Am J Transl Res 2022;14(9):6751-6762
www.ajtr.org /ISSN:1943-8141/AJTR0144790

Original Article

Identification and validation of potential hub genes in rheumatoid arthritis by bioinformatics analysis

Xinling He[1*], Ji Yin[1*], Mingfang Yu[1,2], Jiao Qiu[1], Aiyang Wang[1], Haoyu Wang[1], Xueyi He[1], Xiao Wu[1]

[1]*The Affiliated Traditional Chinese Medicine Hospital of Southwest Medical University, Luzhou, Sichuan, China;* [2]*The Traditional Chinese Medicine Hospital of Luzhou, Luzhou, Sichuan, China.* [*]*Equal contributors.*

Received June 16, 2022; Accepted August 19, 2022; Epub September 15, 2022; Published September 30, 2022

Abstract: Objective: Rheumatoid arthritis (RA) is considered to be a chronic immune disease pathologically characterized by synovial inflammation and bone destruction. At present, the potential pathogenesis of RA is still unclear. Hub genes are recognized to play a pivotal role in the occurrence and progression of RA. Methods: Firstly, we attempted to screen hub genes that are associated with RA, to clarify the underlying pathological mechanisms of RA, and to offer potential treatment methods for RA. We acquired these datasets (GSE12021, GSE55235, and GSE55457) of RA patients and healthy samples from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) were recognized via R software. Then, Gene ontology (GO) functional analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were utilized to deeply explore the underlying biological functions and pathways closely associated with RA. In addition, a protein-protein interaction (PPI) network was built to further evaluate and screen for hub genes. Finally, on the basis of the results of PPI analysis, we confirmed the mRNA expression levels of five hub genes in the synovial tissue of rats modeled with RA. Results: In the human microarray datasets, LCK, JAK2, SOCS3, STAT1, and EGFR were identified as hub genes associated with RA by bioinformatics analysis. Furthermore, we verified the differential expression levels of hub genes in rat synovial tissues via qRT-PCR ($P < 0.05$). Conclusions: Our findings suggest that the hub genes LCK, JAK2, SOCS3, STAT1, and EGFR might have vital roles in the progression of RA and may offer novel therapeutic treatments for RA.

Keywords: Rheumatoid arthritis, bioinformatics, hub gene, protein-protein interaction network, synovial tissue

Introduction

Rheumatoid arthritis (RA) is an autoimmune inflammatory disease characterized by synovial inflammation, and it affects about 1% of the world population [1]. Eventually, RA can result in joint deformity and disability, severely reducing the quality of life for RA patients. Although the etiology of RA is not entirely clarified, environmental and genetic elements have been shown to be implicated in the pathogenesis of RA [2]. At present, medication can ameliorate the clinical symptoms of RA patients, while long-term use has serious side effects [3]. Numerous studies have found that kinase proteins, chemokines, and cytokines might interfere with an organism's immune homeostasis and affect the development of RA [4, 5]. Therefore, identifying the hub genes and pathways of RA might contribute to elucidate the underlying mechanisms and provide novel therapeutic targets for RA.

In recent years, bioinformatics analysis has emerged as an essential approach for investigating etiopathogenesis and screening hub genes for the development and occurrence of disease. In the gene network, hub genes are the genes that interact with many other genes, and commonly play a critical role in biological processes and gene regulation [6]. In addition, hub genes were described as the most closely associated with disease [7]. For example, bioinformatics analysis was used to confirm CXCL9 and CXCL10 as potential hub genes for the recognition of RA and osteoarthritis [8]. A study found that CD53, CD79A, MS4A1, PECAM1 and TAGLN were identified as hub genes associated with the etiology of chronic periodontitis using bioinformatics analysis [9]. Furthermore, FADD, CXCL2, and CXCL8 were considered to be potential hub genes for RA by bioinformatics analytic methods [10]. However, few studies have further validated the results of bioinformatics analysis through experiments. Conse-

quently, there is an urgent need for in-depth analysis and experimental validation of potential hub genes of RA in order to provide effective treatment strategies for RA.

In this study, we concentrated on the changes of genes in the progression of RA to screen hub genes. Firstly, we retrieved the gene expression profiles of RA patients and healthy samples from the Gene Expression Omnibus (GEO) database and identified differentially expressed genes (DEGs) by differential expression analysis. Next, the potential pathways and hub genes relevant to RA progression were explored by Gene Ontology (GO) enrichment analyses, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses, and protein-protein interaction (PPI) network analyses. Lastly, we conducted experimental validation of the screened hub genes using synovial tissue from RA rats, revealing the underlying pathogenesis involved in RA.

Materials and methods

Data preparation

The GEO database (http://www.ncbi.nlm.nih.gov/geo) was performed to download microarray datasets of RA. The GEO database was extracted by the following criteria: "Rheumatoid Arthritis", "Homo sapiens", "Expression profiling by array", "tissue", and "sample count" > 20. The GSE12021 dataset contained 12 RA synovial samples and 9 normal synovial samples [11]. The GSE55235 dataset included 10 RA synovial samples and 10 normal synovial samples [12]. The microarray data of GSE55457 consisted of 13 RA patients' synovial samples and 10 healthy synovial samples [12].

Recognition of differentially expressed genes

The R package limma was employed to identify DEGs [13]. The statistically significant difference of DEGs was defined with adjusted *P* value < 0.05 and $|\log2 FC| \geq 0.585$. Differential expression of DEGs was visualized with the R package ggplot2 [14] and volcano plots were constructed. To further investigate the relationship of DEGs in the GSE12021, GSE55235, and GSE55457 datasets, we respectively performed intersections of the upregulated and downregulated genes in these datasets via the R package venn.

Exploration of functional enrichment analysis

The GO and KEGG enrichment analyses were conducted to investigate the functions and pathways of identified genes using the "clusterProfiler" package ($P < 0.05$) [15].

Construction of protein-protein interaction network

The PPI network was constructed to predict protein-protein interactions of DEGs by the Search Tool for the Retrieval of Interacting Genes database (STRING, http://www.stringdb.org) [16]. The filtering criterion was an interaction score of more than 0.7. Then, we used the CytoNCA 2.1 (http://apps.cytoscape.org/apps/cytonca) [17] from Cytoscape software v3.6.0 [18] to construct the PPI network and analyze the network topological characteristics. The values of Betweenness Centrality (BC), Closeness Centrality (CC), Degree Centrality (DC), Eigenvector Centrality (EC), Local Average Connectivity-based method Centrality (LAC), and Network Centrality (NC) were used as references for the importance of the most significant module. Based on the values of BC, CC, DC, EC, LAC, and NC ranked in the top 50%, the significant modules were filtered. Next, according to the above screening results, we again filtered the most significant modules on the basis of the values of BC, CC, DC, EC, LAC, and NC in the top 50%. In the PPI network, a node represents the protein product of a differentially expressed gene, where Degree Centrality (DC) indicates the number of proteins interacting with the node [19]. A node with high DC is considered a hub node and can be used to assess the importance of a hub gene [20]. The top five genes in the most significant module were selected as the hub genes based on DC value.

