) Nomogram predicting the probability of SLE. (**C**) ROC curve in training cohort. (**D-F**) Calibration curve, DCA curve, and clinical impact curve are used to evaluate the model's performance. (**G**, **H**) ROC curve in validation cohort and testing cohort.
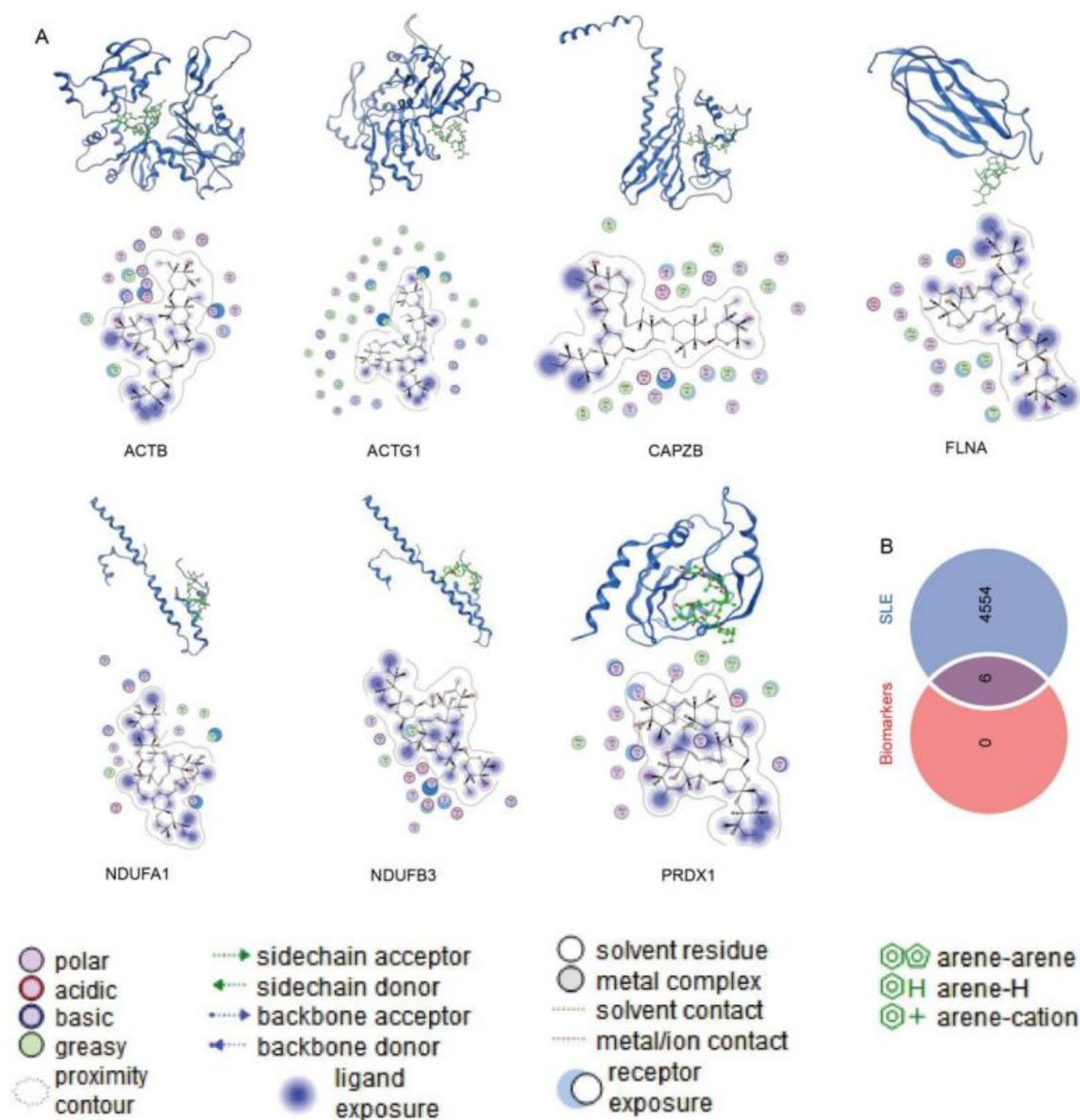
Fig. 6. Molecular docking results. (**A**) Schematic diagrams of the docking complex structures with Doxorubicin and biomarkers, and interaction diagrams between the ligand and receptors. (**B**) Venn diagram illustrating the six small molecule drugs that act on both each biomarker and SLE.

analysis. Future studies with a larger sample size could offer further insights into the role of these markers in SLE pathology.

