Identification of diagnostic genes and vital microRNAs involved in rheumatoid arthritis: based on data mining and experimental verification

Conglin Ren[1,*], Mingshuang Li[2,*], Yang Zheng[1], Fengqing Wu[1], Weibin Du[3] and Renfu Quan[1,3]

[1] The Third Clinical Medical College of Zhejiang Chinese Medical University, Hangzhou, Zhejiang, China
[2] The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, Zhejiang, China
[3] Department of Orthopedics, Xiaoshan Traditional Chinese Medicine Hospital, Hangzhou, Zhejiang, China
[*] These authors contributed equally to this work.

ABSTRACT

Background. The pathogenesis of rheumatoid arthritis (RA) is complex. This study aimed to identify diagnostic biomarkers and transcriptional regulators that underlie RA based on bioinformatics analysis and experimental verification.

Material and Methods. We applied weighted gene co-expression network analysis (WGCNA) to analyze dataset GSE55457 and obtained the key module most relevant to the RA phenotype. We then conducted gene function annotation, gene set enrichment analysis (GSEA) and immunocytes quantitative analysis (CIBERSORT). Moreover, the intersection of differentially expressed genes (DEGs) and genes within the key module were entered into the STRING database to construct an interaction network and to mine hub genes. We predicted microRNA (miRNA) using a web-based tool (miRDB). Finally, hub genes and vital miRNAs were validated with independent GEO datasets, RT-qPCR and Western blot.

Results. A total of 367 DEGs were characterized by differential expression analysis. The WGCNA method divided genes into 14 modules, and we focused on the turquoise module containing 845 genes. Gene function annotation and GSEA suggested that immune response and inflammatory signaling pathways are the molecular mechanisms behind RA. Nine hub genes were screened from the network and seven vital regulators were obtained using miRNA prediction. CIBERSORT analysis identified five cell types enriched in RA samples, which were closely related to the expression of hub genes. Through ROC curve and RT-qPCR validation, we confirmed five genes that were specific for RA, including CCL25, CXCL9, CXCL10, CXCL11, and CXCL13. Moreover, we selected a representative gene (CXCL10) for Western blot validation. Vital miRNAs verification showed that only the differences in has-miR-573 and has-miR-34a were statistically significant.

Conclusion. Our study reveals diagnostic genes and vital microRNAs highly related to RA, which could help improve our understanding of the molecular mechanisms underlying the disorder and provide theoretical support for the future exploration of innovative therapeutic approaches.

INTRODUCTION

Rheumatoid arthritis (RA) is a disorder characterized by symmetrical, aggressive joint inflammation (*Smolen, Aletaha & McInnes, 2016*). The prevalence of RA varies somewhat by ethnicity and geographic region. According to the literature, African Americans have a higher prevalence of the disease by approximately 1.02%, compared to Hispanic Americans, who have a much lower risk of this disease at only 0.45% (*Kawatkar et al., 2012*). The etiology of RA is unclear, and current research suggests that it may result from a combination of factors, including genetics, infection, and sex hormones. A study on twins showed that the genetic probability of RA is approximately 60% and is not influenced by sex, age of onset, or disease severity (*MacGregor et al., 2000*). The changes in the composition and function of gut microbes are closely related to infection by autoimmune diseases. The microbial diversity of the gut is significantly reduced in RA patients compared to healthy individuals, and the proportion of certain rare bacteria like Actinobacteria is elevated (*Chen et al., 2016*). More interestingly, oral infections whose clinical manifestations seem unrelated to RA have also been identified as risk factors in recent studies (*Hajishengallis, 2015*; *Kharlamova et al., 2016*). The pathogenesis of RA is complex and clinical symptoms vary considerably between individuals. There have been no reliable biomarkers identified to date, despite the discovery of multiple susceptibility genes associated with RA (e.g., IL2RA and MMP9) and new advances in the development of therapeutic drugs.

microRNA (miRNA) has become a key node in biomedical research as an important factor regulating gene transcription and post-transcriptional regulation. Numerous findings support the idea that miRNAs are essential for maintaining immune system homeostasis, including the regulation of T-cell activation, immune tolerance and the inflammatory response, and the aberrant expression of miRNAs may lead to increased susceptibility to autoimmune diseases (*O'connell et al., 2010*; *Simpson & Ansel, 2015*). Recently, the deregulation of endogenous miR-155 has been shown to be involved in malignant tumorigenesis and in abnormal autoimmune conditions such as RA and systemic sclerosis (*Leng et al., 2011*). Using miRNA profiling, *Pan et al. (2010)* found that miR-21 and miR-148a were highly expressed in T cells from lupus patients and animal models, which inhibited DNMT1 gene expression to induce cellular hypomethylation (*Pan et al., 2010*). Therefore, an in-depth study of the functional links between miRNAs and mRNAs may reveal the pathogenesis of diseases and develop new therapeutic approaches.

With the advent of microarray and high-throughput sequencing technologies, bioinformatics is becoming a popular discipline in biomedical research, and analysis tools are constantly being developed and upgraded. Weighted gene co-expression network analysis (WGCNA) is an algorithm for detecting correlation between genes (*Langfelder & Horvath, 2008*; *Langfelder & Horvath, 2012*), which facilitates network-based genetic screening methods and is helpful to uncover potential biomarkers and therapeutic
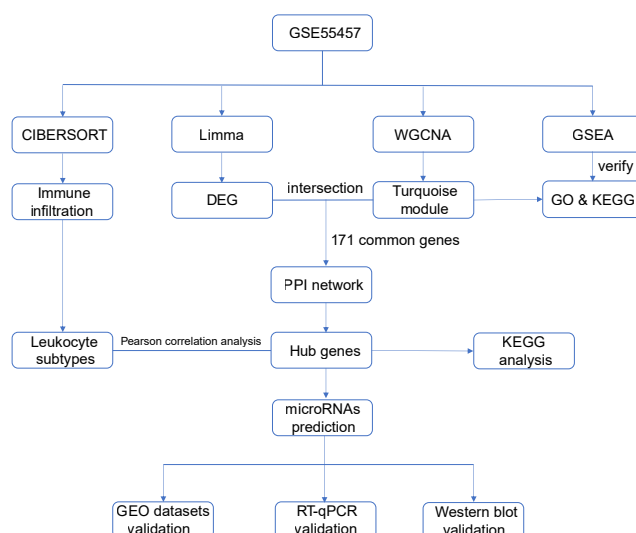
targets for diseases. WGCNA has been successfully practiced in studies of autoimmune diseases, such as juvenile dermatomyositis (*Zhong et al., 2020*), lupus nephritis (*Yang & Li, 2019*), ulcerative colitis (*Zhang et al., 2020*) and Sjögren's syndrome (*Yao et al., 2019*). We integrated multiple bioinformatics methods to analyze microarray datasets derived from the Gene Expression Omnibus (GEO) repository and validated the data mining results in three ways in order to identify diagnostic genes and transcriptional regulators of RA. Figure 1 summarized the entire process of our research.