Establishment of the RA rat model

A total of 16 Sprague-Dawley rats (160-180 g) were supplied by the Experimental Animal of Southwest Medical University (Luzhou, China), half male and half female. All animal experimental procedures were approved by the Institutional Ethics Committee of South Western Medical University (No: 2018030973). After 1 week of adaptive feeding, rats were randomly divided into the control group and RA group (n = 8). Adjuvant arthritis (AA) was induced by Freund's Complete Adjuvant (FCA, Sigma), whi-

Table 1. Specific primer sequences in PCR

Gene	Sequence (5' to 3')	Size (bp)
LCK	Forward: CTCGTCCGGCTTTATGCAGT	85
	Reverse: GGAAATCTACTAGGCTCCCGTT	
STAT1	Forward: GGCTGGGCTTCTATCCTGTG	134
	Reverse: TTGGTGACTGACGAAAACTGC	
JAK2	Forward: ACCTGTGGAATTTATGCGAA	89
	Reverse: TTCCGTTGTTCTTCAGTAGCTC	
SOCS3	Forward: GTGGCTACCCTCCAGCATCT	103
	Reverse: GGTCCAGGAACTCCCGAAT	
EGFR	Forward: ATTGCCCTGAACACCGTGGA	129
	Reverse: CCTAAGCCCAGTTTTGTTGGTT	
β-actin	Forward: CAGGTCATCACTATCGGCAAT	139
	Reverse: CTTTACGGATGTCAACGTCACAC	

ch consists of 1 mg of heat-killed and dried Mycobacterium tuberculosis (strain H37Ra, ATCC 25177), dissolved in 0.85 ml of paraffin oil and 0.15 ml of mannide monooleate [21]. The clinical manifestations, serological indicators, and pathological changes of the AA rat model are similar to those of human rheumatoid arthritis in many aspects, making it an ideal animal model for studying RA [22, 23]. The RA grouped rats were injected with 0.15 ml FCA at the right foot pad to establish the AA rat model [24], while control group rats received the same volume of physiological saline.

Measurement of foot volume

The right paw volume of each rat was periodically measured on day 0 and day 7 of the experiment via using a self-made foot volume measuring device according to the method used in our previously published work [25]. To reduce errors, each rat was measured three times. To reduce the bias caused by subjective visual readings, each operator was fixed.

Measurement of arthritis score

On day 0, 3, and 7 of the experiment, all rats were measured with the arthritis score. Two independent observers who were unaware of the experimental design assessed these rats for signs of adjuvant arthritis. The arthritis score is classified into five grades, ranging from 0 points to 4 points: 0 = normal; 1 = slight swelling of skin and joints; 2 = moderate or mild redness and swelling of feet, pads, or ankles; 3 = severe and moderate redness and swelling of feet, toes, and joints; and 4 = severely and highly swollen of the whole paw and joints [26].

Collection of synovial tissues

Rats were sacrificed on day 8 of the experiment. All rats were anesthetized by intraperitoneal injection of sodium pentobarbital (40 mg/kg). After the rat's pain response had disappeared, the synovial tissues of rats were immediately resected and placed in liquid nitrogen and stored at -80°C for further testing. After synovial tissue sampling, rats were euthanized via cervical dislocation.

Validation of quantitative PCR analysis

To verify the results of the bioinformatics analysis, we measured the expression levels of the top 5 hub genes in the synovial tissues of RA rats and control rats by qRT-PCR. On the basis of the manufacturer's protocol, the total RNA was extracted from synovial tissue by RNAiso Plus (Takara, China). Next, cDNA was synthesized from total RNA using transcriptor cDNA synthesis kit (Roche, Germany). The relative mRNA levels of LCK, JAK2, SOCS3, STAT1, and EGFR were measured with Stormstar SybrGreen qPCR Master Mix (DBI, Germany). The specific primers were designed by the Primer Premier 5.0 software (Premier Biosoft, USA) based on the gene sequences in the NCBI database. The different transcript values were normalized using β-actin as an internal reference gene. The relative mRNA expression levels were calculated with the $2^{-\Delta\Delta}$ct method. All PCR assays were performed in triplicate. **Table 1** shows the primer sequences used for qRT-PCR.

Statistical analysis

Data were presented as the means ± standard deviation (SD). All statistical analyses were conducted by SPSS 17.0 statistical software. Student's t-test statistical analysis was performed for the data analysis. $P < 0.05$ was considered to indicate a statistically significant difference.

Results

Identification of DEGs

We analyzed the DEGs of GSE12021, GSE552-35, and GSE55457 datasets based on adjust-

ed *P* value < 0.05 and |log2 FC| ≥ 0.585. The characteristics of the three datasets are provided in Supplementary Table 1. In the GSE12021 dataset, we identified 433 DEGs, of which 219 were up-regulated and 214 were down-regulated (**Figure 1A**). A total of 1,958 DEGs were screened from the GSE55235 datasets, including 1,046 upregulated and 912 down-regulated genes (**Figure 1B**). In the GSE55457 data set, 351 DEGs were significantly up-regulated and 281 DEGs down-regulated (**Figure 1C**). Then, we used the R package venn to screen for the overlap of DEGs in the three above datasets. The results demonstrated that there were 56 up-regulated and 140 down-regulated overlapping DEGs in the GSE12021, GSE55235, and GSE55457 datasets (**Figure 1D** and **1E**).

GO and KEGG pathway analysis

To characterize the features of the DEGs, we conducted GO and KEGG enrichment analyses. The GO enrichment analysis demonstrated that DEGs were associated with B cell activation in biological process (BP); DEGs were notably enriched in the external side of the plasma membrane in cell compartment (CC); and DEGs were related to G protein-coupled receptor binding in molecular function (MF) (**Figure 2**). In addition, remarkably enriched KEGG pathways of DEGs mainly included cytokine-cytokine receptor interaction, Th17 cell differentiation, chemokine signaling pathway, TNF signaling pathway, Th1 and Th2 cell differentiation, JAK-STAT signaling pathway, and NF-kappaB signaling pathway [27, 28] (**Figure 3**).

Construction of PPI interaction network and recognition of hub genes

The STRING database was applied to elucidate the relationship of DEGs. In total, 92 nodes and 230 edges were recognized in the PPI network (**Figure 4A**). Six topological features of each node in the network were computed to find the hub genes. Then, 27 nodes with the values of BC ≥ 13.928, CC ≥ 0.072, DC ≥ 4.000, EC ≥ 0.066, LAC ≥ 2.000, and NC ≥ 2.667 were selected as the most significant module, namely, CXCL11, CXCL9, PLCG2, LCK, CD79A, BLNK, RAD51, TRIM22, CCNA2, DDX60, CD19, MAPK8, JUN, EGFR, IGF1R, GBP1, MCM5, IFI27, OAS2, PSMB8, STAT1, EGR1, JAK2, CXCL10, SOCS3, CCL5, and IL2RG. The PPI network of the most significant module was constructed, with 27 nodes and 74 edges (**Figure 4B**). Finally, the top 5 genes were recognized as hub genes with values of BC ≥ 4.752, CC ≥ 0.473, DC ≥ 5.000, EC ≥ 0.168, LAC ≥ 2.857, and NC ≥ 4.000. The hub genes were EGFR, STAT1, JAK2, SOCS3, and LCK. A PPI network for these hub genes was constructed with 7 nodes and 13 edges (**Figure 4C**). Therefore, the hub genes were likely to be the most important genes in the development of RA.

Effect of FCA on foot volume and arthritis score variations in rats

As shown in Supplementary Figure 1, compared with the control group, there was no significant difference in foot volume and arthritis score in the model group (P > 0.05) at day 0, indicating that the two groups were comparable at baseline before modeling. On day 7, the model group of right foot volumes and arthritis scores were significantly increased (P < 0.01) compared to the control group, which suggested that creation of the AA rat model was successful.

