

# Identification of Disulfidptosis-Related Genes and Molecular Subgroups in Rheumatoid Arthritis for Diagnostic Model and Patient Stratification

Xinyi Liu<sup>1</sup>, Siyao Wang<sup>2</sup>, Xinru Du<sup>1</sup>, Yulu Wang<sup>3</sup>, Lingfei Mo<sup>1</sup>, Hanchao Li<sup>1</sup>, Zechao Qu<sup>4</sup>, Xiaohao Wang<sup>4</sup>, Jian Sun<sup>5</sup>, Yuanyuan Li<sup>1</sup>, Jing Wang<sup>1</sup>

<sup>1</sup>Department of Rheumatology and Immunology, First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, People's Republic of China; <sup>2</sup>Department of Gastroenterology, First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, People's Republic of China; <sup>3</sup>Xi'an Jiaotong University College of Medicine, Xi'an, People's Republic of China; <sup>4</sup>Department of Spine Surgery, Hong Hui Hospital, Xi'an Jiaotong University, Xi'an, People's Republic of China; <sup>5</sup>Institute of Endemic Diseases, School of Public Health & Key Laboratory of Trace Elements and Endemic Diseases, Xi'an Jiaotong University, Xi'an, People's Republic of China

Correspondence: Jing Wang; Yuanyuan Li, Department of Rheumatology and Immunology, First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, 710061, People's Republic of China, Email kidip@163.com; wudui220@163.com

**Introduction:** Cell death contributes to the pathogenesis of rheumatoid arthritis (RA) through various pathways. Disulfidptosis is a recently discovered novel form of cell death characterized by the abnormal accumulation of intracellular disulfide bonds. It remains unclear for the association between RA and disulfidptosis.

**Methods:** A comprehensive analysis of three GEO datasets was presented in this study. First, the analysis involved the use of weighted gene co-expression network analysis (WGCNA) and differential analysis and were employed to identify the key module genes related to RA and disulfidptosis-related genes. The machine learning algorithms were used to identify the hub genes. Second, a diagnostic model was constructed for RA based on the hub genes. The nomogram and receiver operating characteristic (ROC) curves were utilized to evaluate the diagnostic value of the model. Third, two RA subtypes were identified based on hub genes by using consensus clustering analysis. Then, the disease activity scores, clinical markers, and immune cells were compared between the two RA subgroups. Finally, the differential expression of hub genes was validated between healthy controls and RA patients by qPCR.

**Results:** Four hub genes (KLHL2, POLK, CLEC4D, NXT2) were identified. The expression of the four hub genes was verified to be significantly higher in RA patients compared with healthy controls. The superior diagnostic value of the model was validated, which demonstrated that the model outperforms each hub gene individually. Two subtypes of RA were determined. Patients in cluster A exhibited relatively lower levels of DAS28-CRP, DAS28-ESR, CDAI, SDAI, RF, CRP, and MMP3. In contrast, patients in cluster B had significantly higher levels of the above markers.

**Conclusion:** Four hub genes were identified to provide unique insights into the role of disulfidptosis in RA. Additionally, a promising diagnosis model and patient stratification were established based on the hub genes to assess the risk of RA onset and RA disease activity.

**Keywords:** rheumatoid arthritis, disulfidptosis, machine learning, bioinformatics, diagnostic model

## Introduction

Rheumatoid arthritis (RA) is a systemic autoimmune disease characterized by persistent synovitis and the destruction of joint cartilage and bone.<sup>1</sup> Additionally, it affects the function of various organs, such as the heart, lungs, and kidneys, leading to a decline in quality of life.<sup>2,3</sup> The prognosis of RA is influenced by early diagnosis and proper treatment.<sup>4,5</sup> The biological and targeted synthesis disease modifying anti-rheumatic drugs (DMARDs) have been demonstrated to yield superior outcomes for patients suffering RA compared with conventional DMARDs. However, approximately 30%–40% patients have no response to the current treatment. Additionally, no treatment can be demonstrated to fully cure the disease.<sup>6</sup> In accordance with the well-known heterogeneity of RA, variable prognosis and response to RA treatment are linked to the variable underlying pathogenesis

in individual patients.<sup>7–9</sup> The stratification of RA patients is paramount for the evaluation of prognosis and the selection of appropriate treatment options.<sup>10,11</sup> However, this is challenging due to the lack of clarity regarding the pathogenesis of RA.

It is generally accepted that RA is associated with genetic susceptibility<sup>12–14</sup> and environmental factors.<sup>15,16</sup> Recent studies indicated that a variety of immune cells, including B cells,<sup>17,18</sup> T cells,<sup>19,20</sup> macrophages,<sup>21</sup> dendritic cells,<sup>22,23</sup> mast cells,<sup>24</sup> and neutrophils,<sup>25</sup> were involved in the development and progression of RA. Furthermore, cell death contributes to the pathogenesis of RA through various pathways. The diverse cell death modes, including ferroptosis,<sup>26,27</sup> cuproptosis,<sup>28</sup> and pyroptosis,<sup>29,30</sup> were involved in RA. For instance, the inhibition of ferroptosis may result in a reduction of RA synovium proliferation by means of the alleviation of oxidative stress response, and may also affect the polarization of macrophages to reduce the inflammatory response.<sup>31,32</sup> Liu et al identified disulfidptosis, as a novel form of cell death, was associated with abnormal accumulation of intracellular disulfide bonds, particularly under conditions of glucose deprivation. Cells with high expression of SLC7A11 increase intracellular disulfide bond formation under glucose-starved conditions by promoting cystine uptake. This accumulation depletes intracellular NADPH and then leads to disulfidptosis.<sup>33</sup> Unlike traditional apoptosis or necrosis, disulfidptosis promotes cell death by altering the disulfide bond status of cellular scaffold proteins, a process closely linked to cellular redox balance.<sup>34</sup> It is common for RA synovial tissue characterized by anoxia present of oxidative stress, namely redox imbalance.<sup>35,36</sup> However, to our knowledge, there are no studies on disulfidptosis in RA.

To identify disulfidptosis-related genes associated with RA, we first used transcriptomic data from public databases to confirm disulfidptosis-related genes by weighted gene co-expression network analysis (WGCNA) and differential analysis, and then applied three machine learning algorithms, LASSO regression, SVM-RFE, and Random Forest, to find hub genes. Second, we constructed a prediction model for RA diagnosis using the hub genes. To evaluate the diagnosis value of the model, we performed internal and external validation. Third, we determined two RA subtypes based on the hub genes and compared the disease activity scores, clinical markers, immune checkpoints and immune cells between 2 subgroups of RA patients, trying to indicate the hub genes for RA patient stratification. Finally, the expression of hub genes was experimentally validated between RA patients and healthy controls.

The objective of this study was to construct a promising diagnostic model and validate its diagnostic value, and then provide a patient stratification, effectively distinguishing two RA subtypes by distinct clinical and molecular constituents. This has the potential to identify new therapeutic targets and enhance personalized treatment strategies for RA patients.

## Methods

### Dataset Sources and Preprocessing

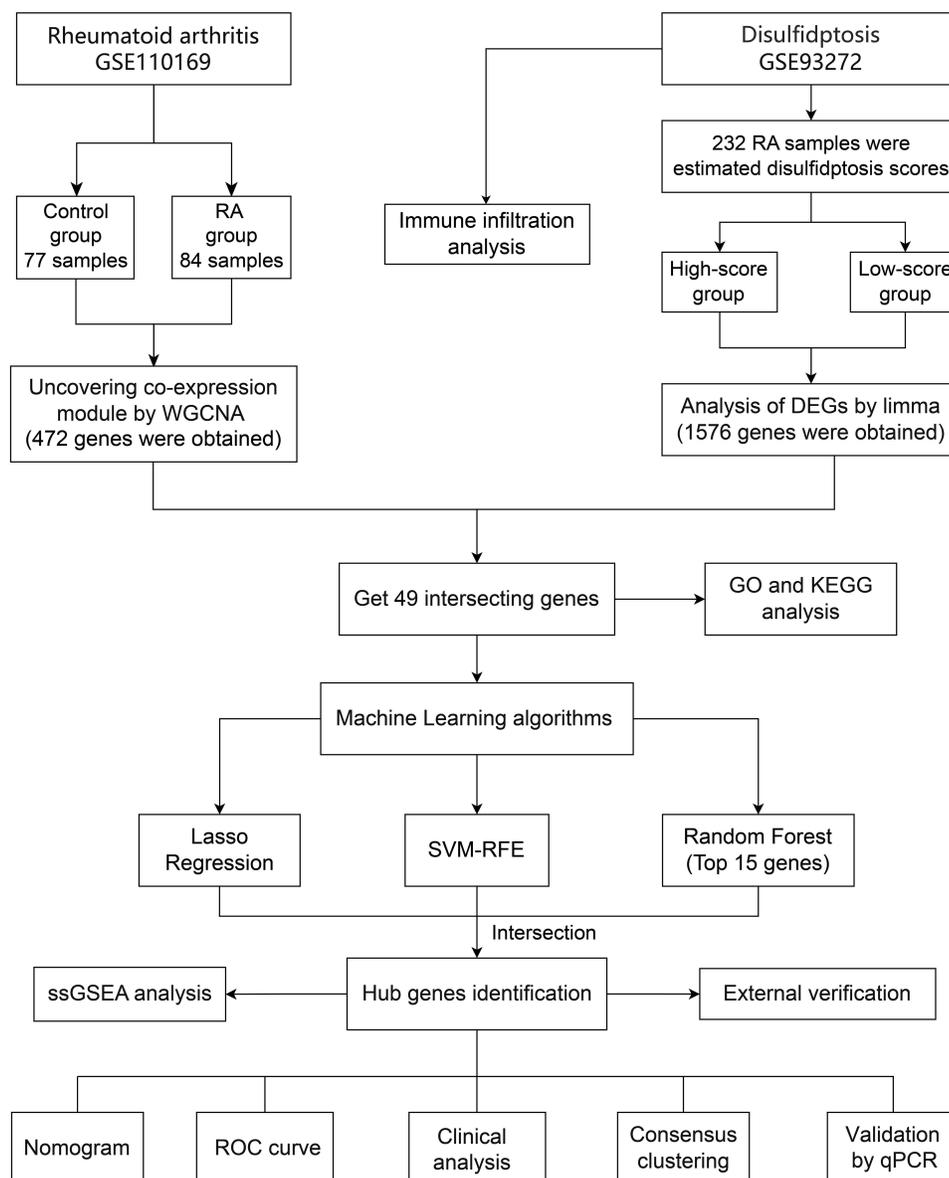
The dataset used in the current study was obtained from the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) using “rheumatoid arthritis” as the search term. Data sets were selected according to the following criteria: gene expression profiles were from human samples; samples included both healthy controls and patients. Three datasets were selected from GSE110169<sup>37</sup> and GSE93272,<sup>38</sup> which contained whole-blood gene expression data using microarray analysis, and GSE89408<sup>39,40</sup> containing synovial biopsy data by high-throughput sequencing. The information of three datasets is shown in Table 1.