MATERIAL AND METHODS

Data download and processing

Gene expression matrices and the clinical information of GSE55457 based on platform GPL96 [HG-U133A], were downloaded from the GEO database. Twenty-three samples were selected for our study. We performed normalization and log2 transformation on original expression matrices and used a platform annotation file to convert probes into gene symbols. If a gene was detected by more than one probe, the average expression value was used for subsequent analysis.

Differentially expressed genes (DEGs) between RA patients and healthy controls were analyzed by limma, which is a tailored R package for expression profiling studies (*Ritchie et al., 2015*). The DEGs filtering thresholds were set to |logFC|>1 and the adjusted *p*-value was < 0.05.

Weighted correlation network analysis

Gene co-expression analysis of dataset GSE55457 was performed in R software (version: 3.6.3) based on the WGCNA R package reference manual. We calculated the median absolute deviation (MAD) for each gene, which reflected the degree of variation in gene expression across samples, and selected the top 3,000 highly variable genes for network

construction and modules detection. First, we clustered samples to detect the presence of outliers. Next, optimal soft-thresholding power (β) was selected, which was a pivotal step in the entire analysis process. In order to reduce interference from spurious correlations, the adjacency matrix was replaced by a topological overlap matrix (TOM), and then the hclust function was used to generate a hierarchical clustering tree (dendrogram) of genes. The detailed parameters used for modules detection were as follows: minModuleSize was set to 30 to obtain large modules; modules with a correlation of 0.75 were merged using mergeCloseModules function. We also quantified module–trait associations and their p-values, which allowed us to pick out modules that were closely related to the trait of interest for next step analysis. The absolute value of correlation coefficient > 0.6 and p-value < 0.05 were cutoffs for module screening. Finally, we measured the gene significance (GS) and module membership (MM) of individual genes and scatterplots were drawn.

Functional annotation and GSEA

R package clusterProfiler (*Yu et al., 2012*) and its dependency annotation package org.Hs.eg.db, sourced from Bioconductor, were used to analyze and visualize the functional profiles (GO and KEGG) of key module genes. To obtain highly credible analysis results, only enrichment items satisfying both p-value < 0.05 and q-value < 0.05 were outputted. Moreover, gene set enrichment analysis (GSEA) (*Mootha et al., 2003*; *Subramanian et al., 2005*), which can be used to determine whether an already-defined gene set exhibits a statistically significant difference in two different traits (e.g., RA patients and healthy individuals), was conducted to explore biological processes and pathways that may be involved in the pathogenesis of RA. The complete gene expression matrix and phenotype information of dataset GSE55457 were uploaded to GSEA and analyzed with hallmark gene sets that are stored in molecular signatures database (MSigDB) (*Liberzon et al., 2015*; *Liberzon et al., 2011*). Gene sets enriched at FDR < 0.25 and p-value < 0.01 were retained.

Building PPI network and mining hub genes

The intersection of DEGs and key module genes was entered into the STRING database to find interrelationships (*Szklarczyk et al., 2015*; *Szklarczyk et al., 2019*; *Szklarczyk et al., 2017*). Due to the large size of the initial protein-protein interaction (PPI) network, we hid isolated nodes and limited the minimum interaction score to 0.7. Then, the network was exported and further optimized in Cytoscape software (*Cline et al., 2007*; *Otasek et al., 2019*; *Shannon et al., 2003*). The MCODE plugin was installed in Cytoscape to find clusters made up of highly related genes in the network (*Bader & Hogue, 2003*). The parameters used in our analysis were default values. Genes in the highest scoring cluster were considered to be hub genes of RA and were re-performed for KEGG analysis to explore vital biological pathways.

Estimating proportion of immune cells in synovial tissue

Understanding the immune microenvironment in RA may provide new insights into the immunotherapy of the disease. The CIBERSORT tool used for estimating the immune cell contents of different samples in a gene expression admixture has a wide range of applicability and high sensitivity, which is important for disease diagnosis and therapeutic

target predictions (*Newman et al., 2015*). A mixture file was generated according to the input format requirements, and we selected default LM22 as signature gene file to distinguish the hematopoietic populations and activation states. The analysis results were filtered based on a threshold of *p*-value < 0.05 to remove samples with a low confidence level. We conducted correlation analysis between the hub genes and immune cell subtypes significantly enriched in RA samples according to Pearson's algorithm.

Prediction of gene-miRNA relationships

MiRNAs are a class of non-coding RNAs that regulate gene expression by making mRNA unstable and inhibiting translation, a process by which mRNA is converted into protein. Here, MCODE-identified hub genes were uploaded to the online analysis tool miRDB; we searched for miRNAs that have regulatory roles in the development of RA (*Chen & Wang, 2020*; *Liu & Wang, 2019*). The parameters were set as follows: only functional miRNAs were considered; gene targets with a score below 60 and miRNAs with more than 2,000 targets were excluded. Cytoscape was then used to construct a gene-miRNA network in which miRNAs targeting two or more genes were considered to be vital regulators involved in RA pathogenesis.

Validating hub genes and miRNAs with independent microarray datasets

The expression matrix GSE12021 was selected as the validation dataset for hub genes. We plotted ROC curves in R software based on the pROC package to test whether screened hub genes have good diagnostic value for RA (*Robin et al., 2011*). The area under the curve (AUC) was computed according to the trapezoidal rule and genes with AUC > 85% were further validated by RT-qPCR. To understand the differences in transcriptional regulators between RA patients and healthy individuals, we validated vital miRNAs with GEO dataset GSE37425, which focused on the expression of miRNAs in synovial tissue using the TaqMan probe approach. The entire gene expression matrix was normalized and then transformed with log2.

Validating hub genes and miRNAs with RT-qPCR

Five RA patients and three healthy individuals who attended the health examination center of Xiaoshan Traditional Chinese Medicine Hospital were included and all subjects signed consent forms. This study was approved by Ethics Committees of Xiaoshan Traditional Chinese Medicine Hospital (NO. 2020012). We collected blood samples from each test subject and isolated the peripheral blood mononuclear cells (PBMCs) using human lymphocyte separation medium (Solarbio, China). Next, the total RNA of PBMCs was extracted using the TRIzol reagent (Invitrogen, USA) and its purity was assessed by NanoDrop (Thermo Scientific, USA). We used a cDNA synthesis kit (Thermo Scientific, USA) to generate the first strand cDNA and cDNA amplification was performed on a 7500 real-time PCR instrument (Applied Biosystems, USA). For the quantitative analysis of miRNA expression, Hairpin-it miRNAs qPCR quantitation kit (Genepharma, China) was applied. We chose GAPDH and U6 as internal reference genes and compared relative

Table 1 RT-qPCR primers used in this study.