Validation of the hub genes

To validate the reliability of the results of bioinformatics analysis, we detected the relative expression levels of LCK, JAK2, SOCS3, STAT1, and EGFR in the synovial tissues of normal rats and RA rats (n = 8) via qRT-PCR. The result showed that compared with the normal rat synovial tissues, the relative expression levels of LCK, JAK2, and STAT1 were significantly increased in RA rats' synovial tissues (P < 0.05), while the relative expression levels of SOCS3 and EGFR were significantly decreased in RA rats' synovial tissues (P < 0.05) (**Figure 5**). The qRT-PCR results supported the bioinformatics analysis findings.

Discussion

RA is an autoimmune disease with complicated etiology characterized by bone destruction, pain, and swelling of synovial joints, eventually leading to physical disability [29]. Studies found that kinase proteins and signaling factors could regulate the transcription and expression of pro-inflammatory cytokines, thereby exacerbating the synovial inflammatory response in RA [30, 31]. Despite numerous stud-

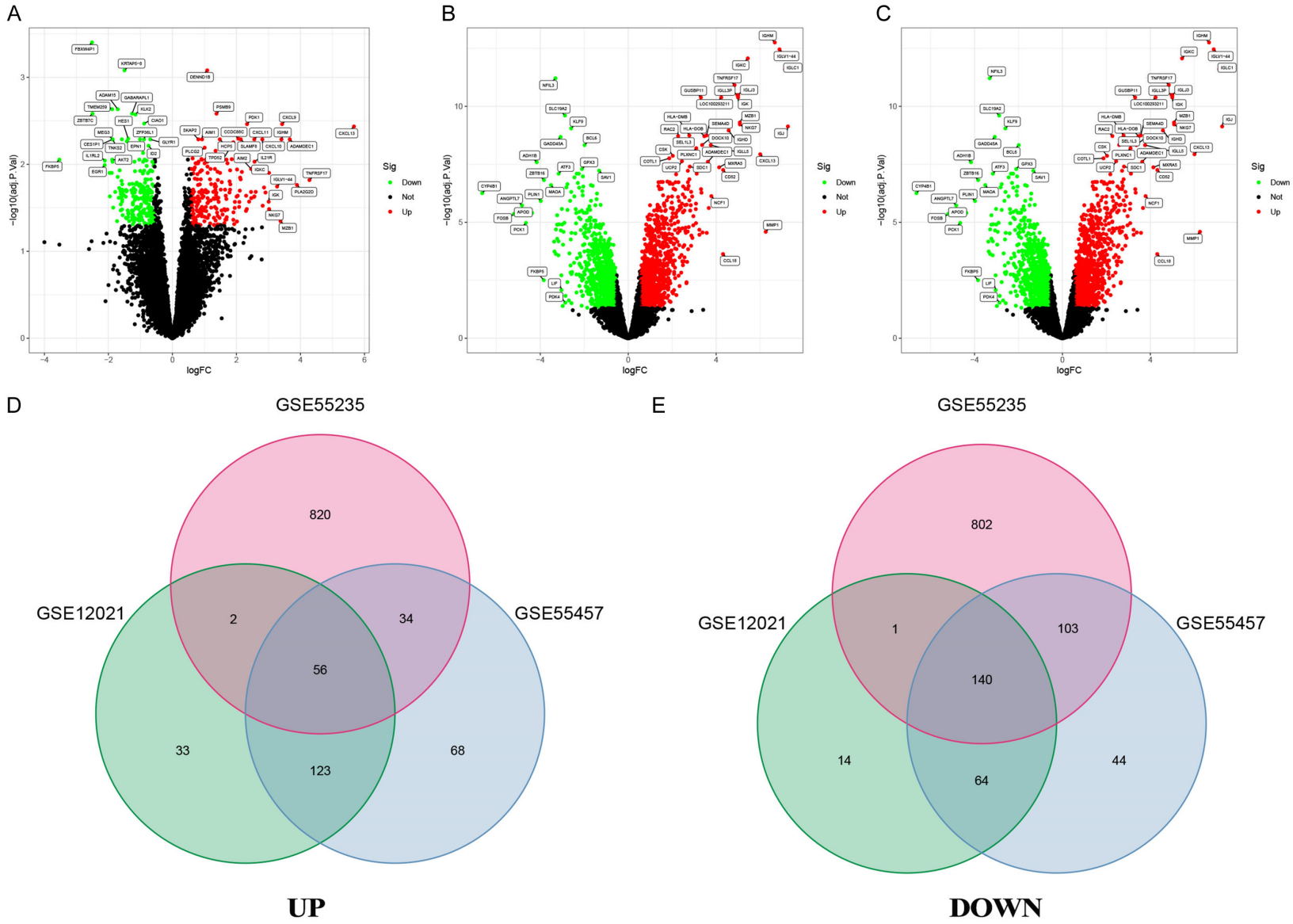


Figure 1. Identification of differentially expressed genes (DEGs). Adjusted *P* value < 0.05 and |log2 FC| ≥ 0.585 were considered as significant difference. The volcano plot shows 219 up-regulated and 214 down-regulated DEGs in GSE12021 (A), 1046 up-regulated and 912 down-regulated DEGs in GSE55235 (B), and 1046 up-regulated and 912 down-regulated DEGs in GSE55235 (C). The venn diagram shows 56 up-regulated overlapping DEGs (D) and 140 down-regulated overlapping DEGs (E).