The study flowchart was shown (Figure 1). The raw datasets were normalized using the Limma package. In instances where multiple probes were mapped to an identical gene, the median of these values was assigned as the gene’s expression level. This process resulted in the formulation of a refined expression matrix by converting the probe identifiers to their respective gene symbols. The GSE89408 dataset was log2 transformed to normalize the variance within gene expression levels and prepared the data suitable for further analytical processes. Meanwhile, the GSE93272

**Table 1** GEO Datasets Information

Datasets	Platform	Disease	Samples	Measurement	Source Type
GSE110169	GPL13667	Rheumatoid Arthritis	84 RA and 77 controls	Microarray	Whole blood
GSE93272	GPL570	Rheumatoid Arthritis	232 RA and 43 controls	Microarray	Whole blood
GSE89408	GPL11154	Rheumatoid Arthritis	152 RA and 28 controls	High-throughput sequencing	Synovial biopsies

**Abbreviation:** RA, rheumatoid arthritis.



**Figure 1** Flowchart of the study.

**Abbreviations:** DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; RA, rheumatoid arthritis; ROC, receiver operating characteristic; ssGSEA, Single-sample Gene Set Enrichment Analysis; SVM-RFE, Support Vector Machine Recursive Feature Elimination; WGCNA, Weighted Gene Co-expression Network Analysis.

dataset, which presented with gaps in clinical data, underwent imputation via the k-nearest neighbor (KNN) algorithm, leveraging the 10 nearest neighbors to surmise the missing entries.<sup>41</sup>

## Weighted Gene Co-Expression Network Analysis and Identification of Significant Module Genes

The Weighted Gene Co-expression Network Analysis (WGCNA) was performed using the WGCNA R package (version 1.70.3). First, sample clustering was conducted to identify potential outliers ([Supplementary Figure 1A](#)). Second, a scale-free co-expression network was constructed using the pickSoftThreshold function in R to compute the soft thresholding power. The soft power was determined to be 14 (scale independence of 0.9) ([Supplementary Figure 1B](#)). The adjacency matrix was then transformed into a topological overlap matrix (TOM), and the gene ratio along with its corresponding dissimilarity (1-TOM) were computed. Then, hierarchical clustering and dynamic tree-cutting methods were applied to

cluster the genes. The minimum number of gene modules was set to 30, and an appropriate cutoff value was chosen to merge similar modules ([Supplementary Figure 1C and D](#)). Finally, gene significance (GS) and module membership (MM) were calculated to explore the relationship between modules and clinical traits. Meanwhile, we obtained a heatmap illustrating the correlation between 20 gene modules and clinical traits.

## Identification of Differentially Expressed Genes (DEGs) Associated With Disulfidptosis

According to the previous research, ten genes were associated with disulfidptosis, including GYS1, LRPPRC, NCKAP1, NDUFS1, NDUFA11, NUBPL, OXSM, SLC3A2, RPN1, and SLC7A11.<sup>33</sup> To identify DEGs, we estimated disulfidptosis scores within the 232 RA samples of the GSE93272 dataset using the single-sample Gene Set Enrichment Analysis (ssGSEA) algorithm in the GSVA package (version 1.46.0). The RA samples were divided into two groups with high and low disulfidptosis scores, based on the median score. Consistent with the previous study,<sup>42</sup> differential expression analysis between the two groups was performed using the limma package (version 3.54.0) with  $P$  value  $< 0.05$  and  $|\log_2FC| > 0.25$ . The genes meeting these criteria were considered as DEGs. These genes were displayed using volcano plot and heatmap, providing a clear visual differentiation of gene expression to facilitate the interpretation of the results.

## Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) Enrichment Analysis

The pathway enrichment analysis provided by KEGG served as a comprehensive knowledge base for systematic gene function analysis.<sup>43</sup> GO analysis, covering Biological Processes (BP), Molecular Functions (MF), and Cellular Components (CC), yielded invaluable insights into gene functionality.<sup>44</sup> Enrichment analysis was performed by the R package “clusterProfiler (version 3.14.3)”,<sup>45</sup> and the top 10 GO terms in each category were visualized using the R package “ggplot2”. These analyses were based on screening criteria:  $P$  value  $< 0.05$ .

## Machine Learning Algorithms and Hub Genes Identification

To identify hub genes, three machine learning algorithms were employed for variable selection in this research. LASSO regression constructed generalized linear models and was used to mitigate model overfitting, which incorporated penalty terms to shrink unimportant feature coefficients to zero, effectively facilitating variable selection.<sup>46</sup> Support Vector Machine Recursive Feature Elimination (SVM-RFE) was applicable to small datasets and was an algorithm that retained variables relevant to outcomes while effectively eliminating redundant factors.<sup>47</sup> Additionally, Random Forest was capable of handling high-dimensional data, allowing predictive models to be built and ranking the importance of feature variables.<sup>48</sup> The intersection of genes obtained from the three machine learning algorithms was hub genes.

## Single-Sample Gene Set Enrichment Analysis (ssGSEA)

We performed ssGSEA to identify the relationship between hub genes and hallmark gene sets.<sup>49</sup> First, we downloaded the h.all.v7.4.symbols.gmt subset from the Molecular Signatures Database (<http://www.gsea-msigdb.org/gsea/downloads.jsp>) to identify relevant pathways and molecular mechanisms. Then, we calculated the enrichment scores of each sample in the gene sets using gene expression profiles. Finally, an enrichment score matrix was obtained, and the results were visualized as a correlation heatmap.

## The Construction and Validation Accuracy of the Nomogram

To construct a diagnostic model, we employed a regression equation in the GSE110169 dataset. Subsequently, we visualized the nomogram using the R package “rms (version 6.2.0)”. The nomogram was based on the regression coefficients of all features. We established scoring criteria for each feature’s different expression levels. By calculating the total score for each sample, we were able to predict the probability of each sample developing RA. Furthermore, we evaluated the accuracy of the model predictions and validated the predictive value of hub genes using ROC curve, calibration curve, and decision curve.

## External Verification of Nomogram and Hub Genes

We performed external validation of the nomogram and hub genes using the GSE93272 and GSE89408 datasets. The GSE93272 dataset included 275 whole blood samples, of which 232 were RA samples and 43 were controls. The GSE89408 dataset comprised 180 synovial biopsy samples, including 152 RA samples and 28 controls. ROC curves were plotted for each dataset to assess the predictive accuracy.

## Immune Infiltration Analysis

The “Cibersort” algorithm was used in the GSE93272 dataset to convert the normalized gene expression matrix into the composition of infiltrating immune cells.<sup>50</sup> We removed immune cells with zero abundance and obtained the abundance of 20 immune cell types in RA samples and controls utilizing the R package “Cibersort”. The relative proportions of the 20 immune cell types in both groups were visualized using a barplot, and the expression differences of immune cells were compared between the two groups using boxplots. The correlations were displayed among the 20 immune cell types using the R package “corrplot”. Additionally, the correlation between hub genes and the 20 immune cell types was assessed.

## Clinical Data Analysis

The clinical scores, including disease activity score in 28 joints (DAS28), clinical disease activity index (CDAI), and simplified disease activity index (SDAI), were commonly used to assess disease activity for RA patients.<sup>51–54</sup> In this analysis, we further clarified the association between hub genes and clinical scores, including DAS28-CRP, DAS28-ESR, CDAI, and SDAI, using data from the GSE93272 dataset. The RA patients were categorized into four groups based on different scoring systems as described in Table 2. We compared the expression differences of hub genes among different groups and performed correlation analyses between hub genes and clinical indicators. Scatter plots were generated to illustrate the correlation between hub genes and DAS28-CRP, DAS28-ESR, CDAI, SDAI, MMP3 and RF.

## Consensus Clustering Analysis

Stratifying patients based on the disease characteristics was essential for personalized treatment of RA.<sup>55</sup> The “ConsensusClusterPlus” package was used to generate a consensus matrix and cumulative distribution function (CDF) plot. The relative change in the area under the CDF curve was assessed, and a tracking plot was used to determine the optimal number of clusters. Then, principal component analysis (PCA) was used to assess the ability to distinguish patients with different subtypes. The clinical markers differences between the two subtypes were compared and the results were displayed using boxplots. The visualization of differential expression of hub genes among different subtypes was achieved using a heatmap and boxplot. Finally, we analyzed the infiltration of 28 immune cells by using ssGSEA. The immune checkpoints were compared among different subtypes.

**Table 2** The Assessment of RA Disease Activity by Different Evaluation Tools

Disease activity	DAS28	CDAI	SDAI
remission	< 2.6	≤ 2.8	≤ 3.3
low activity	≥ 2.6 and < 3.2	> 2.8 and < 10	> 3.3 and < 11
moderate activity	≥ 3.2 and ≤ 5.1	≥ 10 and ≤ 22	≥ 11 and ≤ 26
high activity	> 5.1	> 22	> 26

**Abbreviations:** CDAI, clinical disease activity index; DAS28, disease activity score in 28 joints; RA, rheumatoid arthritis; SDAI, simplified disease activity index.

## Validation of Hub Gene Expression by qPCR Analysis

This study included blood samples from 11 RA patients and 12 healthy individuals. The RA patients met the classification criteria of the ACR/EULAR 2010.<sup>56</sup> All participants signed a written informed consent form, and the study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University (Approval No. XJTU1AF2022LSK-168).

Total RNA was extracted from the whole blood using an RNA isolation kit (5201050, Simgen, China). The RNA was then reverse transcribed into cDNA using the Reverse Transcription Kit (RR037A, Takara, Japan). The qPCR was performed using TB Green Premix Ex Taq II (RR820A, Takara, Japan). The sequences of the primers used were as follows: KLHL2 (forward primer: 5'-AGAGCCGAGCAAAGAGAGTTA-3'; reverse primer: 5'-CCTAAGCAGTTGACAGGGTGA-3'); POLK (forward primer: 5'-TGAGGGACAATCCAGAATTGAAG-3'; reverse primer: 5'-CTGCACGAACACCAAATCTCC-3'); CLEC4D (forward primer: 5'-CTGATACCTTCGGTTATTGCTGT-3'; reverse primer: 5'-GCACTCCTGTGCCTCTCTTAC-3'); NXT2 (forward primer: 5'-GGACAAGGCCACCTTAATATGG-3'; reverse primer: 5'-TGGAACTCACTAGAAGGCAATGT-3'); and 18S (forward primer: 5'-GCAATTATCCCATGAACG-3'; reverse primer: 5'-GGCCTCACTAAACCATCCAA-3'). The 18S gene served as an internal control, and the relative expression of genes was determined using the  $2^{-\Delta\Delta CT}$  method.