Gene symbol	Sequence (5′ to 3′)
CCL25	forward: TATTCTACCTCCCCAAGAGACA reverse: GATGGGATTGCTAAACTTGGAC
CXCL10	forward: CTCTCTCTAGAACTGTACGCTG reverse: ATTCAGACATCTCTTCTCACCC
CCL5	forward: CAGCAGTCGTCCACAGGTCAAG reverse: TTTCTTCTCTGGGTTGGCACACAC
CXCL9	forward: AAGACCTTAAACAATTTGCCCC reverse: TGCTGAATCTGGGTTTAGACAT
CXCL13	forward: CAAGGTGTTCTGGAGGTCTATT reverse: TGAATTCGATCAATGAAGCGTC
CXCL11	forward: GCTGTGATATTGTGTGCTACAG reverse: TTGGGTACATTATGGAGGCTTT
GAPDH	forward: CGGACCAATACGACCAAATCCG reverse: AGCCACATCGCTCAGACACC
miR-34a-5p	forward: TCTGTCTCTCTTGGCAGTGTCTTA reverse: AATGGTTGTTCTCCACTCTCTCTC
U6	forward: CAGCACATATACTAAAATTGGAACG reverse: ACGAATTTGCGTGTCATCC

mRNA or miRNA expression based on the $2^{-\Delta\Delta CT}$ method. The primer sequences are available in Table 1.

Western blot

Total proteins were extracted from PBMCs using RIPA lysis buffer (Beyotime Biotechnology, China), followed by concentration detection with the BCA protein assay kit (Sangon Biotech, China). Equal amounts of protein were separated by 8–20% SDS-PAGE and transferred onto polyvinylidene difluoride membranes. Afterwards, membranes were blocked with 5% skim milk for 1 h and then incubated with primary antibodies including β-actin (1:1000, Cell Signaling Technology (CST), #3700), and CXCL10 (1:1000, CST, #14969) at 4 °C overnight. Anti-mouse IgG (1:2000, CST, #7076) and anti-rabbit IgG (1:2000, CST, #7074) were used to bind the primary antibodies. Finally, membranes were scanned using FluorChem Q (Proteinsimple, USA). Image J software was used for the quantitative analysis of protein bands.

Statistical analysis

Statistical analysis was performed in GraphPad Prism 7.0. Data satisfying normal distribution were analyzed by unpaired t-test; the Mann–Whitney test was used for data that were not normalized. P-value < 0.05 was considered significant.

RESULTS

Screening for DEGs

Our differential expression assessment of dataset GSE55457 identified 367 DEGs, of which 186 genes were highly expressed and 181 genes had low expression levels in RA patients. We created a volcano map and heat map for DEGs as shown in Fig. 2.
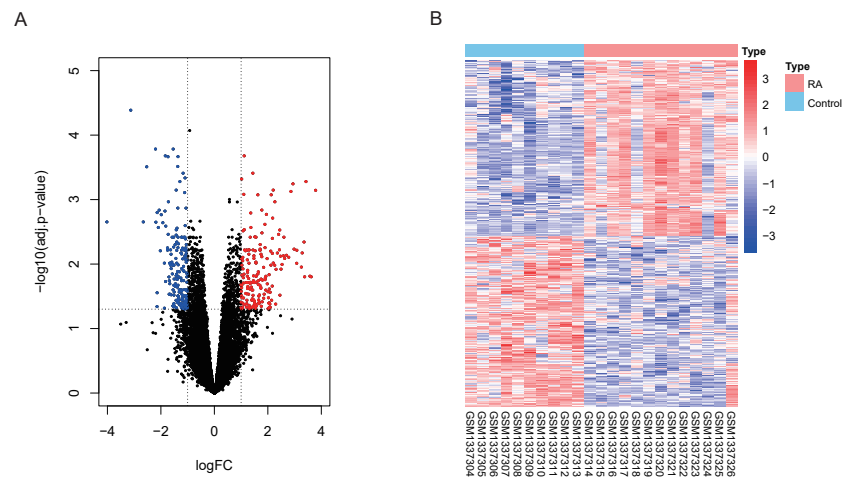
A

B

Figure 2 Visualization of differentially expressed genes (DEGs). (A) DEGs filtered by thresholds were presented in volcano map. Red dots represent up-regulated genes and blue dots represent down-regulated genes. (B) Heatmap showed the expression of DEGs in each sample.

Full-size ◩ DOI: 10.7717/peerj.11427/fig-2

Identification of key module closely related to RA

Two samples (GSM1337307 and GSM1337308) were identified as outliers and were excluded by the cutreeStatic function with parameter cutHeight = 90 (Fig. 3A). The soft-thresholding power (β) was set to 5, for which the scale-free topology fit index (R^2) reached 0.90 (Fig. 3B). Co-expression analysis resulted in 14 modules, and the size of each module is shown in Figs. 3C and 3D. Figure 3E showed the relationships between modules and traits, with four modules (black, turquoise, blue, and purple) having strong positive or negative correlations with disease status (presence or absence of RA). Taking into account gene numbers, correlation coefficients, and p-values of the above four modules, we focused on the turquoise module containing 845 genes and performed an in-depth analysis. The GS and MM plot of the turquoise module (correlation coefficient = 0.79, p-value = 2.7e−181) demonstrated that the closer the gene is to the trait, the more important it is in the module (Fig. 3F).