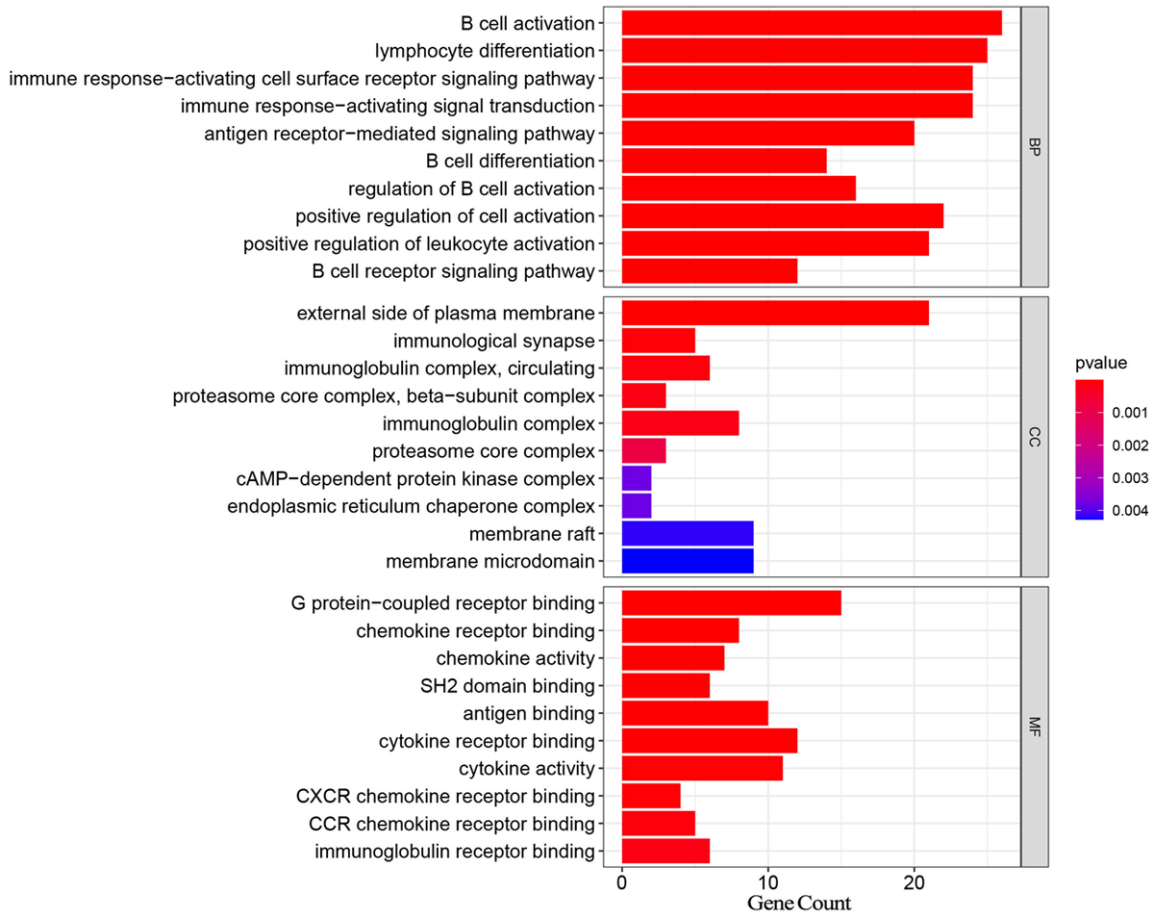


Figure 2. GO functional enrichment of DEGs. The gradually changing color represents the *P*-value ($P < 0.05$). The redder the colour, the stronger the correlation.

ies having explored the formation and development of RA, the underlying pathogenesis is unclear. Consequently, we identified the hub genes and genetic features of RA by bioinformatics analysis to clarify the underlying pathogenesis of RA and offer novel approaches for RA treatment.

Firstly, we conducted differential expression analysis to recognize a total of 196 overlapping RA-relate DEGs (56 up-regulated and 140 down-regulated genes). The GO and KEGG functional enrichment analyses indicated that the DEGs were markedly implicated in cytokine receptor binding, cytokine-cytokine receptor interaction, Th17 cell differentiation, TNF signaling pathway, and JAK-STAT signaling pathway. Previous studies indicated that these potential signaling pathways are not only involved in participating in the inflammatory response and immune reactions, but also the physiological and pathological processes of RA [27, 28, 32, 33]. In addition, we constructed a PPI network for these DEGs on the basis of the STRING database, and five hub genes were recognized, including LCK, JAK2, STAT1, SOCS3, and EGFR. The reliability of the above results was further supported by qRT-PCR.

The signal transducer and activator of transcription 1 (STAT1) is a member of the STAT transcription factor family activated by IFN-γ and IL-6 and it plays a crucial role in the immune response [34]. STAT1 has been found to be associated with numerous autoimmune diseases, such as RA, asthma, and inflammatory bowel disease [35]. Some studies have shown that the expression of STAT1 in RA patients' synovial tissue was notably higher compared to the healthy control [36]. In addition, the protein expression level of STAT1 was higher in collagen-induced arthritis (CIA) rat synovial cells [37]. The Janus kinase 2 (JAK2) is a non-receptor protein tyrosine kinase belongs to the

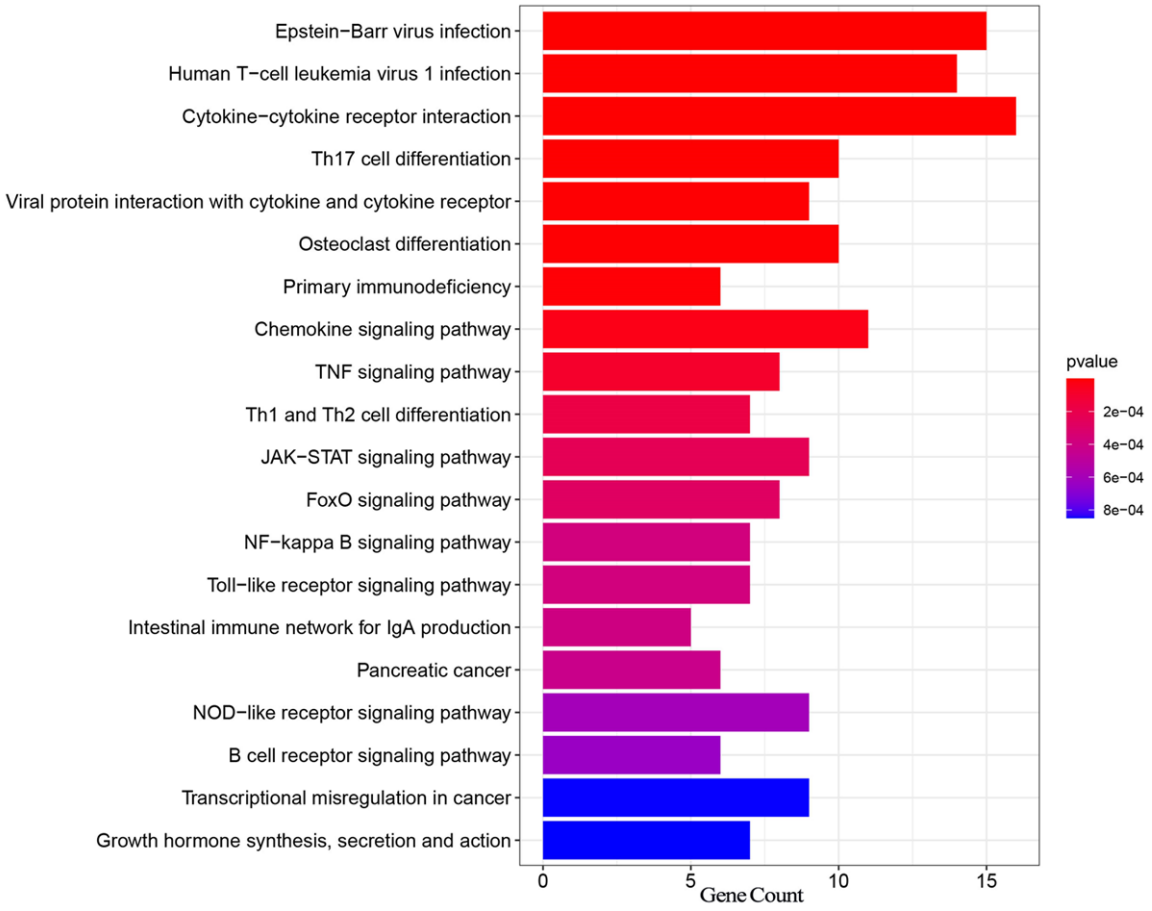


Figure 3. KEGG functional enrichment of DEGs. The gradually changing color represents the *P*-value ($P < 0.05$). The redder the colour, the stronger the correlation.