Small molecule agents screening and Docking analysis

In the CTD database, we identified 4,554 small molecule drugs related to SLE and also searched for those interacting with each diagnostic biomarker. We discovered six small molecule drugs that interact with all the diagnostic biomarkers and are also associated with SLE, with IDs D000082(Acetaminophen), C006780(bisphenol A), D004317(Doxorubicin), D007559(Ivermectin), D013749(Tetrachlorodibenzodioxin), D014635(Valproic Acid) (Fig. 6B). We speculate that these six small molecule drugs could potentially serve as therapeutic strategies for SLE by affecting the aforementioned biomarkers. Molecular docking using MOE revealed that among all the proteins, the small molecule Doxorubicin had the strongest affinity. In ACTB, Doxorubicin forms hydrogen

bonds with Lys214 and Lys335; in ACTG1, it forms hydrogen bonds and aromatic interactions with Phe374 in both chains A and B; in CAPZB, FLNA, and NDUFA1, Doxorubicin does not form polar bonds but exhibits strong hydrophobic interactions; in NDUFB3, it forms a hydrogen bond with Arg93; and in PRDX1, it forms hydrogen bonds with Lys34 and Asn101, and an aromatic interaction with Tyr (Fig. 6A). Notably, the interactions of Doxorubicin with ACTG1 and CAPZB exhibited the highest affinity, as indicated by the S values of -8.6954 and − 8.4585, respectively. This suggests that the interaction mechanisms of these two proteins with Doxorubicin may represent promising therapeutic targets for SLE treatment development.

Discussion

Systemic lupus erythematosus (SLE) is a prototypical autoimmune disease that can lead to chronic inflammation and damage in various organs and tissues[35]. Patients with SLE often encounter mitochondrial dysfunction, which plays a critical role in energy metabolism and is frequently compromised by the excessive production of reactive oxygen species and oxidative stress[36]. This dysfunction is further exacerbated by iron deficiency, a key factor for maintaining mitochondrial health, thereby intensifying the associated issues. Furthermore, ferroptosis—an iron-dependent form of cell death—emerges as a potential link between mitochondrial dysfunction and iron deficiency, suggesting its role as a significant mechanism in the pathogenesis of SLE[37]. Our research integrates a multi-omics analysis to pinpoint the genetic and cell-type-specific mechanisms that drive the pathogenesis of SLE, with the goal of enhancing prognosis and guiding the development of more effective treatment strategies.

Here, we used single-cell RNA sequencing technology to delve into the heterogeneity of immune cell subpopulations in patients with SLE. We successfully identified 19 distinct cell clusters and annotated five major cell types, including CD4 memory T cells, B cells, CD8 effector T cells, dendritic cells, and megakaryocytes. Through calculating mitochondrial-related genes (MRGs) and ferroptosis-related genes (FRGs) scores, we found FRGs were most highly expressed in megakaryocytes, while MRGs were most prominent in B cells. These findings align with those of Shigeru Iwata and colleagues, who reported that B cells drive immune dysregulation in SLE through the activation of the mTORC1 signaling pathway and the enhancement of glutaminolysis. This metabolic shift results in mitochondrial dysfunction and an increased generation of ROS, which subsequently promotes the differentiation of plasma cells and the production of autoantibodies. Consequently, this cascade of events intensifies the immunological imbalance associated with SLE[38]. Ferroptosis is speculated to be an integral component in the vicious cycle of immune dysfunction, inflammation, and tissue damage in lupus[5,39]. Recently, it has been shown that ferroptosis regulated several fundamental cellular processes including cell proliferation and differentiation[40]. Song et al. found that ferroptosis may influence the differentiation process and function of megakaryocytes by affecting glutathione (GSH) levels, the accumulation of lipid reactive oxygen species (ROS), and the expression of genes associated with ferroptosis[41]. To the best of our knowledge, the current study is the first to explore the role of ferroptosis in the context of megakaryocytes and its association with immune dysregulation in SLE. This is a significant finding that requires further investigation in future research.