## Data Analysis

The bioinformatics analyses and data visualization were performed using R software (version 4.2.0). Statistical analysis was conducted using GraphPad Prism (version 9.0). Group comparisons were made using Student's *t*-test for two groups, and one-way ANOVA for three or more groups (with Holm-Sidak's multiple comparison), with statistical significance set at  $P < 0.05$ .

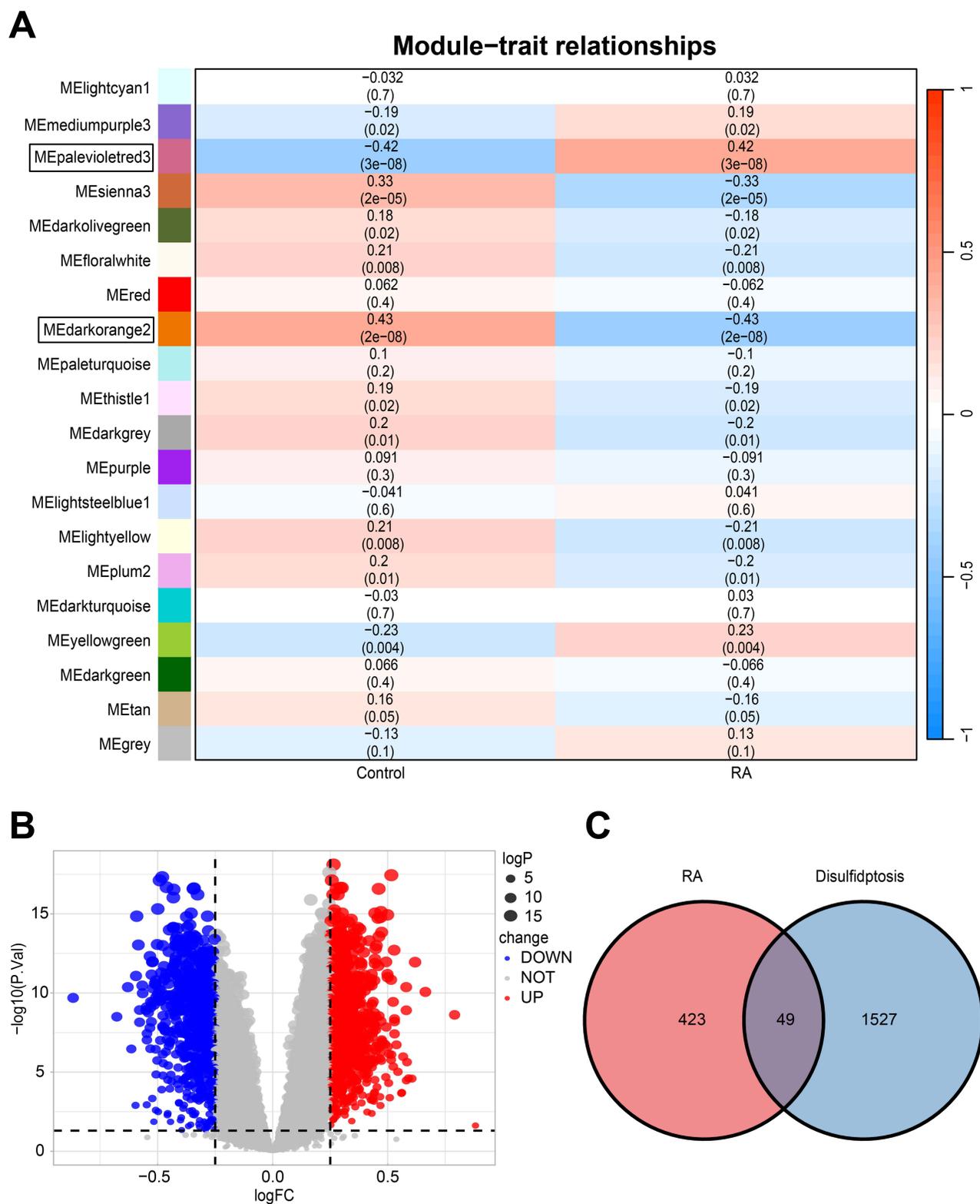
## Results

### Identification of Disulfidptosis-Related Genes in RA and Functional Enrichment Analysis

To identify significant module genes in RA, we performed WGCNA analysis using the GSE110169 dataset. The correlation heatmap revealed that the palevioletred3 module was the most significantly positively correlated with RA ( $r = 0.42$ ,  $P = 3 \times 10^{-8}$ ), while the darkorange2 module was the most significantly negatively correlated ( $r = -0.43$ ,  $P = 2 \times 10^{-8}$ ) (Figure 2A). Furthermore, the Module Membership and Gene Significance for the palevioletred3 and darkorange2 modules showed significant correlation (Supplementary Figure 2A and B). Genes were selected based on criteria of geneTraitSignificance  $> 0.40$  and geneModuleMembership  $> 0.60$  to further analyze these modules. Genes meeting these conditions were considered important for the modules. A total of 472 genes, identified as important genes from both modules, were merged for subsequent analysis.

To explore the relationship between RA and disulfidptosis, we calculated the disulfidptosis score for each sample of RA. A total of 1576 disulfidptosis-related DEGs were identified comparing the high-scoring group with the low-scoring group. The volcano plot and heatmap displayed all DEGs, with 871 genes upregulated and 705 genes downregulated (Figure 2B, Supplementary Figure 3). Finally, we identified 49 important genes by intersecting the DEGs associated with disulfidptosis and the significant module genes in RA (Figure 2C).

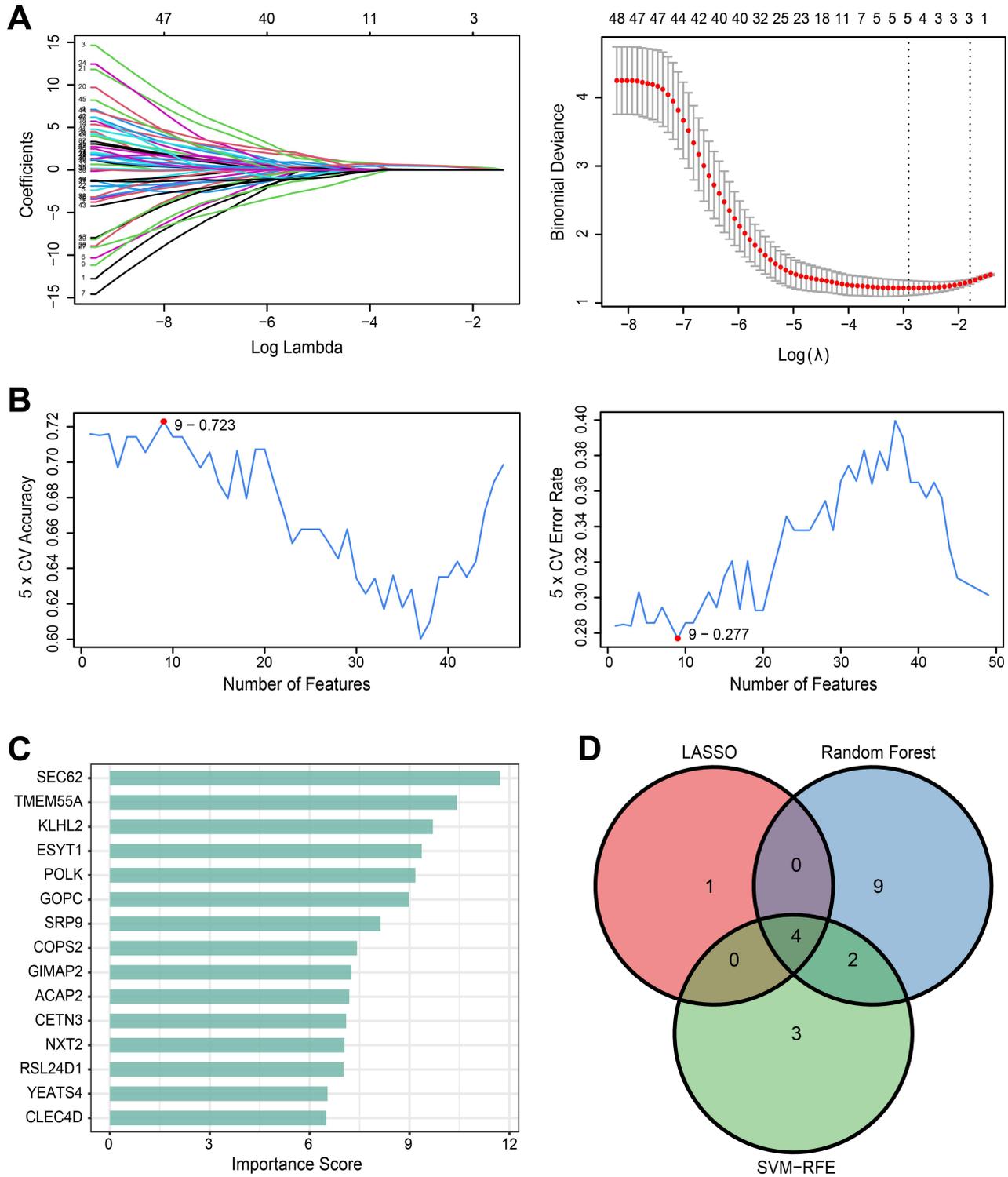
The KEGG pathway enrichment analysis revealed the association between the intersecting genes and signaling pathways. The analysis results indicated that the intersecting genes were primarily enriched in pathways such as "Ribosome", "Cardiac muscle contraction", "Protein export", "Thermogenesis", and "Oxidative phosphorylation" (Supplementary Figure 4A). GO analysis, covering Biological Processes (BP), Cellular Components (CC), and Molecular Functions (MF), was employed to assess the functional enrichment of genes. In the BP term, significant enrichments included "amide biosynthetic process", "protein localization to membrane", and "establishment of protein localization to membrane" (Supplementary Figure 4B). Notable differences in CC were observed in "organelle envelope", "envelope", and "ribosomal subunit" (Supplementary Figure 4C). Within the MF category, prominent enrichments were seen in "structural constituent of ribosome", "cytochrome-c oxidase activity", and "heme-copper terminal oxidase activity" (Supplementary Figure 4D).