Janus family of kinases (JAKs). JAK2 can activate many cytokines such as interleukin-6, thereby may be involved in the pathogenesis of RA [38]. A clinical study found that JAK2 inhibitors could significantly relieve pain symptoms in RA patients [39]. A new therapeutic approach for RA has been suggested as JAK2 relates closely to the pathogenesis of RA [40]. The suppression of cytokine signaling 3 (SOCS3) is a negative regulator of cytokine and growth factor signaling, which plays a vital role in cell proliferation, differentiation, and immunity [41]. A study found that local osteoclast production and bone destruction were significantly increased in SOCS3-deficient RA rats [42]. Research showed that inducing high expression of SOCS3 in CIA rats could alleviate synovial inflammation [43]. The transcription and expression levels of SOCS3 were significantly increased in RA rats [44]. Studies have proven that JAK2, SOCS3 and STAT1 were important proteins involved in the JAK/STAT pathway, which play vital roles in the regulation of immunity and inflammation [45]. Besides, KEGG analysis demonstrated that the JAK/STAT pathway was significantly enriched in RA. The JAK/STAT signaling pathway is involved in the signaling of multiple cytokines and has a crucial function in the pathogenesis of RA [46]. In addition, studies have confirmed that joint inflammation and bone destruction in RA rats were dramatically improved via inhibiting the JAK/STAT signaling pathway [47].

Lymphocyte-specific protein tyrosine kinase (LCK) is a member of the Src family of tyrosine kinases that is mainly expressed in T cells and natural killer (NK) cells [48]. Study has confirmed that LCK had remarkably higher expression in synovial tissue of RA patients [49]. LCK inhibitors are thought to be an effective therapeutic strategy in inflammatory immune diseases, including RA [50]. Previous bioinformatics analyses have also shown that LCK might

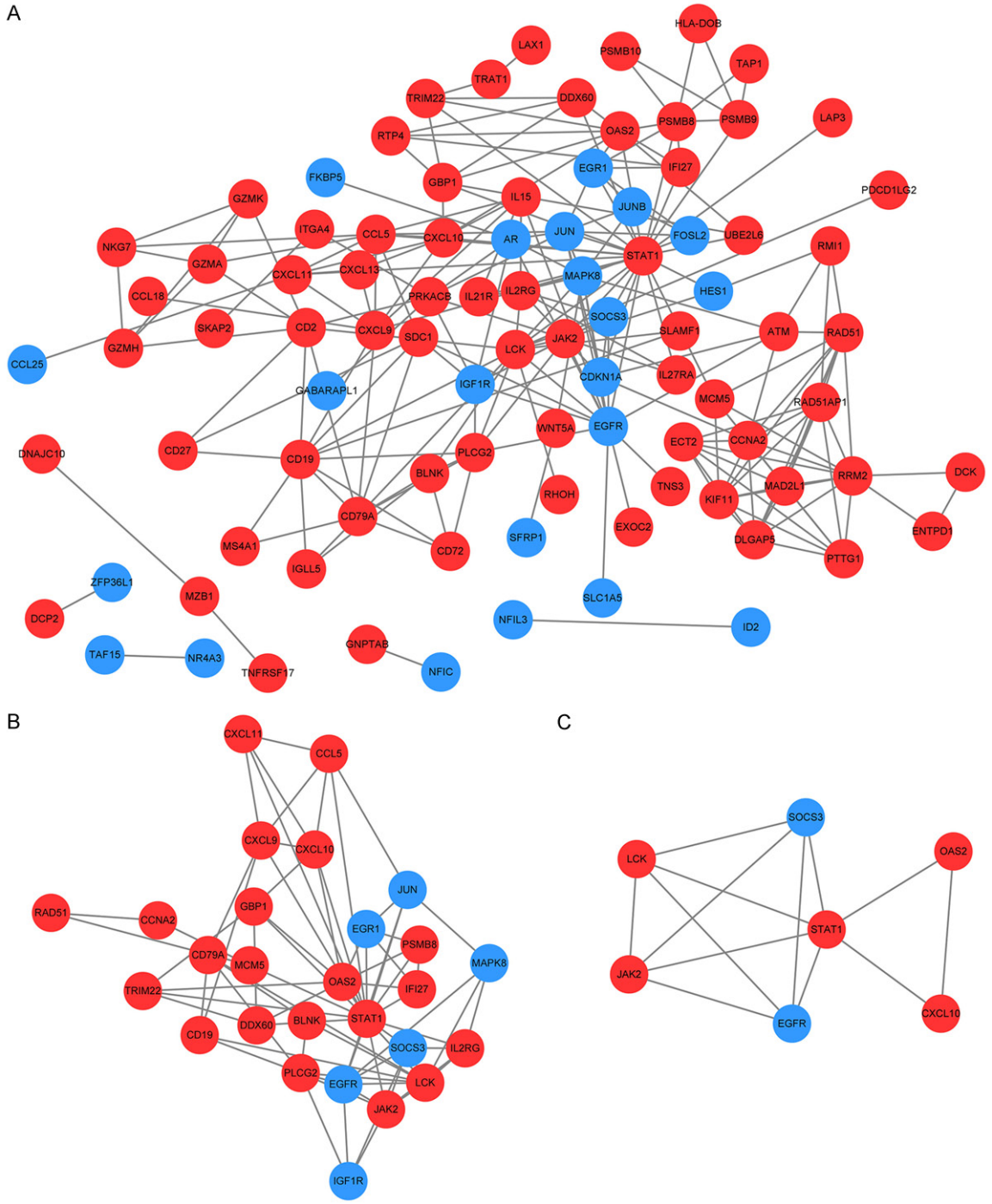


Figure 4. PPI network of DEGs. The edge denotes the interaction between two genes. The degree is utilized to depict the significance of the protein nodes in the network. Node represents gene. Red indicates up-regulated genes and blue indicates down-regulated genes. A. A total of 92 nodes and 230 edges were recognized in the PPI network. B. A total of 27 nodes and 74 edges were recognized in the PPI network. C. A total of 7 nodes and 13 edges were recognized in the PPI network.

be a potential hub gene for RA [31]. The epidermal growth factor receptor (EGFR) is part of the ErbB family of tyrosine kinase receptors, which is a transmembrane glycoprotein [51]. In the pathogenesis of RA, EGFR can facilitate the production of cytokines in synovial fibroblasts and inhibit osteoclast formation, thereby aggravating the severity of RA [52]. Earlier studies