In our study, we applied the hdWGCNA, PPI network and machine learning approaches to identify eight diagnostic genes—ACTB, ACTG1, CAPZB, FLNA, MYL6, NDUFA1, NDUFB3, and PRDX1—that are significantly associated with SLE. Among the eight potential diagnostic markers, several have been identified with a robust correlation to the pathophysiology of SLE. For example, filamin A (FLNA), an actin-binding protein, exhibited increased expression in the amniotic fluid of SLE patients, which was significantly correlated with the occurrence of adverse pregnancy outcomes[42].The expression of PRDX1, an antioxidant gene, was found to be upregulated in sulforaphane-treated lupus-prone MRL/lpr mice, indicating that PRDX1 plays a protective role in mitigating the pathogenesis of SLE[43].

Other genes are involved in immune regulation. NDUFA1 and NDUFB3 are both components of the mitochondrial electron transport chain (ETC), which demonstrates heightened activity, particularly at Complex I, contributing to increased oxidative stress in SLE patients[44]. As for ACTB (Actin Beta)), ACTG1 (Actin Gamma 1), CAPZB (Capping Actin Protein Of Muscle Z-Line Subunit Beta), and MYL6 (Myosin Light Chain 6), they are all components that contribute to the structure and function of the actin and myosin cytoskeleton. A study has demonstrated that actin and myosin play a crucial role in the formation of ring-like actin networks within T cells, which are important for immune cell signaling[45]. Regardless, the specific effects of these genes on SLE necessitate further investigation in future studies.

Utilizing drug screening and molecular docking analysis, our research revealed that doxorubicin not only has a significant association with SLE but also exhibits the strongest interaction affinity with all the identified diagnostic biomarkers. Doxorubicin (DOX), a member of the anthracycline class of antibiotics, primarily targets mitochondria and serves as an effective chemotherapeutic agent for the treatment of various cancers[46]. Although the role of doxorubicin in SLE has not been studied in the literature, it has been employed in other autoimmune diseases. For instance, Khan et al. found that doxorubicin may serve as a potential treatment for NETosis dysregulation in rheumatoid arthritis (RA). The drug's capacity to inhibit NETosis without compromising neutrophil ROS production indicates its potential therapeutic value in RA[47,48]. Molecular docking results suggest that doxorubicin may be a target agent for SLE patients by targeting ACTG1 and CAPZB, but the specific mechanism requires further experimental verification.

Our results can serve as a theoretical foundation for pathophysiological mechanisms in SLE. But there are some limitations in our research. Firstly, our analysis focused only on specific cell types (B cells and megakaryocytes), which may have overlooked global features to some extent. However, multiple cells are involved in the immune dysregulation process of SLE, such as neutrophils. Secondly, during the scRNA-seq analysis for subpopulation annotation, there may have been instances where certain cells were not accurately annotated, which could potentially affect the comprehensiveness of our assessment of cellular heterogeneity. Thirdly, although we constructed a diagnostic model and identified relationship of immune infiltration based on diagnostic genes, further validation by prospective studies with a large sample number is needed before clinical

application. Lastly, our study lacks in vivo or in vitro functional validation for the diagnostic genes identified, such as CRISPR gene editing or siRNA knockdown. This limitation should be addressed in future research to confirm the biological relevance and impact of the identified genes.