**Figure 2** Identification of disulfidptosis-related genes in RA. **(A)** The heatmap showed the relationship between 2 traits and 20 modules. **(B)** The volcano plot displayed DEGs associated with disulfidptosis. **(C)** The Venn plot showed the intersection of DEGs associated with disulfidptosis and significant module genes in RA. **Abbreviations:** DEGs, differentially expressed genes; RA, rheumatoid arthritis.

# Identification of Hub Genes via Machine Learning Algorithms and ssGSEA

To select hub genes for subsequent model construction, we employed three machine learning algorithms for feature selection. Lasso regression results revealed 5 genes (CLEC4D, KLHL2, MRPL50, NXT2, POLK) with the lowest binomial deviation



**Figure 3** Identification of hub genes via machine learning algorithms. **(A)** The five genes were selected as the features with the lowest binomial deviation by the LASSO regression. **(B)** The nine genes were identified as the features with the lowest error rate and highest accuracy via the SVM-RFE. **(C)** The top fifteen genes were ranked according to their importance score based on the Random Forest. **(D)** The four hub genes were identified through the intersection of three machine learning algorithms. **Abbreviations:** CV, Cross Validation; SVM-RFE, Support Vector Machine Recursive Feature Elimination.

**Table 3** Genes Selected by Three Machine Learning Algorithms

Machine learning algorithm	Gene Symbol
Lasso regression	CLEC4D, KLHL2, MRPL50, NXT2, POLK
SVM-RFE	ACAP2, ACSL4, CD46, C5orf28, CLEC4D, KLHL2, NXT2, POLK, TMEM55A
Random Forest	ACAP2, CETN3, CLEC4D, COPS2, ESYT1, GIMAP2, GOPC, KLHL2, NXT2, POLK, RSL24D1, SEC62, SRP9, TMEM55A, YEATS4

**Abbreviation:** SVM-RFE, Support Vector Machine Recursive Feature Elimination.

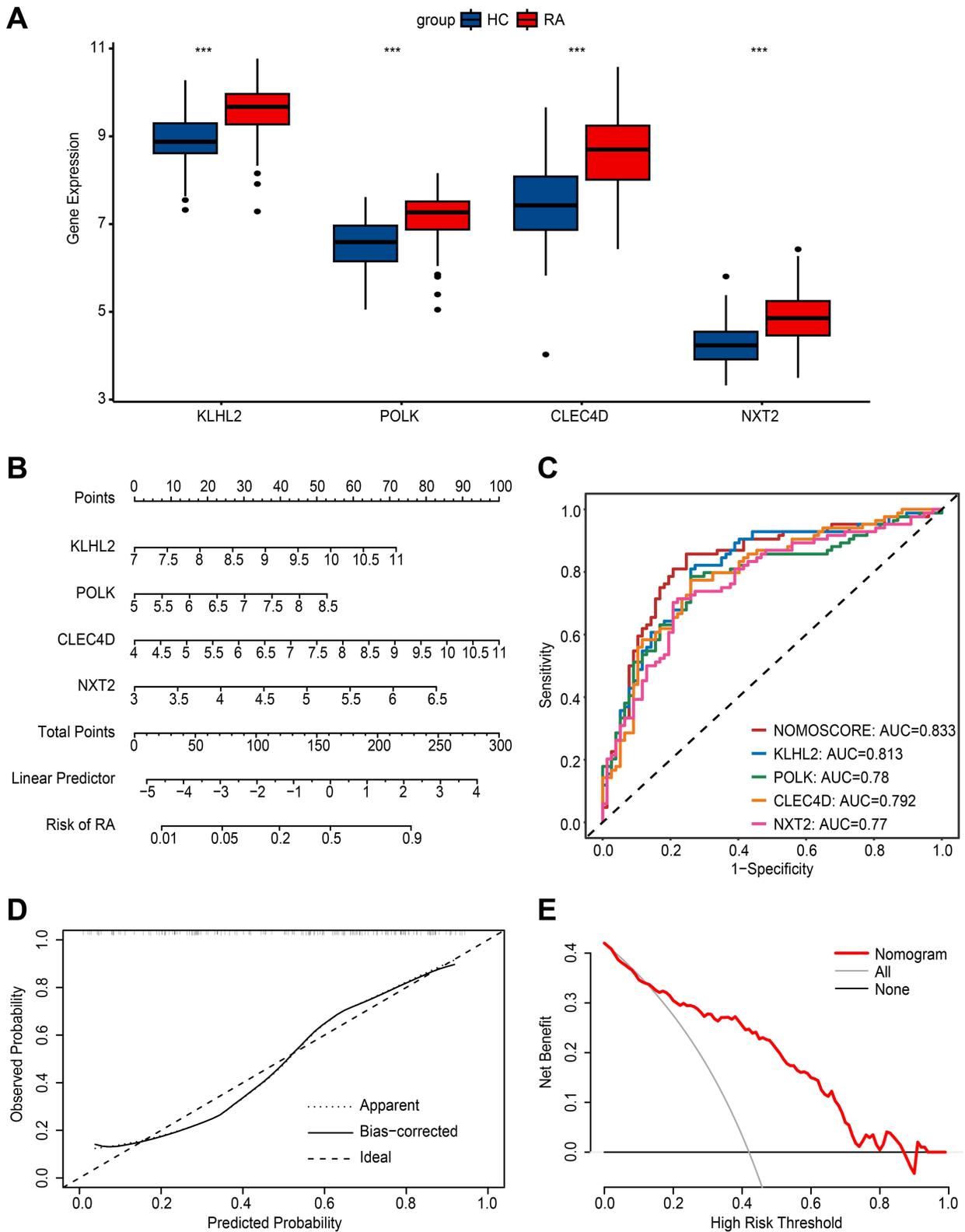
(Figure 3A). SVM-RFE results indicated 9 genes (ACAP2, ACSL4, CD46, C5orf28, CLEC4D, KLHL2, NXT2, POLK, TMEM55A) with the highest accuracy and the lowest error rate (Figure 3B). Random Forest analysis displayed the top 15 important genes, including ACAP2, CETN3, CLEC4D, COPS2, ESYT1, GIMAP2, GOPC, KLHL2, NXT2, POLK, RSL24D1, SEC62, SRP9, TMEM55A, and YEATS4 (Figure 3C). Table 3 summarizes all results. As shown in Figure 3D, 4 genes, including KLHL2, POLK, CLEC4D and NXT2, were identified by the overlapping areas of the three machine learning algorithms. Furthermore, we investigated the relationship between hub genes and hallmark gene sets. A correlation heatmap depicted that the associations of the four genes had a significant correlation with “IL6\_JAK\_STAT3\_SIGNALING” and “WNT\_BETA\_CATENIN\_SIGNALING” (Supplementary Figure 5). These hallmark gene sets were regarded as the signaling pathways associated with RA.

## The Model Construction Based on Hub Genes and the External Verification

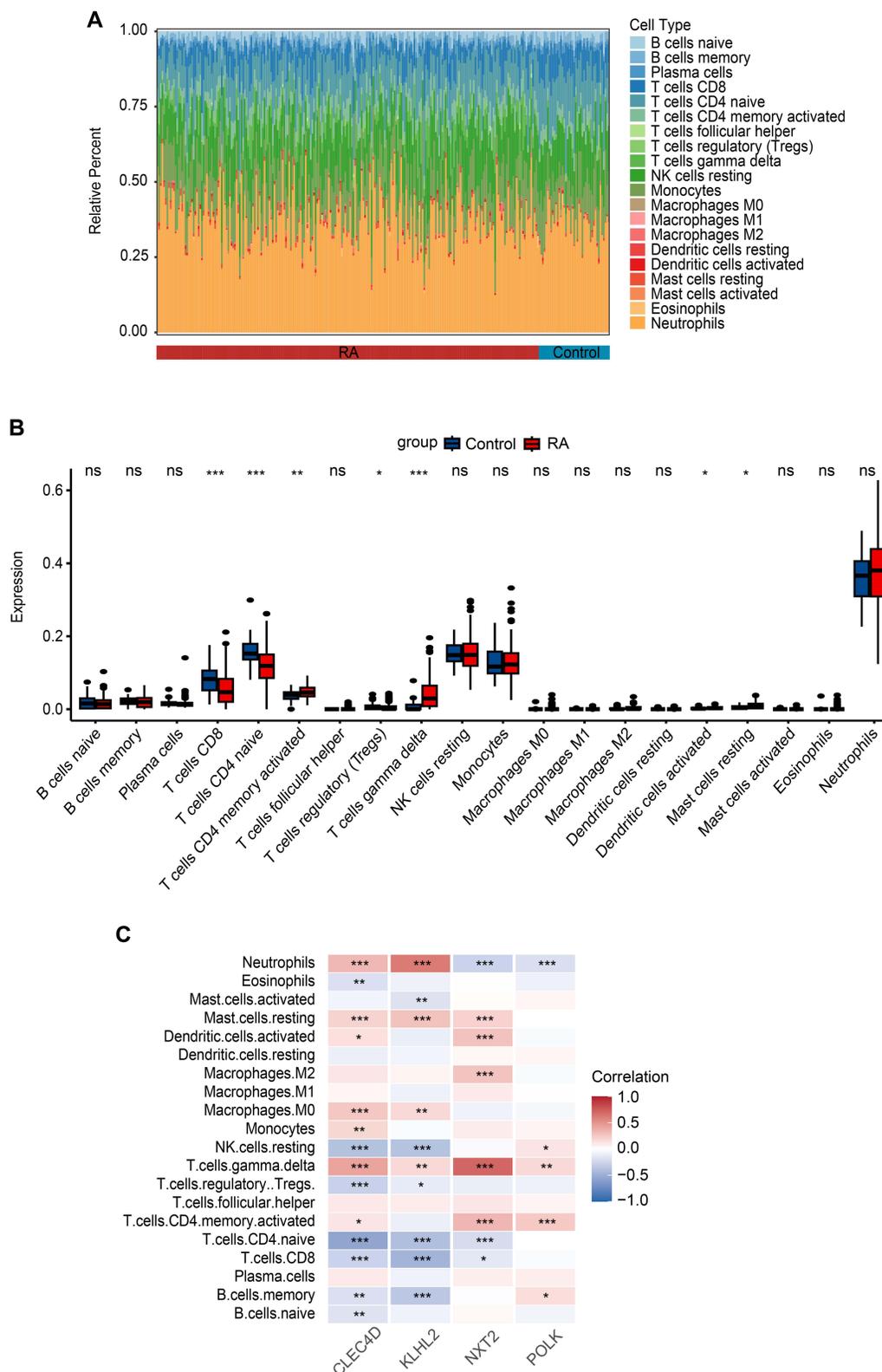
To assess the value of the hub genes for RA diagnosis, we used data from three datasets: GSE110169 as the training set, and GSE89408 and GSE93272 as external validation sets, for the model construction and validation. Initially, we observed significantly higher expression of hub genes in RA compared with Healthy control (HC) across all three datasets (Figure 4A, Supplementary Figure 6A and B). Subsequently, the model was constructed and visualized using the nomogram in GSE110169. The nomogram displayed scoring criteria for each feature’s different expression levels. By calculating the total score for each sample, we predicted the probability of each sample developing RA (Figure 4B). ROC curves demonstrated that the model exhibited worthy predictive performance and the model’s predictive performance surpassed that of each hub genes (Figure 4C). Similarly, in an external validation set that included data from both whole blood and synovial biopsy samples, the model also showed promising predictive value (Supplementary Figure 6C and D). In the GSE110169 dataset, the AUC values and 95% confidence intervals for NOMOSCORE (AUC: 0.833, 95% CI: 0.768–0.899), KLHL2 (AUC: 0.813, 95% CI: 0.745–0.880), CLEC4D (AUC: 0.792, 95% CI: 0.722–0.862), NXT2 (AUC: 0.77, 95% CI: 0.696–0.843), and POLK (AUC: 0.78, 95% CI: 0.708–0.853) were shown (Supplementary Figure 7A). Supplementary Figure 7B shows that in the GSE93272 dataset, the NOMOSCORE model (AUC: 0.895, 95% CI: 0.847–0.943) demonstrated superior predictive power compared to each hub gene (Supplementary Figure 7B). Furthermore, we were somewhat surprised to find that in the GSE89408 dataset, which contained synovial biopsy samples, NOMOSCORE exhibited markedly higher predictive power (AUC: 0.977, 95% CI: 0.952–1.000) compared to each hub gene (Supplementary Figure 7C). The calibration curve indicated relatively high accuracy of the nomogram (Figure 4D). Decision curve analysis showed that the net benefit of the constructed model exceeded that of the default method (Figure 4E).