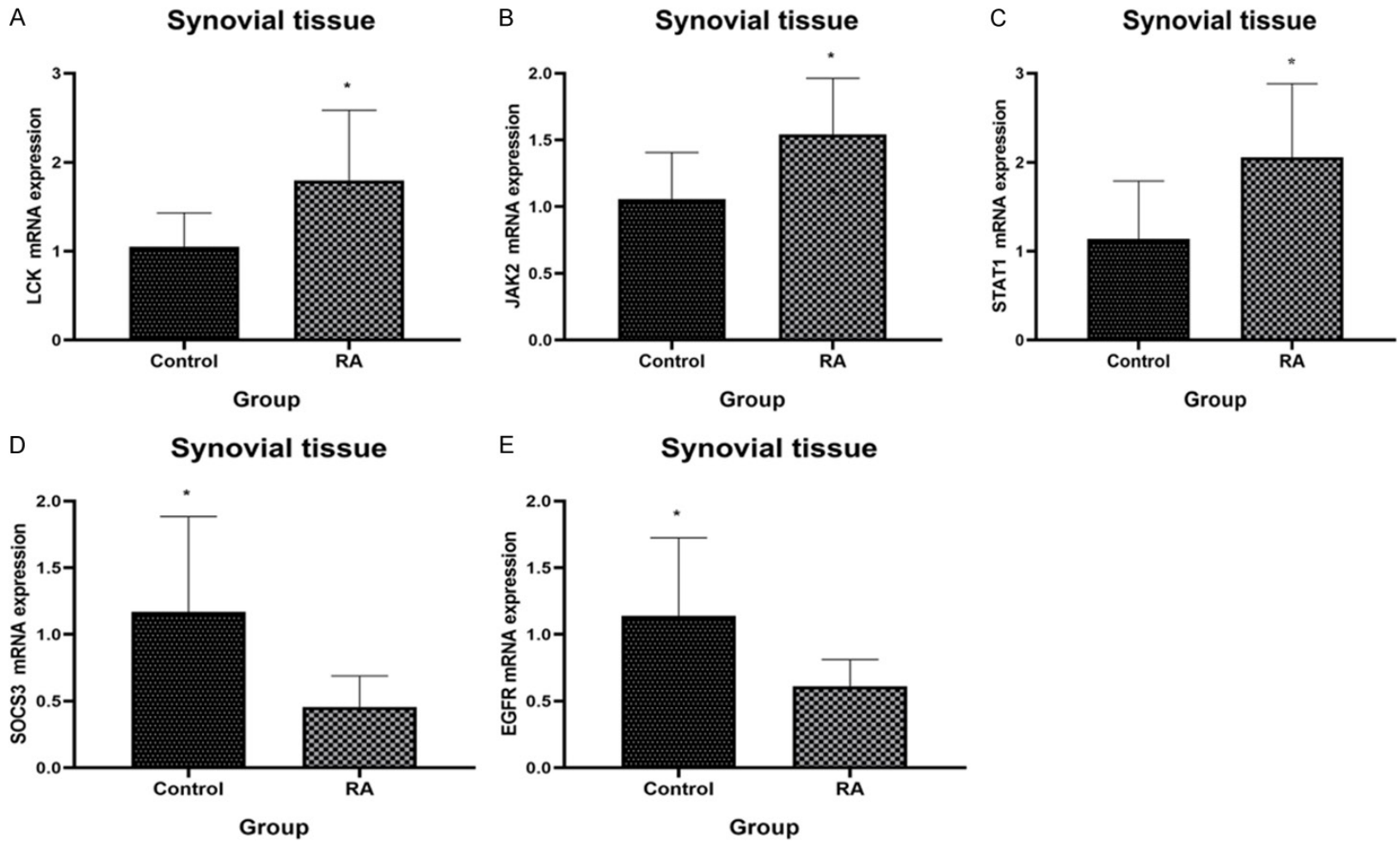


Figure 5. The expression level of hub genes in RA and control rat synovial tissues. A-E. Validation of the top five hub genes by qRT-PCR between the RA group and the control group. * represents $P < 0.05$.

found that an EGFR inhibitor (Herstatin) could significantly improve the clinical symptoms and reduce joint damage in CIA mice [53]. A study demonstrated that an EGFR inhibitor (Erlotinib) could effectively inhibit osteoclast formation and reduce bone loss and bone erosion in CIA mice [52]. A study reported that EGFR was found to be highly expressed in the synovial of CIA mice and RA patients [52]. Besides, EGFR was recognized as a potential therapeutic target for rheumatoid arthritis, which was consistent with our screening findings [54].

However, there are some deficiencies in our study. On the one hand, the limited number of relevant samples obtained from the GEO database may lead to some bias in the bioinformatics analysis. We need to incorporate more samples to further evaluate the reliability of predicted hub genes. On the other hand, due to the difficulty in obtaining human synovial tissue samples, we used synovial tissues from RA rats for experimental validation. Nevertheless, the pathogenesis of RA in humans is complex and diverse, and animal models only partially explain the pathophysiologic manifestations of RA. In the future, we will endeavour to collect human synovial tissue samples to investigate and verify the expression of hub genes in healthy controls and RA patients. Furthermore, the above results need to be further verified in *in vitro* experiments and large-sample clinical trials.

In conclusion, we identified LCK, JAK2, SOCS3, STAT1, and EGFR as hub genes associated with the occurrence and development of RA in human microarray datasets by bioinformatics and validated them in animal experiments, which might offer novel candidate genes for the treatment of RA.

Acknowledgements

The study was supported by the National Natural Science Foundation of China (No. 81804208) and the Sichuan Provincial Administration of Traditional Chinese Medicine (No. 2021MS084).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Xiao Wu, The Affiliated Traditional Chinese Medicine Hospital of Southwest Medical University, No. 182 Chunhui Road, Longmatan District, Luzhou 646000, Sichuan, China. E-mail: 29434230@qq.com

References

[1] Kuchler-Bopp S, Mariotte A, Strub M, Po C, De Cauwer A, Schulz G, Van Bellinghen X, Fioretti F, Clauss F, Georgel P, Benkirane-Jessel N and Bornert F. Temporomandibular joint damage in K/BxN arthritic mice. Int J Oral Sci 2020; 12: 5.

[2] Ansari MA, Nadeem A, Bakheet SA, Attia SM, Shahid M, Alyousef FS, Alswailem MA, Alqinyah M and Ahmad SF. Chemokine receptor 5 antagonism causes reduction in joint inflammation in a collagen-induced arthritis mouse model. Molecules 2021; 26: 1839.

[3] Bindu S, Mazumder S and Bandyopadhyay U. Non-steroidal anti-inflammatory drugs (NSAIDs) and organ damage: a current perspective. Biochem Pharmacol 2020; 180: 114147.

[4] Fang Q, Zhou C and Nandakumar KS. Molecular and cellular pathways contributing to joint damage in rheumatoid arthritis. Mediators Inflamm 2020; 2020: 3830212.

[5] Li XF, Chen X, Bao J, Xu L, Zhang L, Huang C, Meng XM and Li J. PTEN negatively regulates the expression of pro-inflammatory cytokines and chemokines of fibroblast-like synoviocytes in adjuvant-induced arthritis. Artif Cells Nanomed Biotechnol 2019; 47: 3687-3696.

[6] Yu D, Lim J, Wang X, Liang F and Xiao G. Enhanced construction of gene regulatory networks using hub gene information. BMC Bioinformatics 2017; 18: 186.

[7] Li Y, He XN, Li C, Gong L and Liu M. Identification of candidate genes and MicroRNAs for acute myocardial infarction by weighted gene coexpression network analysis. Biomed Res Int 2019; 2019: 5742608.

[8] Zuo B, Zhu J, Xiao F, Wang C, Shen Y and Chen X. Identification of novel biomarkers and candidate small molecule drugs in rheumatoid arthritis and osteoarthritis based on bioinformatics analysis of high-throughput data. Bioscience Rep 2020; 40: BSR20193823.

[9] Suzuki A, Horie T and Numabe Y. Investigation of molecular biomarker candidates for diagnosis and prognosis of chronic periodontitis by bioinformatics analysis of pooled microarray gene expression datasets in Gene Expression Omnibus (GEO). BMC Oral Health 2019; 19: 52.