Conclusion

This study has successfully identified eight diagnostic biomarkers for SLE using single-cell RNA sequencing and bioinformatics analysis, revealing the intricate roles of mitochondria and ferroptosis in the pathogenesis of SLE. We have developed a predictive diagnostic model with high accuracy. Furthermore, our findings suggest that doxorubicin may play a role in modulating the pathogenesis and progression of SLE by affecting mitochondrial function and iron metabolism. These insights emphasize the significance of mitochondrial dysfunction and the ferroptotic process in SLE, offering new avenues for future research and clinical interventions.

Data availability

The datasets generated and analysed during the current study are available from the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) database. And the data used to support the findings of this study are available from the corresponding author upon request.

References

1. Tsokos, G. C. Systemic lupus erythematosus. *N. Engl. J. Med.* **365** (22), 2110–2121. https://doi.org/10.1056/NEJMra1100359 (2011).
2. Tian, J., Zhang, D., Yao, X., Huang, Y. & Lu, Q. Global epidemiology of systemic lupus erythematosus: A comprehensive systematic analysis and modelling study. *Ann. Rheum. Dis.* **82** (3), 351–356. https://doi.org/10.1136/ard-2022-223035 (2023).
3. Zeng, L. et al. Advances in research on immunocyte iron metabolism, ferroptosis, and their regulatory roles in autoimmune and autoinflammatory diseases. *Cell Death Dis.* **15** (7), 481. https://doi.org/10.1038/s41419-024-06807-2 (2024).
4. Zheng, Y., Yan, F., He, S. & Luo, L. Targeting ferroptosis in autoimmune diseases: Mechanisms and therapeutic prospects. *Autoimmun. Rev.* **23** (11), 103640. https://doi.org/10.1016/j.autrev.2024.103640 (2024).
5. Chen, Q. et al. The potential role of ferroptosis in systemic lupus erythematosus. *Front. Immunol.* **13**, 855622. https://doi.org/10.3389/fimmu.2022.855622 (2022).
6. Xu, Y., Shen, J. & Ran, Z. Emerging views of mitophagy in immunity and autoimmune diseases. *Autophagy* **16** (1), 3–17. https://doi.org/10.1080/15548627.2019.1603547 (2020).
7. Askeland, G. et al. Increased nuclear DNA damage precedes mitochondrial dysfunction in peripheral blood mononuclear cells from Huntington's disease patients. *Sci. Rep.* **8** (1), 9817. https://doi.org/10.1038/s41598-018-27985-y (2018).
8. Wei, Q., Podocyte in Various International Society of Nephrology/Renal Pathology Society Class Lupus Nephritis Patients. & et al Mitochondrial dynamics of. *DNA Cell Biol.* **42** (7), 358–363. https://doi.org/10.1089/dna.2022.0636 (2023).
9. Dixon, S. J. et al. Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell* **149** (5), 1060–1072. https://doi.org/10.1016/j.cell.2012.03.042 (2012).
10. Ta, N. et al. Mitochondrial outer membrane protein FUNDC2 promotes ferroptosis and contributes to doxorubicin-induced cardiomyopathy. *Proc. Natl. Acad. Sci. U. S. A.* **119** (36), e2117396119. https://doi.org/10.1073/pnas.2117396119 (2022).
11. Friedman, J. R. & Nunnari, J. Mitochondrial form and function. *Nature* **505** (7483), 335–343. https://doi.org/10.1038/nature12985 (2014).