Functional annotation of turquoise module genes and GSEA

The top 10 results for GO enrichment are shown in Fig. 4A and mainly involved leukocyte activation and cell adhesion. To associate a specific gene with its assigned GO terms, we adopted the GOChord function to generate a circle diagram (Fig. 4B). Figure 4C pointed out that inflammation-related signaling pathways, such as chemokine signaling pathway, NF-kappa B signaling pathway, Th17 cell differentiation and T cell receptor signaling pathway, were significantly linked to the pathogenesis of RA. Similarly, a circle plot was displayed to visualize the correspondence between genes and pathways (Fig. 4D). We identified five gene sets based on a reference database derived from MSigDB using the GSEA method. These were associated with the immune response or inflammatory pathways that were highly enriched in RA, which verified the GO and KEGG results (Figs. 5A–5E).
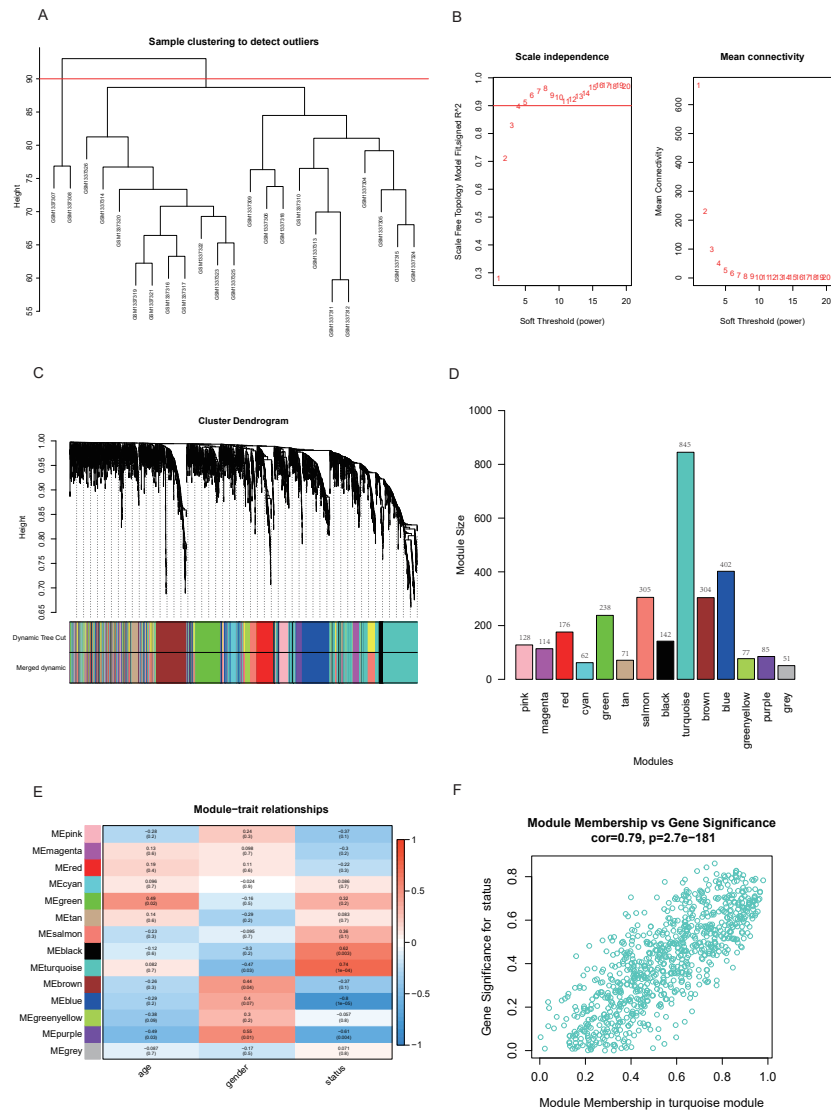
Figure 3 Identification of key module closely related to RA. (A) Clustering of samples and removal of outliers. (B) $\beta = 5$ was chosen as appropriate soft-thresholding power, for which scale-free topology fit index (R^2) reached 0.90. (C) Co-expression analysis resulted in 14 modules. Each branch represents one gene, and every color represents a co-expression module. (D) Number of genes in each co-expression module. (E) Heatmap showed the correlation between module eigengenes and clinical traits. (F) Gene significance (GS) and module membership (MM) of each gene in turquoise module.

Full-size 🖼 DOI: 10.7717/peerj.11427/fig-3

Constructing PPI network to find hub genes

A total of 171 intersection genes were entered into the STRING database to establish the network (Fig. 6A). The network was adjusted using Cytoscape software, resulting in 52 nodes and 169 relationship pairs. The MCODE plugin uncovered five clusters, with the highest scoring cluster containing nine nodes, which were potential hub genes for RA (Fig. 6B, Table 2). Furthermore, we re-performed KEGG analysis on the nine genes to investigate the signaling pathways in which they were involved. As shown in Fig. 6C, we
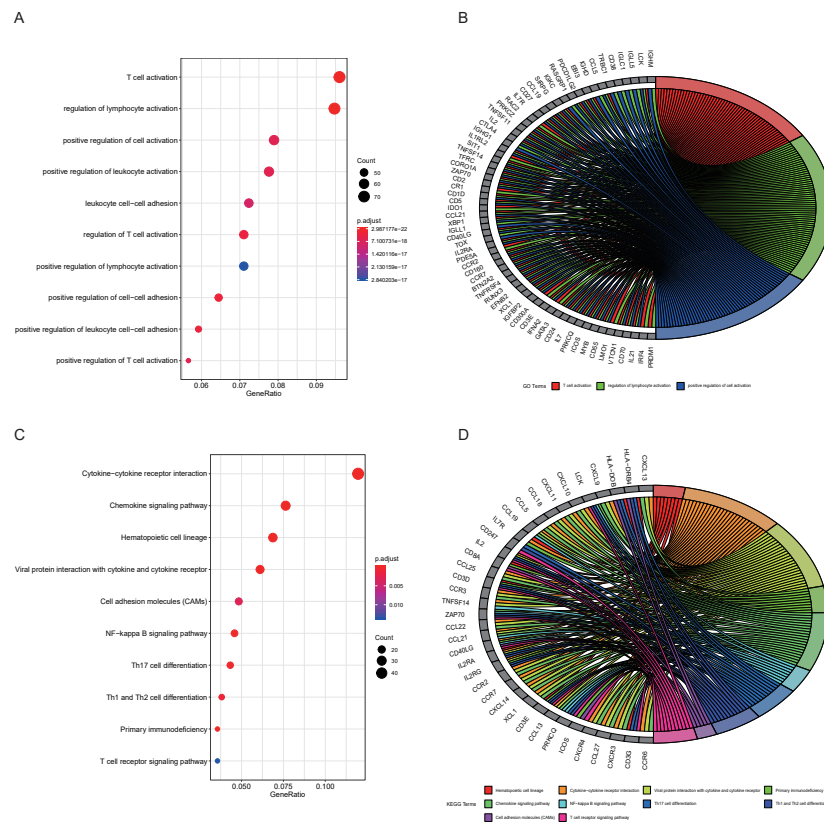
Figure 4 **Results of functional enrichment analysis.** (A) GO analysis of turquoise module genes. (B) Circle plot was displayed to visualize correspondence between genes and top three GO terms. (C) KEGG analysis of turquoise module genes. (D) Circle plot was displayed to visualize correspondence between genes and top 10 KEGG pathways.

Full-size 🖼 DOI: 10.7717/peerj.11427/fig-4

Table 2 Detailed information of each cluster.

Cluster	Score	Nodes	Edges	Node IDs
1	9.000	9	36	CCL25, CCL5, CXCL10, CXCL11, CXCL13, CXCL9, CXCR6, GPR18, PNOC
2	6.286	8	22	BLNK, CD19, CD72, CD79A, CD79B, IGLL5, ITK, PLCG2
3	4.400	6	11	CD8A, MAP4K1, CD247, TRAT1, CD3D, GZMA
4	4.000	4	6	BIRC3, PSMB10, PSMB9, TNFSF11
5	3.000	3	3	CD2, CD48, CD52

found that nine hub genes played important roles in inflammatory or immune-related pathways, including the chemokine signaling pathway, toll-like receptor signaling pathway, and TNF signaling pathway.