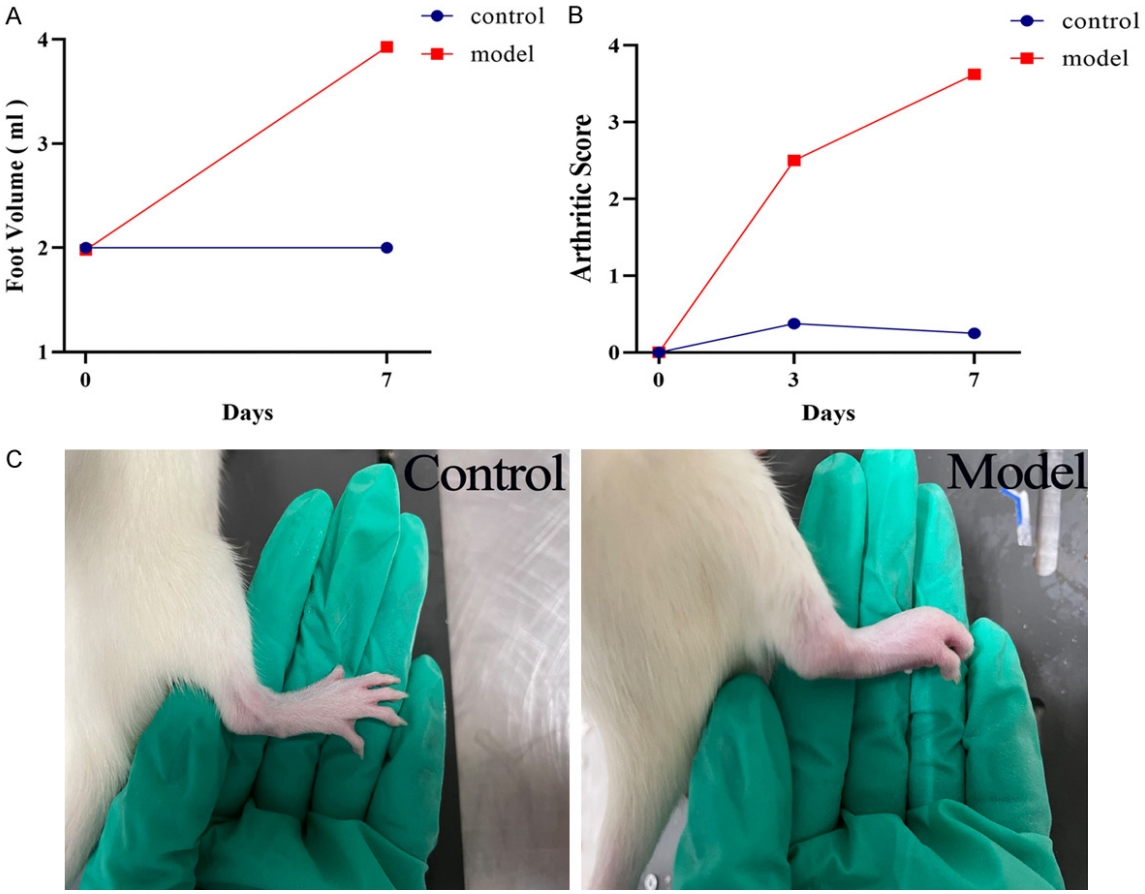
[10] Chen Y, Liao R, Yao Y, Wang Q and Fu L. Machine learning to identify immune-related biomarkers of rheumatoid arthritis based on WGCNA network. Clin Rheumatol 2022; 41: 1057-1068.

[11] Huber R, Hummert C, Gausmann U, Pohlers D, Koczan D, Guthke R and Kinne RW. Identification of intra-group, inter-individual, and gene-specific variances in mRNA expression profiles in the rheumatoid arthritis synovial membrane. Arthritis Res Ther 2008; 10: R98.
[12] Woetzel D, Huber R, Kupfer P, Pohlers D, Pfaff M, Driesch D, Haupl T, Koczan D, Stiehl P, Guthke R and Kinne RW. Identification of rheumatoid arthritis and osteoarthritis patients by transcriptome-based rule set generation. Arthritis Res Ther 2014; 16: R84.
[13] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47.
[14] Wickham H. Ggplot2: elegant graphics for data analysis. New York: Springer; 2009.
[15] Yu G, Wang L, Han Y and He Q. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287.
[16] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ and von Mering C. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015; 43: D447-D452.
[17] Tang Y, Li M, Wang J, Pan Y and Wu FX. CytoNCA: a cytoscape plugin for centrality analysis and evaluation of protein interaction networks. Biosystems 2015; 127: 67-72.
[18] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.
[19] Yu H, Han Z, Xu Z, An C, Xu L and Xin H. RNA sequencing uncovers the key long non-coding RNAs and potential molecular mechanism contributing to XAV939-mediated inhibition of non-small cell lung cancer. Oncol Lett 2019; 17: 4994-5004.
[20] Gu H, Huang Z, Chen G, Zhou K, Zhang Y, Chen J, Xu J and Yin X. Network and pathway-based analyses of genes associated with osteoporosis. Medicine (Baltimore) 2020; 99: e19120.
[21] Mandal G, Chatterjee C and Chatterjee M. Evaluation of anti-inflammatory activity of methanolic extract of leaves of Bougainvillea spectabilis in experimental animal models. Pharmacognosy Res 2015; 7: 18-22.
[22] Kim EY, Durai M, Mia Y, Kim HR and Moudgil KD. Modulation of adjuvant arthritis by cellular and humoral immunity to hsp65. Front Immunol 2016; 7: 203.
[23] Sun S, Li S, Du Y, Wu C, Zhang M, Li J and Zhang X. Anti-inflammatory effects of the root, stem and leaf extracts of Chloranthus serratus on adjuvant-induced arthritis in rats. Pharm Biol 2020; 58: 528-537.
[24] Uroos M, Abbas Z, Sattar S, Umer N, Shabbir A, Shafiq-Ur-Rehman and Sharif A. Nyctanthes arbor-tristis ameliorated FCA-induced experimental arthritis: a comparative study among different extracts. Evid Based Complement Alternat Med 2017; 2017: 4634853.
[25] Wu X, Liu X, Jing Z, Chen Y, Liu H and Ma W. Moxibustion benignantly regulates circadian rhythm of REV-ERBα in RA rats. Am J Transl Res 2020; 12: 1459-1468.
[26] Linghang Q, Yiyi X, Guosheng C, Kang X, Jiyuan T, Xiong L, Guangzhong W, Shuiqing L and Yanju L. Effects of atractylodes oil on inflammatory response and serum metabolites in adjuvant arthritis rats. Biomed Pharmacother 2020; 127: 110130.
[27] Ailioaie LM and Litscher G. Molecular and cellular mechanisms of arthritis in children and adults: new perspectives on applied photobiomodulation. Int J Mol Sci 2020; 21: 6565.
[28] Guo MF, Dai YJ, Gao JR and Chen PJ. Uncovering the mechanism of astragalus membranaceus in the treatment of diabetic nephropathy based on network pharmacology. J Diabetes Res 2020; 2020: 5947304.
[29] Buwembo W, Munabi IG, Kaddumukasa M, Kiryowa H, Mbabali M, Nankya E, Johnson WE, Okello E and Sewankambo NK. Non-surgical oral hygiene interventions on disease activity of rheumatoid arthritis patients with periodontitis: a randomized controlled trial. J Dent Res Dent Clin Dent Prospects 2020; 14: 26-36.
[30] Palmroth M, Kuuliala K, Peltomaa R, Virtanen A, Kuuliala A, Kurttila A, Kinnunen A, Leirisalo-Repo M, Silvennoinen O and Isomäki P. In VivoTofacitinib suppresses several JAK-STAT pathways in rheumatoid arthritis and baseline signaling profile associates with treatment response. Front Immunol 2021; 12: 738481.
[31] Li Z, Xu M, Li R, Zhu Z, Liu Y, Du Z, Zhang G and Song Y. Identification of biomarkers associated with synovitis in rheumatoid arthritis by bioinformatics analyses. Bioscience Rep 2020; 40: BSR20201713.
[32] Elhaj Mahmoud D, Kaabachi W, Sassi N, Mokhtar A, Ben Ammar L, Rekik S, Tarhouni L, Kallel-Sellami M, Cheour E and Laadhar L. Expression of extracellular matrix components and cytokine receptors in human fibrocytes during rheumatoid arthritis. Connect Tissue Res 2021; 62: 720-731.
[33] Yang P, Zhang M, Wang X, Xu AL, Shen M, Jiang B, Zhou X and Zhou L. MicroRNA let-7g-5p alleviates murine collagen-induced arthritis by