12. Wójcik, P. et al. Oxidative stress and lipid mediators modulate immune cell functions in autoimmune diseases. *Int. J. Mol. Sci.* **22** (2), 723. https://doi.org/10.3390/ijms22020723 (2021).
13. Zhang, L. et al. Investigation of ferroptosis-associated molecular subtypes and immunological characteristics in lupus nephritis based on artificial neural network learning. *Arthritis Res. Therapy*. **26** (1), 126. https://doi.org/10.1186/s13075-024-03356-z (2024).
14. Papalexi, E. & Satija, R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* **18** (1), 35–45. https://doi.org/10.1038/nri.2017.76 (2018).
15. Yang, J. et al. Fibronectin-1 is a dominant mechanism for rheumatoid arthritis via the mediation of synovial fibroblasts activity. *Front. Cell. Dev. Biol.* **10**, 1010114. https://doi.org/10.3389/fcell.2022.1010114 (2022).
16. Wang, Y. et al. The shared biomarkers and pathways of systemic lupus erythematosus and metabolic syndrome analyzed by bioinformatics combining machine learning algorithm and single-cell sequencing analysis. *Front. Immunol.* **13**, 1015882. https://doi.org/10.3389/fimmu.2022.1015882 (2022).
17. Cui, Y. et al. Exploring the shared molecular mechanisms between systemic lupus erythematosus and primary Sjögren's syndrome based on integrated bioinformatics and single-cell RNA-seq analysis. *Front. Immunol.* **14**, 1212330. https://doi.org/10.3389/fimmu.2023.1212330 (2023).
18. Chiche, L. et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. *Arthritis Rheumatol.* **66** (6), 1583–1595. https://doi.org/10.1002/art.38628 (2014).
19. Carpintero, M. F. et al. Diagnosis and risk stratification in patients with anti-RNP autoimmunity. *Lupus*. **24** (10), 1057–1066 (2015). https://doi.org/10.1177/0961203315575586
20. Kennedy, W. P. et al. Association of the interferon signature metric with serological disease manifestations but not global activity scores in multiple cohorts of patients with SLE. *Lupus Sci. Med.* **2** (1), e000080. https://doi.org/10.1136/lupus-2014-000080 (2015).
21. Nehar-Belaid, D. et al. Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* **21** (9), 1094–1106. https://doi.org/10.1038/s41590-020-0743-0 (2020).
22. Rath, S. et al. MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Res.* **49** (D1), D1541–D1547. https://doi.org/10.1093/nar/gkaa1011 (2021).
23. Zhou, N. & Bao, J. FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database (Oxf.)* **2020**, baaa021 (2020). https://doi.org/10.1093/database/baaa021
24. Andreatta, M. & Carmona, S. J. UCell: Robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* **19**, 3796–3798. https://doi.org/10.1016/j.csbj.2021.06.043 (2021).
25. Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F. & Theis, F. J. Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods*. **13** (10), 845–848. https://doi.org/10.1038/nmeth.3971 (2016).
26. Jin, S. et al. Inference and analysis of cell-cell communication using cellchat. *Nat. Commun.* **12** (1), 1088. https://doi.org/10.1038/s41467-021-21246-9 (2021).
27. Aibar, S. et al. SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods*. **14** (11), 1083–1086. https://doi.org/10.1038/nmeth.4463 (2017).

28. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E. & Swarup, V. High dimensional co-expression networks enable discovery of transcriptomic drivers in complex biological systems. *BioRxiv* https://doi.org/10.1101/2022.09.22.509094 (2022).
29. Gene Ontology Consortium. Gene ontology consortium: Going forward. *Nucleic Acids Res.* D1049–D1056. https://doi.org/10.1093/nar/gku1179 (2015).
30. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51** (D1), D587–D592. https://doi.org/10.1093/nar/gkac963 (2023).
31. Szklarczyk, D. et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47** (D1), D607–D613. https://doi.org/10.1093/nar/gky1131 (2019).
32. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods.* **12** (5), 453–457. https://doi.org/10.1038/nmeth.3337 (2015).
33. Hochberg, M. C. Updating the American college of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* **40** (9), 1725. https://doi.org/10.1002/art.1780400928 (1997).
34. DDavis, A. P. et al. Comparative toxicogenomics database (CTD): Update 2023. *Nucleic Acids Res.* **51** (D1), D1257–D1262. https://doi.org/10.1093/nar/gkac833 (2023).
35. Dörner, T. & Furie, R. Novel paradigms in systemic lupus erythematosus. *Lancet* **393** (10188), 2344–2358. https://doi.org/10.1016/S0140-6736(19)30546-X (2019).
36. Doherty, E., Oaks, Z. & Perl, A. Increased mitochondrial electron transport chain activity at complex I is regulated by N-acetylcysteine in lymphocytes of patients with systemic lupus erythematosus. *Antioxid. Redox Signal.* **21** (1), 56–65. https://doi.org/10.1089/ars.2013.5702 (2014).
37. Wincup, C., Sawford, N. & Rahman, A. Pathological mechanisms of abnormal iron metabolism and mitochondrial dysfunction in systemic lupus erythematosus. *Expert Rev. Clin. Immunol.* **17** (9), 957–967. https://doi.org/10.1080/1744666X.2021.1953981 (2021).
38. Iwata, S., Hajime Sumikawa, M. & Tanaka, Y. B Cell activation via immunometabolism in systemic lupus erythematosus. *Front. Immunol.* **14**, 1155421. https://doi.org/10.3389/fimmu.2023.1155421 (2023).
39. Tang, D., Chen, X., Kang, R. & Kroemer, G. Ferroptosis: Molecular mechanisms and health implications. *Cell. Res.* **31** (2), 107–125. https://doi.org/10.1038/s41422-020-00441-1 (2021).
40. Wang, D., Xie, N., Gao, W., Kang, R. & Tang, D. The ferroptosis inducer Erastin promotes proliferation and differentiation in human peripheral blood mononuclear cells. *Biochem. Biophys. Res. Commun.* **503** (3), 1689–1695. https://doi.org/10.1016/j.bbrc.2018.07.100 (2018).
41. Song, B. et al. Inhibition of ferroptosis promotes megakaryocyte differentiation and platelet production. *J. Cell. Mol. Med.* **26** (12), 3582–3585. https://doi.org/10.1111/jcmm.17289 (2022).
42. Jeon, H. S. et al. Proteomic biomarkers in mid-trimester amniotic fluid associated with adverse pregnancy outcomes in patients with systemic lupus erythematosus. *PLoS ONE.* **15** (7), e0235838. https://doi.org/10.1371/journal.pone.0235838 (2020).
43. Du, P. et al. Sulforaphane ameliorates the severity of psoriasis and SLE by modulating effector cells and reducing oxidative stress. *Front. Pharmacol.* **13**, 805508. https://doi.org/10.3389/fphar.2022.805508 (2022).
44. Doherty, E., Oaks, Z. & Perl, A. Increased mitochondrial electron transport chain activity at complex I is regulated by N-acetylcysteine in lymphocytes of patients with systemic lupus erythematosus. *Antioxid. Redox Signal.* **21** (1), 56–65. https://doi.org/10.1089/ars.2013.5702 (2014).
45. Ni, Q. et al. A tug of war between filament treadmilling and myosin induced contractility generates actin rings. *ELife* **11**, e82658 (2022). https://doi.org/10.7554/eLife.82658
46. Wu, B. B., Leung, K. T. & Poon, E. N. Mitochondrial-targeted therapy for doxorubicin-induced cardiotoxicity. *Int. J. Mol. Sci.* **23** (3), 1912. https://doi.org/10.3390/ijms23031912 (2022).
47. Khan, M. A., D'Ovidio, A., Tran, H. & Palaniyar, N. Anthracyclines suppress both NADPH oxidase- dependent and -Independent NETosis in human neutrophils. *Cancers (Basel).* **11** (9), 1328. https://doi.org/10.3390/cancers11091328 (2019).
48. Khandpur, R. et al. NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis. *Sci. Transl. Med.* **5** (178), 178ra40. https://doi.org/10.1126/scitranslmed.300558 (2013).

Author contributions

Chen Zhihan, and Zhang li contributed to the study design and critical revision of the manuscript. Dai Yunfeng and Liu Jianwen carried out the study and drafted the manuscript. Dai Yunfeng, Liu Jianwen, Lai Yongxing, Gao Fei, He Lin, Zhang li, and Chen Zhihan analyzed the data. All authors read and approved the final manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Fujian Provincial Hospital (No.K2024-03-016). The patients/participants provided their written informed consent to participate in this study. All methods were performed in accordance with the relevant guidelines and regulations.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-93872-y.

Correspondence and requests for materials should be addressed to L.Z. or Z.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.