Immune infiltration and Pearson correlation analysis

One sample (GSM1337307) was excluded from analysis results because its *p*-value did not meet the threshold we set. The contents of various immune cells in each sample were
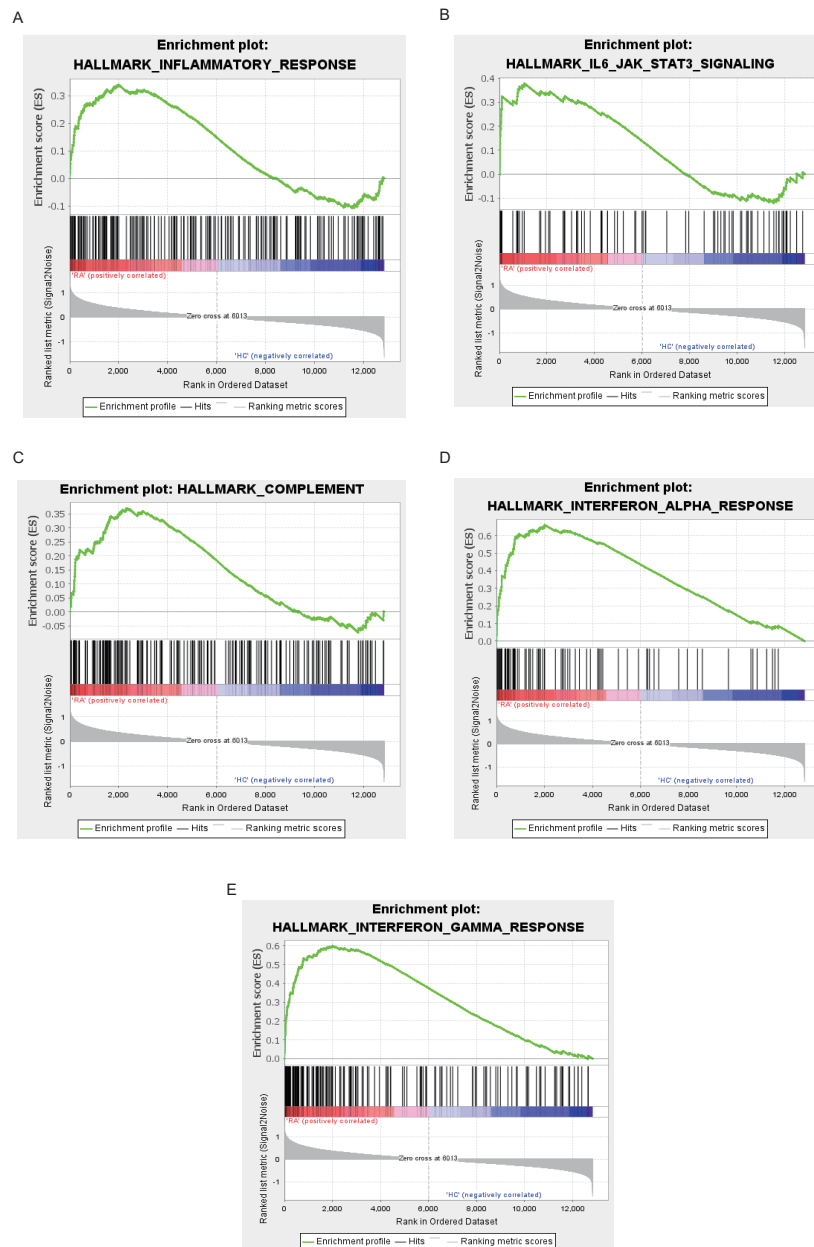
Figure 5 **Results of GSEA analysis.** Enrichment plots for inflammatory response (A), IL6-JAK-STAT3 signaling (B), complement (C), interferon alpha response (D), interferon gamma response (E).

Full-size ⬛ DOI: 10.7717/peerj.11427/fig-5

presented by stacked bar chart (Fig. 7A). M2 macrophages accounted for the highest proportion in both RA and normal tissue. We conducted a Wilcoxon test to further clarify whether the level of each immune cell type was statistically different between two groups. From the violin plot (Fig. 7B), we inferred that five cell types, including plasma cells, CD8 T cells, follicular helper T cells, $\gamma\delta$ T cells, and M1 macrophages, were abundant in RA synovial tissue compared to the normal group. More importantly,
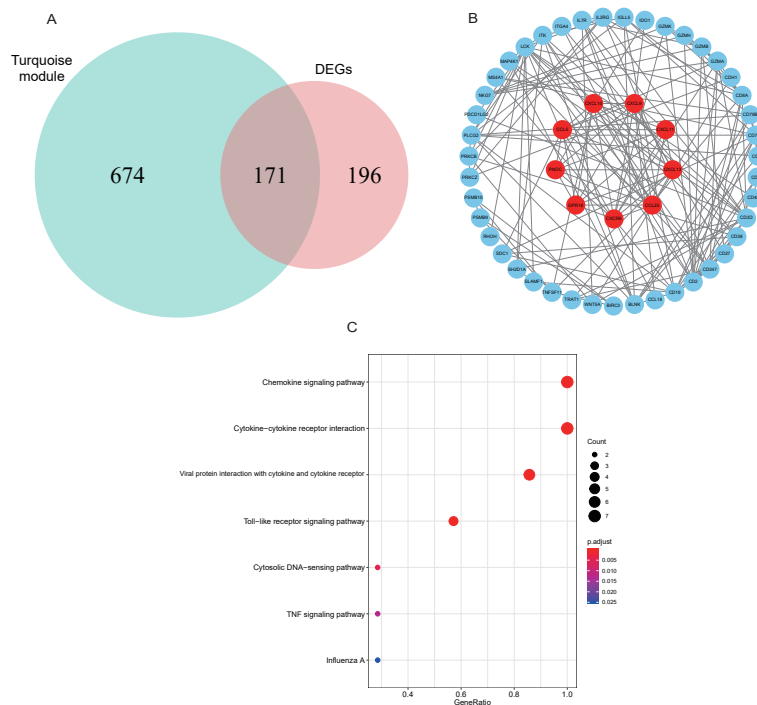
Figure 6 **PPI network construction and mining hub genes.** (A) The intersection of DEGs and turquoise module genes was presented in Venn diagram. (B) The PPI network of intersection genes was constructed in Cytoscape. Genes in the highest scoring cluster were marked in red. (C) KEGG analysis of nine hub genes.

Full-size ☒ DOI: 10.7717/peerj.11427/fig-6

Pearson correlation analysis confirmed that, in addition to CCL25, eight other hub genes identified by MCODE played pro-inflammatory roles. Their overexpression may recruit large numbers of inflammatory cells to lesion sites, which is a critical pathological basis for RA (Fig. 7C).

Predicting miRNAs that regulate nine hub genes

Normally, miRNAs can fine-tune gene expression to maintain the homeostasis of life, while abnormal miRNAs expression is thought to be associated with diseases, including benign or malignant tumors, and immune system dysfunction. Our prediction of miRNAs was made in the miRDB database and 72 miRNA-gene regulatory relationships were obtained. We then created an interactive network in Cytoscape and analyzed the number of genes targeted by each miRNA (Fig. 8). The miRNAs possessing two target genes are listed in Table 3 and were vital regulators in the pathogenesis of RA.