## Immune Infiltration Analysis and the Relevance With Hub Genes

The imbalance of immune cells played a pivotal place in the pathogenesis of RA, which necessitated a comprehensive analysis of immune infiltration in each sample. Utilizing the Cibersort analysis, we obtained profiles of 20 immune cell types after excluding two cells with zero abundance. The barplot illustrated the relative percentages of these 20 immune cells (Figure 5A). Then, significant differences were demonstrated among 7 immune cell types, including T cells CD8, T cells CD4 naïve, T cells CD4 memory activated, Tregs, T cells gamma delta, dendritic cells activated and mast cells resting among different groups (Figure 5B). Furthermore, correlation analysis revealed significant associations between hub genes and the 20 immune cell types. For example, CLEC4D demonstrated a significant positive correlation with



**Figure 4** Validation of hub genes in RA. **(A)** The different expression of the four hub genes in RA and HC; \*\*\*,  $P < 0.001$ . **(B)** The hub genes were utilized in the construction of the nomogram to predict the risk of RA. **(C)** The predicted value of the nomogram and the hub genes using the ROC curve. **(D)** The calibration curve showed the accuracy of the nomogram. **(E)** The decision curve analysis indicated that the net benefit of the constructed model surpassed that of the default method. **Abbreviations:** AUC, area under the curve; HC, Healthy control; RA, rheumatoid arthritis; ROC, receiver operating characteristic.



**Figure 5** Immune infiltration analysis from RA to Control and correlation between hub genes and immune cells. **(A)** The relative percent of 20 immune cells in each sample. **(B)** The different expression of 20 immune cells between RA and control; ns: no significance; \*:  $P < 0.05$ ; \*\*:  $P < 0.01$ ; \*\*\*:  $P < 0.001$ . **(C)** Correlation analysis of immune cells with hub genes.

**Abbreviation:** RA, rheumatoid arthritis.

T cells gamma delta and a significant negative correlation with T cells CD4 naïve. Similarly, KLHL2 exhibited a significant positive correlation with neutrophils and a significant negative correlation with T cells CD8. Moreover, NXT2 displayed a significant positive correlation with T cells gamma delta and a significant negative correlation with neutrophils. Lastly, POLK showed a significant positive correlation with T cells CD4 memory activated and a significant negative correlation with neutrophils (Figure 5C). The correlation heatmap displayed notable associations (Supplementary Figure 8), including dendritic cells resting exhibiting the highest positive correlation with macrophages M1 ( $r = 0.74$ ), while B cells memory showed the greatest negative correlation with B cells naïve ( $r = -0.5$ ).

## The Correlation Between Hub Genes Expression and RA Disease Activity

The scoring systems, including DAS28, CDAI and SDAI, were the most used to assess disease activity in RA. The RA patients were divided for 4 groups according to disease activity. Then, the expression of hub genes was compared among the 4 groups. Overall, the expression of CLEC4D, KLHL2 and NXT2 showed an elevated trend with increasing disease activity, conversely, the expression of POLK exhibited a decreasing trend in the patients stratified by DAS28-CRP (Figure 6A) as well as DAS28-ESR, CDAI, and SDAI (Supplementary Figures 9A and B, Supplementary Figures 10). The correlation heatmaps demonstrated the association between hub genes and clinical indicators (Figure 6B, Supplementary Figure 11). We observed significant positive correlations between CLEC4D, KLHL2, NXT2 and DAS28-CRP scores ( $r = 0.29$ ,  $P = 6.6 \times 10^{-6}$ ;  $r = 0.23$ ,  $P = 3.51 \times 10^{-4}$ ;  $r = 0.19$ ,  $P = 3.04 \times 10^{-3}$ , respectively), while POLK showed a significant negative correlation with DAS28-CRP scores ( $r = -0.19$ ,  $P = 3.04 \times 10^{-3}$ , Figure 6C). The scatter plots also illustrated the same correlations between hub genes and DAS28-ESR, CDAI, SDAI, MMP3, and RF (Supplementary Figure 12A–E).

## Two RA Subtypes Based on Hub Genes Had Different Disease Activity

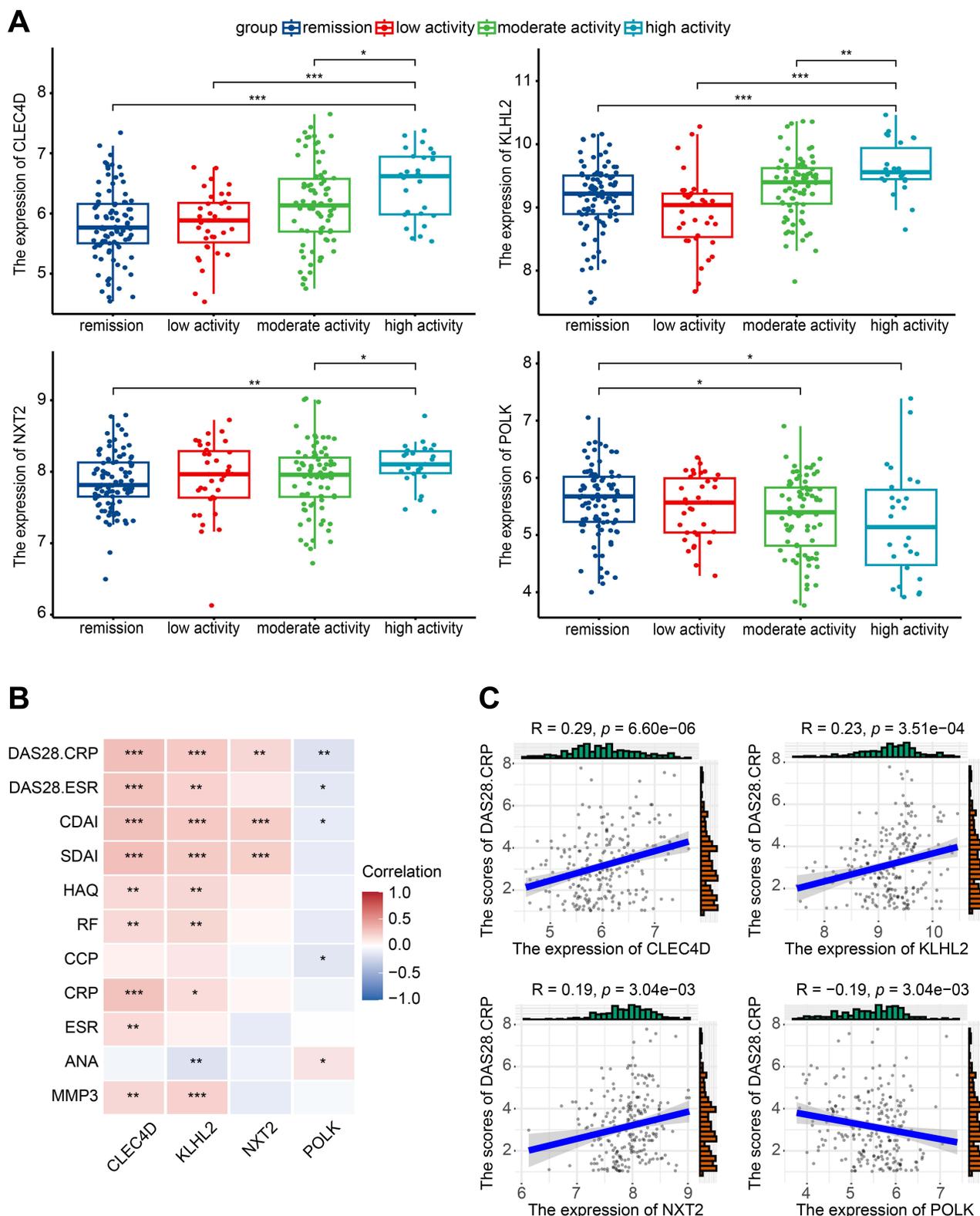
In RA samples, consensus clustering analysis was performed using the expression profiles of hub genes. Based on the consensus matrix, CDF plot, relative alterations in the area under the CDF curve, and tracking plot results, we determined the two subtypes (Supplementary Figure 13A–D). PCA results demonstrated significant differences between most RA samples between the two subtypes (Figure 7A). Boxplot and heatmap revealed significant differences in the expression of hub genes between the two subtypes. Specifically, cluster B had high expression of KLHL2, CLEC4D and NXT2 and low expression of POLK, on the contrary, cluster A had high expression of POLK as well as low expression of KLHL2, CLEC4D and NXT2 (Figure 7B, Supplementary Figure 14A). Furthermore, the markers associated with RA disease activity were compared between two groups, the patients of cluster B had significantly higher levels of DAS28-CRP, DAS28-ESR, CDAI, SDAI, RF, CRP, and MMP3 than the subjects of cluster A (Figure 7C). Additionally, we observed significant differences in 13 immune checkpoints and 15 immune cells between the two subtypes, mainly in activated dendritic cell, effector memory CD8 T cell, Eosinophil, Gamma delta T cell, natural killer cell and associated activation markers (Supplementary Figure 14B, Figure 7D).