inhibiting Th17 cell differentiation. Biochem Pharmacol 2020; 174: 113822.
[34] Xiong J, Zhou H, Lu D, Wang Z, Liu H, Sun Y, Xu J, Feng Y and Xing A. Levetiracetam reduces early inflammatory response after experimental intracerebral hemorrhage by regulating the janus kinase 2 (JAK2)-signal transducer and activator of transcription 3 (STAT3) signaling pathway. Med Sci Monit 2020; 26: e922741.
[35] Yu Y, Xu F, Shen H and Wu J. Chronic Candida infection, bronchiectasis, immunoglobulin abnormalities, and stunting: a case report of a natural mutation of STAT1 (c.986C>G) in an adolescent male. BMC Infect Dis 2021; 21: 38.
[36] Di Benedetto P, Ruscitti P, Berardicurti O, Panzera N, Grazia N, Di Vito NM, Di Francesco B, Navarini L, Maurizi A, Rucci N, Teti AM, Zazzeroni F, Guggino G, Ciccia F, Dolo V, Alesse E, Cipriani P and Giacomelli R. Blocking Jak/STAT signalling using tofacitinib inhibits angiogenesis in experimental arthritis. Arthritis Res Ther 2021; 23: 213.
[37] Yang Y, Dong Q and Li R. Matrine induces the apoptosis of fibroblast-like synoviocytes derived from rats with collagen-induced arthritis by suppressing the activation of the JAK/STAT signaling pathway. Int J Mol Med 2017; 39: 307-316.
[38] Honda S and Harigai M. The safety of baricitinib in patients with rheumatoid arthritis. Expert Opin Drug Saf 2020; 19: 545-551.
[39] Simon LS, Taylor PC, Choy EH, Sebba A, Quebe A, Knopp KL and Porreca F. The Jak/STAT pathway: a focus on pain in rheumatoid arthritis. Semin Arthritis Rheum 2021; 51: 278-284.
[40] Shibata M, Toyoshima J, Kaneko Y, Oda K and Nishimura T. A drug-drug interaction study to evaluate the impact of peficitinib on OCT1- and MATE1-mediated transport of metformin in healthy volunteers. Eur J Clin Pharmacol 2020; 76: 1135-1141.
[41] Seif F, Khoshmirsafa M, Aazami H, Mohsenzadegan M, Sedighi G and Bahar M. The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. Cell Commun Signal 2017; 15: 23.
[42] Wong PK, Egan PJ, Croker BA, O'Donnell K, Sims NA, Drake S, Kiu H, Mcmanus EJ, Alexander WS, Roberts AW and Wicks IP. SOCS-3 negatively regulates innate and adaptive immune mechanisms in acute IL-1-dependent inflammatory arthritis. J Clin Invest 2006; 116: 1571-1581.
[43] Li T, Wu S, Li S, Bai X, Luo H and Zuo X. SOCS3 participates in cholinergic pathway regulation of synovitis in rheumatoid arthritis. Connect Tissue Res 2018; 59: 287-294.
[44] Meng C, Qi H, Wang X, Wu X, Huang W, Wang H and He Y. Expression and methylation levels of suppressor of cytokine signaling 3 in rheumatic arthritis synovial fibroblasts. Exp Mol Pathol 2020; 113: 104361.
[45] Khan MZ, Khan A, Xiao J, Ma Y, Ma J, Gao J and Cao Z. Role of the JAK-STAT pathway in bovine mastitis and milk production. Animals (Basel) 2020; 10: 2107.
[46] Colafrancesco S, Scrivo R, Barbati C, Conti F and Priori R. Targeting the immune system for pulmonary inflammation and cardiovascular complications in COVID-19 patients. Front Immunol 2020; 11: 1439.
[47] Bao Y, Sun YW, Ji J, Gan L, Zhang CF, Wang CZ and Yuan CS. Genkwanin ameliorates adjuvant-induced arthritis in rats through inhibiting JAK/STAT and NF-kappaB signaling pathways. Phytomedicine 2019; 63: 153036.
[48] Ma J, Zhang JK, Yang D and Ma XX. Identification of novel prognosis-related genes in the endometrial cancer immune microenvironment. Aging (Albany NY) 2020; 12: 22152-22173.
[49] Aihaiti Y, Tuerhong X, Ye JT, Ren XY and Xu P. Identification of pivotal genes and pathways in the synovial tissue of patients with rheumatoid arthritis and osteoarthritis through integrated bioinformatic analysis. Mol Med Rep 2020; 22: 3513-3524.
[50] Farag AK, Elkamhawy A, Londhe AM, Lee KT, Pae AN and Roh EJ. Novel LCK/FMS inhibitors based on phenoxypyrimidine scaffold as potential treatment for inflammatory disorders. Eur J Med Chem 2017; 141: 657-675.
[51] Singh D, Attri BK, Gill RK and Bariwal J. Review on EGFR inhibitors: critical updates. Mini Rev Med Chem 2016; 16: 1134-1166.
[52] Swanson CD, Akama-Garren EH, Stein EA, Petralia JD, Ruiz PJ, Edalati A, Lindstrom TM and Robinson WH. Inhibition of epidermal growth factor receptor tyrosine kinase ameliorates collagen-induced arthritis. J Immunol 2012; 188: 3513-3521.
[53] Sumariwalla PF, Jin P, Zhang J, Ni I, Crawford D, Shepard HM, Paleolog EM and Feldmann M. Antagonism of the human epidermal growth factor receptor family controls disease severity in murine collagen-induced arthritis. Arthritis Rheum 2008; 58: 3071-3080.
[54] Yuan FL, Li X, Lu WG, Sun JM, Jiang DL and Xu RS. Epidermal growth factor receptor (EGFR) as a therapeutic target in rheumatoid arthritis. Clin Rheumatol 2013; 32: 289-292.

Supplementary Table 1. The detail information on the gene expression profiles of patients with RA

Datasets	Normal				Rheumatoid arthritis				Platform
	Female	Male	Total	Age (years)	Female	Male	Total	Age (years)	
GSE12021	3	6	9	50.2±21.9	9	3	12	64.8±10.4	GPL96
GSE55235	NA	NA	10	NA	NA	NA	10	NA	GPL96
GSE55457	2	8	10	51.0±19.7	10	3	13	64.6±9.9	GPL96



Supplementary Figure 1. A. The right foot swelling variation of rats. B. The arthritis scores variation of rats. C. Comparison between control group and model group after modeling (day 7).