Validation of hub genes by ROC curves

The expression matrix of dataset GSE12021 was chosen to test the diagnostic efficacy of hub biomarkers for RA. Probability curves (ROC curves) were plotted and the AUC reflected the degree or measure of separability. A higher AUC indicated that the gene is better at discriminating between diseased and non-diseased individuals. As shown in Fig. 9, six genes had AUC values greater than 85%, including CCL25, CXCL10, CCL5, CXCL9, CXCL13,
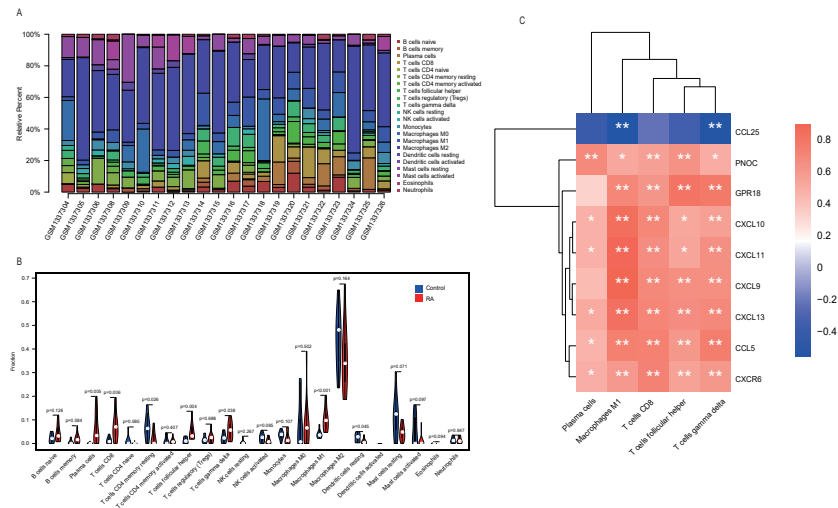
Figure 7   Immune infiltration analysis performed by CIBERSORT. (A) Stacked bar chart showed contents of various immune cells in each sample. (B) The contents of 22 types of immune cells in normal (blue color) and RA (red color) groups were compared. *P*-value < 0.05 was considered statistically significant. (C) Heat map of correlation between hub genes and leukocyte subtypes. The shades of color represent the magnitude of correlation coefficient. * *p*-value < 0.05; ** *p*-value < 0.01.

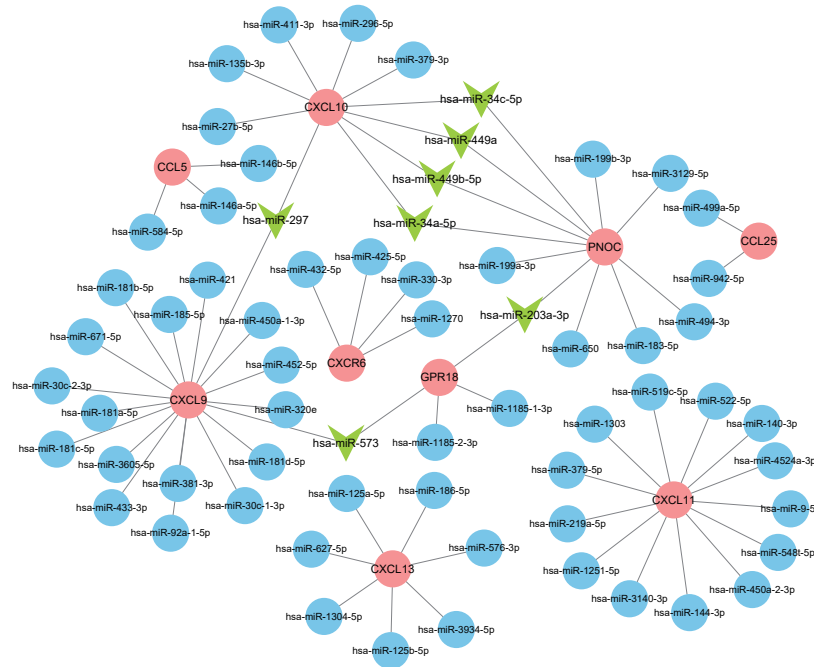Full-size 🖼 DOI: 10.7717/peerj.11427/fig-7

Figure 8   Interaction network between hub genes and miRNAs. Hub genes were colored in red; miRNAs with one target gene were indicated by blue nodes; core miRNAs targeting two genes were indicated by green arrows.

Full-size 🖼 DOI: 10.7717/peerj.11427/fig-8

Table 3 Vital miRNAs and its target genes.

miRNA	Target genes	Count
hsa-miR-297	CXCL9, CXCL10	2
hsa-miR-573	CXCL9, GPR18	2
hsa-miR-203a-3p	GPR18, PNOC	2
hsa-miR-449b-5p	PNOC, CXCL10	2
hsa-miR-34c-5p	PNOC, CXCL10	2
hsa-miR-449a	PNOC, CXCL10	2
hsa-miR-34a-5p	PNOC, CXCL10	2

and CXCL11, indicating that these genes were potential biomarkers for the diagnosis of RA.

Validation of hub genes by RT-qPCR and Western blot

We collected peripheral blood from subjects for RT-qPCR validation to further evaluate the six genes with AUC > 85% identified by ROC curves. Five genes were differentially expressed between RA patients and normal subjects (Fig. 10A). Among them, CXCL9, CXCL10, CXCL11, and CXCL13 were up-regulated genes, while CCL25 was down-regulated. Moreover, to detect the expression of genes in protein level, we chose a representative gene (CXCL10) for Western blot validation. As shown in Figs. 10B and 10C, the protein and mRNA level of gene expression were consistent.

Vital miRNAs verification

The vital miRNAs possessing two target genes were validated with GSE37425. Among the six miRNAs investigated, has-miR-449b, has-miR-573, and has-miR-203 expression was suppressed in RA patients (Figs. 11A–11F). It is worth noting that only the differences in has-miR-573 and has-miR-34a were statistically significant. Previous studies have reported that miR-34a played an important role in the pathogenesis of RA, but the findings were controversial (*Niederer et al., 2012*; *Xie et al., 2019*). Hence, we analyzed the expression of miR-34a-5p in PBMCs using RT-qPCR and found that miR-34a-5p was significantly down-regulated in RA patients compared to health controls (Fig. 11G).