## The Expression Levels of Hub Genes in the in-House Validation Cohort

To validate the differential expression of hub genes between RA patients and healthy controls, we compared the expression levels of CLEC4D, KLHL2, NXT2, and POLK in 11 RA patients and 12 healthy controls. The results showed that CLEC4D, POLK, KLHL2, and NXT2 were more highly expressed in RA patients than those in the healthy controls (Figure 8), which was consistent with validation from public database analyses.

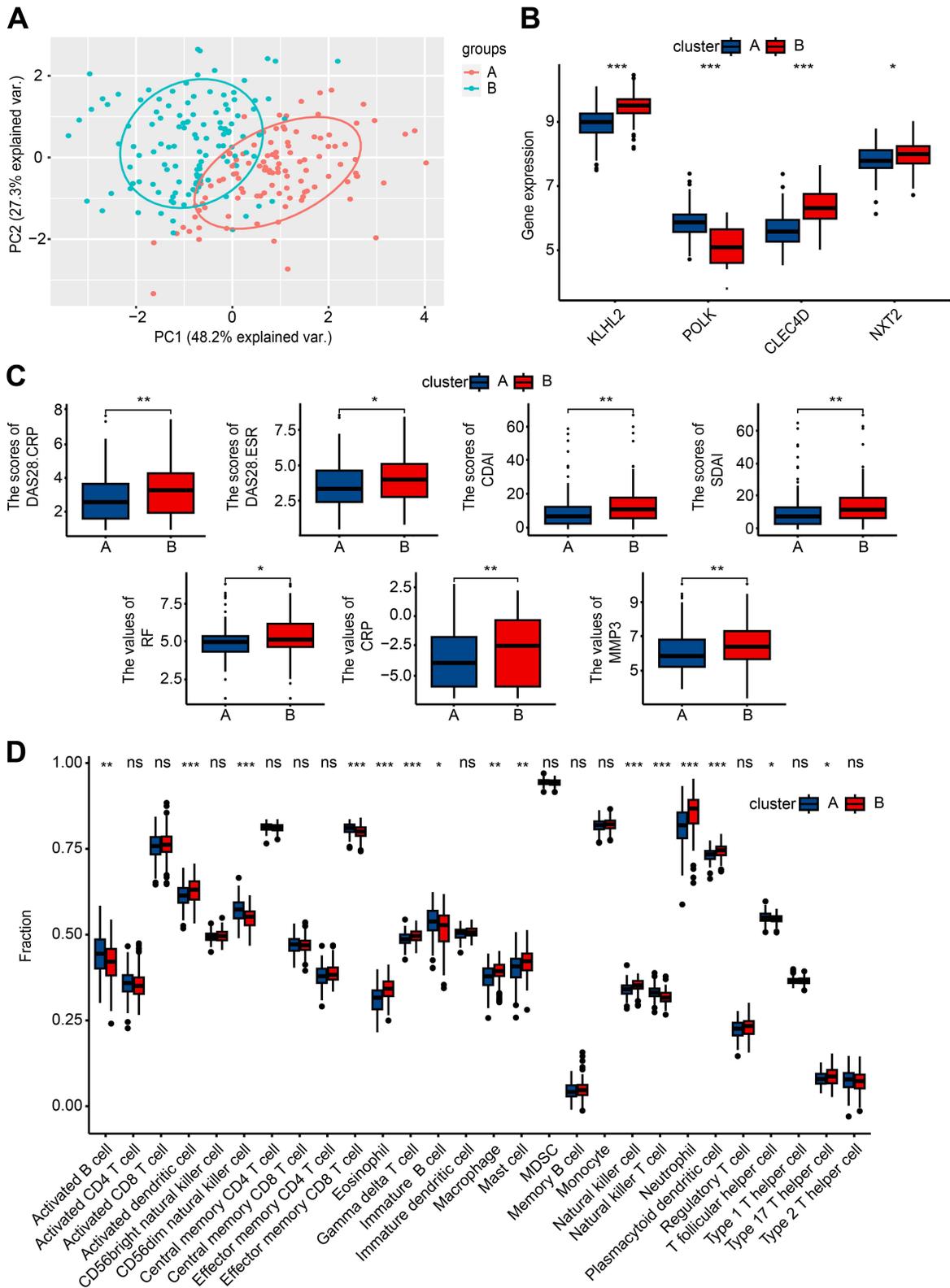
## Discussion

The heterogeneity of RA has posed significant challenges in the early diagnosis and optimal treatment selection for RA.<sup>57</sup> A comprehensive understanding of the underlying pathogenesis of RA has been instrumental in developing patient stratification strategies. Recently, disulfidptosis, a novel form of cell death triggered by intracellular accumulation of excess cysteine, has been identified as a potential mechanism in the pathogenesis of RA.<sup>33,34</sup> Disulfidptosis modulated the equilibrium of redox reactions and cytoskeletal stability in a variety of cells, and may be involved in immunity activation and inflammatory response, which influenced the onset and progression of RA. However, there is currently a paucity of research on disulfidptosis in RA. To the best of our knowledge, this study represented the inaugural attempt



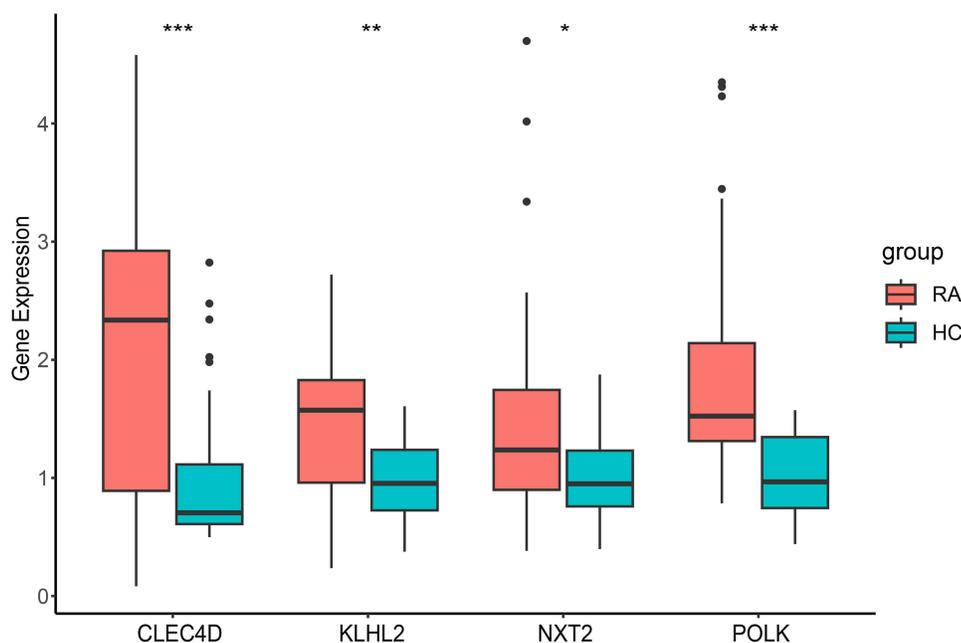
**Figure 6** The correlation between hub genes expression and RA disease activity. **(A)** The expression levels of four hub genes were compared among four different disease activity stages based on DAS28.CRP scores in RA; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ . **(B)** The heatmap displayed the correlation and significant difference between the expression patterns of four genes and clinical markers. **(C)** The scatter plots demonstrated a significant correlation between the expression of the four hub genes and DAS28.CRP.

**Abbreviations:** CDAI, clinical disease activity index; DAS28, disease activity score in 28 joints; HAQ, health assessment questionnaire; RA, rheumatoid arthritis; SDAI, simplified disease activity index.



**Figure 7** Identification of two subtypes based on the hub genes in RA samples. **(A)** The PCA plot showed that RA samples were categorized as two subtypes. **(B)** The different expression of the hub genes between the two subtypes. **(C)** The boxplots displayed the differences of clinical markers between the two subtypes. **(D)** The different expression of 28 immune cells between the two subtypes; ns: no significance; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ .

**Abbreviations:** CDAI, clinical disease activity index; DAS28, disease activity score in 28 joints; PCA, principal component analysis; RA, rheumatoid arthritis; SDAI, simplified disease activity index.



**Figure 8** The expression differences of four hub genes (KLHL2, POLK, CLEC4D, NXT2) between RA and HC; \*:  $P < 0.05$ ; \*\*:  $P < 0.01$ ; \*\*\*:  $P < 0.001$ .  
**Abbreviations:** HC, Healthy control; RA, rheumatoid arthritis.

to investigate the relationship between RA and disulfidptosis using bioinformatics techniques and machine learning algorithms. It encompassed comprehensive analysis of hub genes in both peripheral blood and synovial samples and evaluated the diagnostic value of the model constructed using the hub genes for RA. Based on the differential expression of the 4 hub genes, two subtypes of RA patients were determined, and different disease activity was compared between the two subgroups, which lay the groundwork for future approaches to patient stratification.