DISCUSSION

We integrated multiple bioinformatics analysis methods to attain diagnostic genes and vital microRNAs of RA. A total of 14 modules were identified using the WGCNA algorithm. Of these, the turquoise module was large in size (containing 845 genes) and highly correlated with RA, and was selected for further study. We found that the main functions of genes were to participate in leukocyte activation and regulate cell–cell adhesion by performing GO annotation on the turquoise module. KEGG results showed that these genes were highly focused on cytokine-cytokine receptor interaction, chemokine signaling pathway, NF-kappa B signaling pathway, and T cell receptor signaling pathway, in addition to regulating helper T (Th) cell differentiation. Cytokines are small peptides or glycoproteins produced by stimulated immune cells and some stromal cells (e.g., endothelial cells) with a wide range
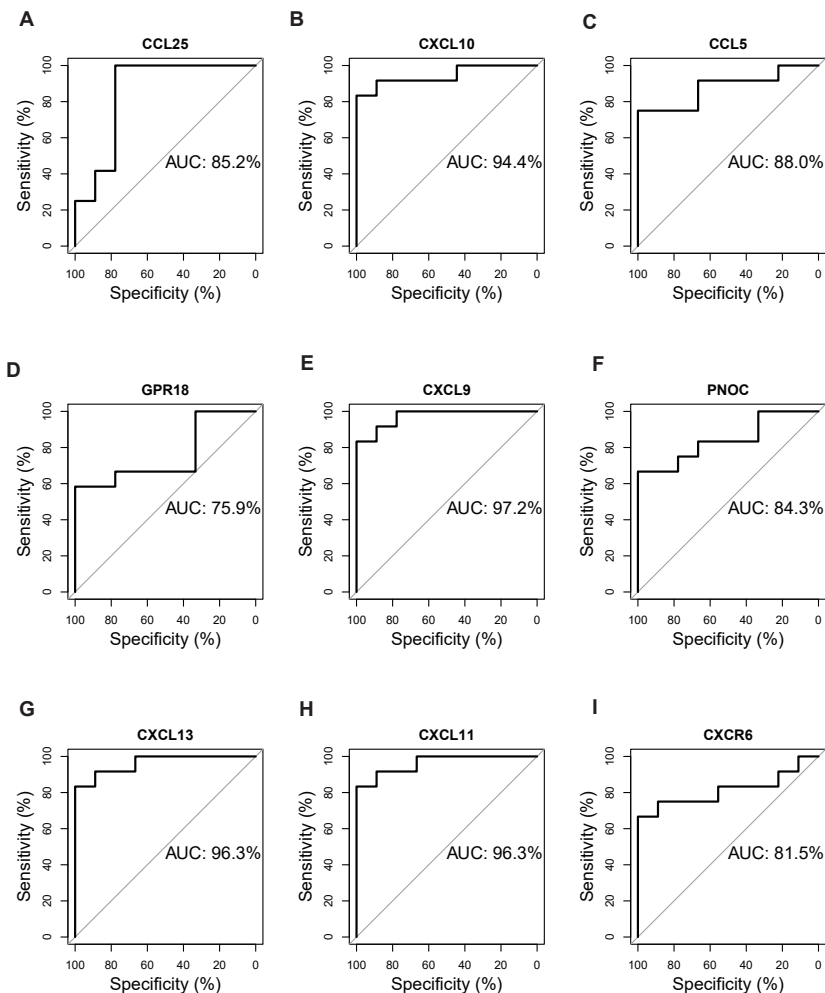
Figure 9 **Validation of hub genes by ROC curve analysis.** AUC value of each plot was calculated. Genes with AUC values greater than 85% were considered as potential biomarkers for the diagnosis of RA. (A) CCL25, (B) CXCL10, (C) CCL5, (D) GPR18, (E) CXCL9, (F) PNOC, (G) CXCL13, (H) CXCL11, (I) CXCR6.

Full-size 🖼 DOI: 10.7717/peerj.11427/fig-9

of biological activities (*Lackie, 2010*). They can mediate long-lasting tissue inflammation and injury by autocrine or paracrine systems and have been shown to be implicated in the pathogenesis of many host autoimmune diseases, such as RA, inflammatory bowel disease, and ankylosing spondylitis (*Schett et al., 2013*). It is believed that TNF and IL-6 are critical pro-inflammatory cytokines that promote osteoclast maturation leading to cartilage degeneration and matrix degradation, and can induce the release of other stimulating factors (such as IL-1), resulting in a vicious cycle of unrelieved synovitis in RA (*Bertolini et al., 1986*; *Ohshima et al., 1998*). The NF-kappa B signaling pathway, activated by the receptor activator of nuclear factor κB ligand (RANKL), is also worthy of attention for its pathogenic effect on RA. *Hirayama et al. (2005)*. found that blocking the NF-kappa B pathway with STAT-6 fusion protein significantly alleviated joint inflammation and
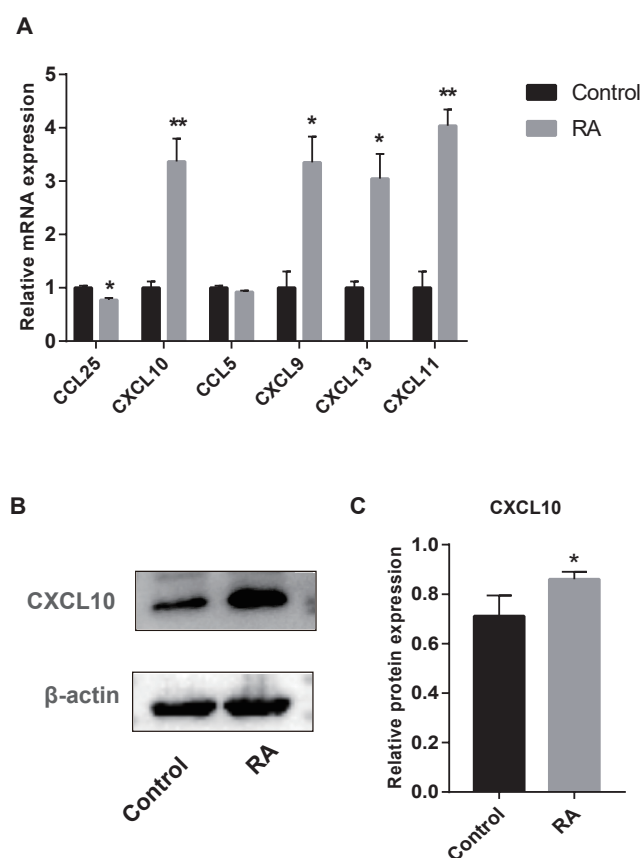
Figure 10 **Validation of hub genes by RT-qPCR and Western blot.** (A) RT-qPCR validation of the hub gene between RA and normal controls. (B) Expression of CXCL10 in PBMCs was detected with Western blot. (C) The relative expression of CXCL10 was quantified via normalization to β-actin. * p-value < 0.05; ** p-value < 0.01.