To investigate the potential correlation between RA and disulfidptosis, we identified 49 key genes by overlapping the key module genes in RA and disulfidptosis-related DEGs using WGCNA and differential analysis. Then, we identified 4 hub genes (KLHL2, POLK, CLEC4D and NXT2) among the 49 key genes by three machine learning algorithms. The 4 hub genes were associated with the autoimmune disease by involvement in the inflammatory response as well as differentiation and development of immune cells. Specifically, KLHL2, binding to With-No-Lysine Kinases (WNKs), modulated the activity of the NLRP3 inflammasome by sensing intracellular chloride and potassium ion levels, thereby leading to the production of IL-1 $\beta$  and other cytokines.<sup>58,59</sup> Moreover, KLHL2 was considered to be involved in B cell activation by regulating the B cell receptor (BCR) signaling pathway as well as inhibition of Treg-mediated tolerance by promoting FOXP3 deacetylation and destabilization.<sup>60,61</sup> CLEC4D was demonstrated to initiate signal cascades that induce robust innate immune and inflammatory responses by activating spleen tyrosine kinase (Syk) and the NF- $\kappa$ B signal pathway.<sup>62–64</sup> Furthermore, CLEC4D was linked to the differentiation of Th17 cells in certain inflammatory diseases.<sup>65</sup> The existing literature on NXT2 in autoimmune diseases is limited. In the current study, NXT2 was found to be enriched in multiple pathways, including the IL6-JAK-STAT3 pathway, which suggested a potential role for NXT2 in the pathogenesis of RA. Contrarily, POLK, a specialized DNA polymerase, may mitigate premature aging and functional decline of T cells by facilitating DNA damage response repair, thereby suppressing inflammation.<sup>66–69</sup> In summary, four hub genes played a potential role in immune responses and inflammation of RA.

In this study, we performed validation using an in-house cohort to confirm the significantly higher expression of 4 hub genes in RA compared with the healthy controls, which improved the reliability of the results. To assess the diagnostic value of four hub genes, we developed a model using machine learning algorithms and performed external verification to confirm the model power for RA diagnosis. Due to prediction performance improvement, error rates decline and dependability elevation of machine learning, the model was demonstrated superior diagnostic value compared with each hub gene individually by ROC analysis, calibration curve and decision curve analysis. Synovitis is the most basic pathological manifestation of RA, as is well known.<sup>1</sup> In the external validation set, the model exhibited a markedly enhanced predictive capacity in synovial biopsy

samples, thereby reinforcing the pivotal role of the four hub genes in RA pathogenesis. To further determine the role of hub genes in RA, we compared the expression of the 4 gene among the RA patients with different disease activity. Consistent with the biological function of the genes mentioned before, the expression of CLEC4D, KLHL2, and NXT2 exhibited an elevated trend with increasing RA disease activity, while that of POLK exhibited a decreasing trend in patients stratified by four scoring systems of disease activity assessment in RA. Based on differential expression of the 4 hub genes, we further determined the two subtypes of RA patients by consensus clustering analysis, consensus matrix, CDF plot, relative alterations in the area under the CDF curve, and tracking plot results. Among the RA patients, the cluster A subgroup exhibited high expression of POLK and low expression of KLHL2, CLEC4D, and NXT2, whereas the cluster B subgroup exhibited high expression of KLHL2, CLEC4D, and NXT2 and low expression of POLK. As anticipated, the patients in the cluster B subgroup exhibited significantly higher RA disease activity compared with those in the cluster A subgroup, which provided a new approach to RA patients stratification.

This study inevitably has some limitations. Although we demonstrated the association of the 4 hub genes with clinical markers, cell or animal model was deficient. Integrating single-cell transcriptomic analysis may enhance the reliability of our findings and provide further insights into the role of the hub genes in RA.

## Conclusion

This study systematically delineated the complex correlation between disulfidptosis and RA and identified four hub genes (KLHL2, CLEC4D, NXT2, POLK). Additionally, a promising diagnostic model and stratification strategy for RA was developed to assess the risk of RA onset and RA disease activity, providing unique insights into the role of disulfidptosis in RA, which may contribute to improving early diagnosis and personalized treatment strategies.

## Data Sharing Statement

Datasets used in the study (GSE110169, GSE93272, GSE89408) can be downloaded without restriction from the public GEO database. GSE110169: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110169>; GSE93272: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE93272>; GSE89408: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89408>. Upon reasonable request, data analysis procedures are available from the first and corresponding authors.

## Ethics Approval and Consent to Participate

The study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University (No. XJTU1AF2022LSK-168), Xi'an, China. Written informed consent was obtained from all participants before study enrollment.

## Acknowledgments

We appreciate Sangerbox for providing us with the platform for partly data analysis.

## Funding

The work was supported by the Shaanxi Province Natural Science Foundation (No. 2022JQ-962 and No. 2023-JC-YB-754), and First Affiliated Hospital of Xi'an Jiaotong University Foundation (No. 2022MS-04).

## Disclosure

All authors declare that they have no disclosures.

## References

1. Lee DM, Weinblatt ME. Rheumatoid arthritis. *Lancet*. 2001;358(9285):903–911. doi:10.1016/S0140-6736(01)06075-5
2. Aletaha D, Funovits J, Smolen JS. Physical disability in rheumatoid arthritis is associated with cartilage damage rather than bone destruction. *Ann Rheum Dis*. 2011;70(5):733–739. doi:10.1136/ard.2010.138693
3. McInnes Iain B, Georg S. The Pathogenesis of Rheumatoid Arthritis. *N Engl J Med*. 2011;365(23):2205–2219. doi:10.1056/NEJMra1004965

4. Combe B, Landewe R, Lukas C, et al. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheumatic Dis.* 2006;66(1):34–45. doi:10.1136/ard.2005.044354
5. Aletaha D, Smolen JS. Diagnosis and management of rheumatoid arthritis: a review. *JAMA.* 2018;320(13):1360–1372. doi:10.1001/jama.2018.13103
6. Radu AF, Bungau SG, Negru PA, Marcu MF, Andronie-Cioara FL. In-depth bibliometric analysis and current scientific mapping research in the context of rheumatoid arthritis pharmacotherapy. *Biomed. Pharmacother.* 2022;154:113614. doi:10.1016/j.biopha.2022.113614
7. Tak PP. A personalized medicine approach to biologic treatment of rheumatoid arthritis: a preliminary treatment algorithm. *Rheumatology (Oxford).* 2012;51(4):600–609. doi:10.1093/rheumatology/ker300
8. Myngbay A, Bexetov Y, Adilbayeva A, et al. CTHRC1: a new candidate biomarker for improved rheumatoid arthritis diagnosis. *Front Immunol.* 2019;10:1353. doi:10.3389/fimmu.2019.01353
9. van der Pouw Kraan TCTM, Wijbrandts CA, van Baarsen LGM, et al. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. *Ann Rheum Dis.* 2007;66(8):1008–1014. doi:10.1136/ard.2006.063412
10. Martinez-Prat L, Martinez-Taboada V, Santos C, Lopez-Hoyos M, Mahler M. Anti-protein-arginine deiminase 4 IgG and IgA delineate severe rheumatoid arthritis. *Diagnostics.* 2022;12(9):2187. doi:10.3390/diagnostics12092187
11. Suter LG, Fraenkel L, Braithwaite RS. Cost-effectiveness of adding magnetic resonance imaging to rheumatoid arthritis management. *Archives of Internal Medicine.* 2011;171(7):657–667. doi:10.1001/archinternmed.2011.115
12. MacGregor AJ, Snieder H, Rigby AS, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum.* 2000;43(1):30–37. doi:10.1002/1529-0131(200001)43:1<30::AID-ANR5>3.0.CO;2-B
13. 1958 Birth Cohort Controls Jones Richard W. 18 McArdle Wendy L. 18 Ring Susan M. 18 Strachan David P. 19 Pembrey Marcus 18 20, Type 1 Diabetes Clayton David G. 2 Dunger David B. 2 41 Nutland Sarah 2 Stevens Helen E. 2 Walker Neil M. 2 Widmer Barry 2 41 Todd John A. 2. Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature.* 2007;447(7145):661–678.
14. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. an approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum.* 1987;30(11):1205–1213. doi:10.1002/art.1780301102
15. Symmons DPM, Bankhead CR, Harrison BJ, et al. Blood transfusion, smoking, and obesity as risk factors for the development of rheumatoid arthritis. Results from a primary care-based incident case-control study in Norfolk, England. *Arthritis Rheum.* 1997;40(11):1955–1961. doi:10.1002/art.1780401106
16. Klareskog L, Padyukov L, Lorentzen J, Alfredsson L. Mechanisms of Disease: genetic susceptibility and environmental triggers in the development of rheumatoid arthritis. *Nat Clin Pract Rheumatol.* 2006;2(8):425–433. doi:10.1038/nclrheum0249
17. Schröder AE, Greiner A, Seyfert C, Berek C. Differentiation of B cells in the nonlymphoid tissue of the synovial membrane of patients with rheumatoid arthritis. *Proc Natl Acad Sci U S A.* 1996;93(1):221–225. doi:10.1073/pnas.93.1.221
18. Ohata J, Zvaifler NJ, Nishio M, et al. Fibroblast-like synoviocytes of mesenchymal origin express functional B cell-activating factor of the TNF family in response to proinflammatory cytokines. *J Immunol.* 2005;174(2):864–870. doi:10.4049/jimmunol.174.2.864
19. Behrens F, Himsel A, Rehart S, et al. Imbalance in distribution of functional autologous regulatory T cells in rheumatoid arthritis. *Ann Rheum Dis.* 2007;66(9):1151–1156. doi:10.1136/ard.2006.068320
20. Miossec P, Korn T, Kuchroo VK. Interleukin-17 and Type 17 helper T Cells. *N Engl J Med.* 2009;361(9):888–898. doi:10.1056/NEJMra0707449
21. Haringman J, Gerlag D, Zwiderman A, et al. Synovial tissue macrophages: a sensitive biomarker for response to treatment in patients with rheumatoid arthritis. *Ann Rheum Dis.* 2005;64(6):834–838. doi:10.1136/ard.2004.029751
22. Wehr P, Purvis H, Law SC, Thomas R. Dendritic cells, T cells and their interaction in rheumatoid arthritis. *Clin Exp Immunol.* 2019;196(1):12–27. doi:10.1111/cei.13256
23. Lebre MC, Jongbloed SL, Tas SW, Smeets TJM, McInnes IB, Tak PP. Rheumatoid arthritis synovium contains two subsets of CD83–DC–LAMP–dendritic cells with distinct cytokine profiles. *Am J Pathol.* 2008;172(4):940–950. doi:10.2353/ajpath.2008.070703
24. Nigrovic PA, Lee DM. Synovial mast cells: role in acute and chronic arthritis. *Immunol Rev.* 2007;217(1):19–37. doi:10.1111/j.1600-065X.2007.00506.x
25. Cascão R, Rosário HS, Souto-Carneiro MM, Fonseca JE. Neutrophils in rheumatoid arthritis: more than simple final effectors. *Autoimmunity Rev.* 2010;9(8):531–535. doi:10.1016/j.autrev.2009.12.013
26. Long L, Guo H, Chen X, et al. Advancement in understanding the role of ferroptosis in rheumatoid arthritis. *Front Physiol.* 2022;13:1036515. doi:10.3389/fphys.2022.1036515
27. Wu J, Feng Z, Chen L, et al. TNF antagonist sensitizes synovial fibroblasts to ferroptotic cell death in collagen-induced arthritis mouse models. *Nat Commun.* 2022;13(1):676. doi:10.1038/s41467-021-27948-4
28. Zhou Y, Li X, Ng L, et al. Identification of copper death-associated molecular clusters and immunological profiles in rheumatoid arthritis. *Front Immunol.* 2023;14:1103509. doi:10.3389/fimmu.2023.1103509
29. Zhao J, Jiang P, Guo S, Schrodi SJ, He D. Apoptosis, autophagy, netosis, necroptosis, and pyroptosis mediated programmed cell death as targets for innovative therapy in rheumatoid arthritis. *Front Immunol.* 2021;12:809806. doi:10.3389/fimmu.2021.809806
30. Chadha S, Behl T, Bungau S, et al. Mechanistic insights into the role of pyroptosis in rheumatoid arthritis. *Curr Res Translational Med.* 2020;68(4):151–158. doi:10.1016/j.retram.2020.07.003
31. Ling Y, Yang Y, Ren N, et al. Jinwu Jiangu capsule attenuates rheumatoid arthritis via the SLC7A11/GSH/GPX4 pathway in M1 macrophages. *Phytomedicine.* 2024;135:156232. doi:10.1016/j.phymed.2024.156232
32. Liu Y, Liang J, Sha Z, Yang C. Inhibition of oxidative stress-induced ferroptosis can alleviate rheumatoid arthritis in human. ciccacci C, editor. *J Immunol Res.* 2024;2024(1):9943747. doi:10.1155/2024/9943747
33. Liu X, Nie L, Zhang Y, et al. Actin cytoskeleton vulnerability to disulfide stress mediates disulfidptosis. *Nat Cell Biol.* 2023;25(3):404–414. doi:10.1038/s41556-023-01091-2
34. Liu X, Olszewski K, Zhang Y, et al. Cystine transporter regulation of pentose phosphate pathway dependency and disulfide stress exposes a targetable metabolic vulnerability in cancer. *Nat Cell Biol.* 2020;22(4):476–486. doi:10.1038/s41556-020-0496-x

35. Kamanlı A, Nazıroğlu M, Aydılek N, Hacıevliyagil C. Plasma lipid peroxidation and antioxidant levels in patients with rheumatoid arthritis. *Cell Biochem. Funct.* 2004;22(1):53–57. doi:10.1002/cbf.1055
36. Remans PHJ, Gringhuis SI, van Laar JM, et al. Rap1 signaling is required for suppression of ras-generated reactive oxygen species and protection against oxidative stress in T Lymphocytes. *J Immunol.* 2004;173(2):920–931. doi:10.4049/jimmunol.173.2.920
37. Hu Y, Carman JA, Holloway D, et al. Development of a molecular signature to monitor pharmacodynamic responses mediated by in vivo administration of glucocorticoids. *Arthritis Rheumatol.* 2018;70(8):1331–1342. doi:10.1002/art.40476
38. Tasaki S, Suzuki K, Kassai Y, et al. Multi-omics monitoring of drug response in rheumatoid arthritis in pursuit of molecular remission. *Nat Commun.* 2018;9(1):2755. doi:10.1038/s41467-018-05044-4
39. Guo Y, Walsh AM, Fearon U, et al. CD40L-dependent pathway is active at various stages of rheumatoid arthritis disease progression. *J Immunol.* 2017;198(11):4490–4501. doi:10.4049/jimmunol.1601988
40. Walsh AM, Wechalekar MD, Guo Y, et al. Triple DMARD treatment in early rheumatoid arthritis modulates synovial T cell activation and plasmablast/plasma cell differentiation pathways. Kuwana M, editor. *PLoS One.* 2017;12(9):e0183928. doi:10.1371/journal.pone.0183928
41. Idri A, Kadi I, Abnane I, Fernandez-Aleman JL. Missing data techniques in classification for cardiovascular dysautonomias diagnosis. *Med Biol Eng Comput.* 2020;58(11):2863–2878. doi:10.1007/s11517-020-02266-x
42. Li L, Fang H, Li F, et al. Regulation mechanisms of disulfidptosis-related genes in ankylosing spondylitis and inflammatory bowel disease. *Front Immunol.* 2024;15:1326354. doi:10.3389/fimmu.2024.1326354
43. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30. doi:10.1093/nar/28.1.27
44. The Gene Ontology Consortium. The gene ontology resource: 20 years and still going strong. *Nucleic Acids Res.* 2019;47(D1):D330–8. doi:10.1093/nar/gky1055
45. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–287. doi:10.1089/omi.2011.0118
46. Tibshirani R. Regression shrinkage and selection via the lasso. *J Royal Statistical Soc.* 1996;58(1):267–288. doi:10.1111/j.2517-6161.1996.tb02080.x
47. Huang ML, Hung YH, Lee WM, Li RK, Jiang BR. SVM-RFE based feature selection and Taguchi parameters optimization for multiclass SVM classifier. *ScientificWorldJournal.* 2014;2014:795624. doi:10.1155/2014/795624
48. Blanchet L, Vitale R, van Vorstenbosch R, et al. Constructing bi-plots for random forest: tutorial. *Anal. Chim. Acta.* 2020;1131:146–155. doi:10.1016/j.aca.2020.06.043
49. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinf.* 2013;14:14. doi:10.1186/1471-2105-14-14
50. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods mol Biol.* 2018;1711:243–259.
51. Smolen JS, Breedveld FC, Schiff MH, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology.* 2003;42(2):244–257. doi:10.1093/rheumatology/keg072
52. Aletaha D, Nell VPK, Stamm T, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther.* 2005;7(4):R796–806. doi:10.1186/ar1740
53. Felson DT, Anderson JJ, Boers M, et al. American college of rheumatology. preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum.* 1995;38(6):727–735. doi:10.1002/art.1780380602
54. Van Gestel AM, Prevoo MLL, Hof MA V, Van Rijswijk MH, De Putte LBA V, Riel PLCM V. Development and validation of the European league against rheumatism response criteria for rheumatoid arthritis: comparison with the preliminary American college of rheumatology and the world health organization/international league against rheumatism criteria. *Arthritis Rheum.* 1996;39(1):34–40.
55. Aletaha D. Precision medicine and management of rheumatoid arthritis. *J Autoimmun.* 2020;110:102405. doi:10.1016/j.jaut.2020.102405
56. Kay J, Upchurch KS. ACR/EULAR 2010 rheumatoid arthritis classification criteria. *Rheumatology.* 2010;51(suppl\_6):vi5–9.
57. Lindstrom TM, Robinson WH. Biomarkers for rheumatoid arthritis: making it personal. *Scand J Clin Lab Invest Suppl.* 2010;242(sup242):79–84. doi:10.3109/00365513.2010.493406
58. Takahashi D, Mori T, Wakabayashi M, et al. KLHL2 interacts with and ubiquitinates WNK kinases. *Biochem. Biophys. Res. Commun.* 2013;437(3):457–462. doi:10.1016/j.bbrc.2013.06.104
59. Koumangoye R. The role of Cl<sup>-</sup> and K<sup>+</sup> efflux in NLRP3 inflammasome and innate immune response activation. *Am J Physiol Cell Physiol.* 2022;322(4):C645–52. doi:10.1152/ajpcell.00421.2021
60. Hayward DA, Vanes L, Wissmann S, et al. B cell-intrinsic requirement for WNK1 kinase in antibody responses in mice. *J Exp Med.* 2023;220(3):e20211827. doi:10.1084/jem.20211827
61. Yang JM, Ren Y, Kumar A, et al. NAC1 modulates autoimmunity by suppressing regulatory T cell-mediated tolerance. *Sci Adv.* 2022;8(26):eabo0183. doi:10.1126/sciadv.abo0183
62. Reis E, Sousa C, Yamasaki S, Brown GD. Myeloid C-type lectin receptors in innate immune recognition. *Immunity.* 2024;57(4):700–717. doi:10.1016/j.immuni.2024.03.005
63. Miyake Y, Toyonaga K, Mori D, et al. C-type Lectin MCL Is an Fcγ-coupled receptor that mediates the adjuvanticity of mycobacterial cord factor. *Immunity.* 2013;38(5):1050–1062. doi:10.1016/j.immuni.2013.03.010
64. Zhao XQ, Zhu LL, Chang Q, et al. C-type lectin receptor dectin-3 mediates trehalose 6,6'-Dimycolate (TDM)-induced mincle expression through CARD9/Bcl10/MALT1-dependent Nuclear Factor (NF)-κB Activation. *J Biol Chem.* 2014;289(43):30052–30062. doi:10.1074/jbc.M114.588574
65. Du X. Expressions and Roles of CLEC4D and CARD9 in inflammatory bowel disease. *Chin J Gastroenterol.* 2019;144–147.
66. Sweasy JB, Lauper JM, Eckert KA. DNA polymerases and human diseases. *Radiation Res.* 2006;166(5):693–714. doi:10.1667/RR0706.1
67. Hakura A, Sui H, Seki Y, et al. DNA polymerase κ suppresses inflammation and inflammation-induced mutagenesis and carcinogenic potential in the colon of mice. *Genes and Environ.* 2023;45(1):15. doi:10.1186/s41021-023-00272-7
68. Hakura A, Sui H, Sonoda J, Matsuda T, Nohmi T. DNA polymerase kappa counteracts inflammation-induced mutagenesis in multiple organs of mice. *Environ and mol Mutagen.* 2019;60(4):320–330. doi:10.1002/em.22272
69. Shao L. DNA damage response signals transduce stress from rheumatoid arthritis risk factors into T cell dysfunction. *Front Immunol.* 2018;9:3055. doi:10.3389/fimmu.2018.03055

**Journal of Inflammation Research**

**Publish your work in this journal**

The Journal of Inflammation Research is an international, peer-reviewed open-access journal that welcomes laboratory and clinical findings on the molecular basis, cell biology and pharmacology of inflammation including original research, reviews, symposium reports, hypothesis formation and commentaries on: acute/chronic inflammation; mediators of inflammation; cellular processes; molecular mechanisms; pharmacology and novel anti-inflammatory drugs; clinical conditions involving inflammation. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-inflammation-research-journal>

**Dovepress**  
Taylor & Francis Group