Full-size ⬛ DOI: 10.7717/peerj.11427/fig-10

bone destruction in mice (*Hirayama et al., 2005*). Moreover, GSEA analysis identified five inflammation-associated gene sets that were up-regulated in the RA phenotype, such as IL6-JAK-STAT3 signaling, interferon α/γ response and complement, which were highly consistent with the results of GO and KEGG annotation. The Janus kinase (JAK) family, consisting of four JAK proteins, transduces inflammatory cytokines signals (e.g., IL-6 and interferon) through the JAK-STAT pathway. A double-blind clinical study showed promising clinical improvement in RA patients after oral administration of tofacitinib, a drug that blocks the JAK1-mediated cytokine signaling and inhibits STAT phosphorylation (*Boyle et al., 2015*). We used CIBERSORT to quantify the composition and content of immune cells in synovial tissue to further understand the immune microenvironment of RA. We found that the RA specimens contained higher levels of leukocyte subtypes, such as plasma cells and CD8 T cells, than normal tissue using the Wilcox test. The influx of immune cells into joint forms a complicated inflammatory network that activates fibroblast-like synoviocytes (FLS), inducing synovial vascular proliferation, and recruiting more immune cells to inflammatory sites (*Chen et al., 2019*).
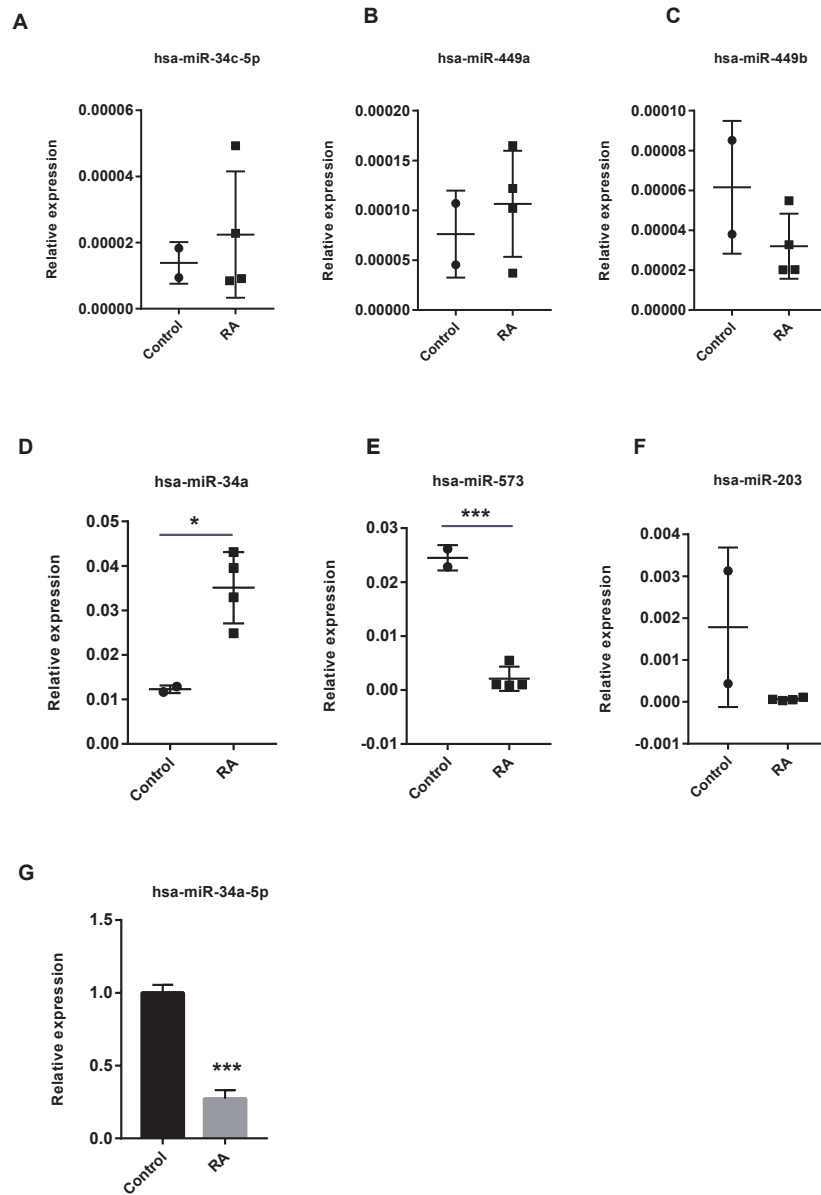
Figure 11 **Validation of vital microRNAs.** (A-F) The relative expressions of hsa-miR-34c-5p, hsa-miR-449a, hsa-miR-449b, hsa-miR-34a, hsa-miR-573, and hsa-miR-203 in synovial tissue. Data were obtained from the GEO dataset GSE37425 and presented as log2 expression of probe intensity (health control: $n = 2$, RA patients: $n = 4$). (G) hsa-miR-34a-5p expression in PBMCs was detected with RT-qPCR. Data were relative to U6 expression. * p-value < 0.05, *** p-value < 0.001.

Full-size 🖻 DOI: 10.7717/peerj.11427/fig-11

We identified five clusters in the PPI network and the most significant one contained nine hub genes, five of which were confirmed by ROC curves analysis and RT-qPCR, including CCL25, CXCL9, CXCL10, CXCL11, and CXCL13. Chemokines can be divided into homeostatic subgroup (e.g., CXCL12) and inflammatory subgroup (such as CXCL9, CXCL10) depending on their function(*Vergunst & Tak, 2005*). However, this functional

classification is not static, and chemokines act in a context-dependent manner based on different tissues and disease stages. A recent study found that plasma levels of chemokines, such as CXCL9, CXCL10, and CXCL13, are much higher in patients with early untreated RA than in normal subjects. Among these discriminators, CXCL10 is associated with indicators of disease activity (e.g., DAS28), which may help determine whether RA patients are in an active phase of the disease (*Pandya et al., 2017*). CXCL10 is mainly produced by interferon-γ stimulated T lymphocytes, neutrophils, and monocytes, and is therefore also known as 10 kDa IP-10, which exerts its biological function by binding to the receptor CXCR3 (*Antonelli et al., 2014*). High levels of CXCL10 in humoral components, such as peripheral blood and joint fluid, are typical of various autoimmune diseases, especially those with a predominance of Th1 cells (*Lee, Lee & Song, 2009*). On the one hand, CXCL10 and CXCR3 play an important role in the homing of leukocytes to inflamed tissues (*Romagnani et al., 2002*); on the other hand, the massive accumulation of inflammatory cells, especially Th1 lymphocytes, enhances the secretion of IFN-γ and TNF-α, which in turn stimulates CXCL10 production, thus amplifying the feedback loop and leading to persistent inflammation and tissue damage (*Antonelli et al., 2008*). Chemokine receptors, which bind to chemokines, have also been shown to be involved in Th1-type and Th2-type inflammatory responses (*Norii et al., 2006*). Specific or non-specific targeting of chemokines and their receptors can be used to treat RA, and many targeted therapeutics have been tested in clinical trials with satisfactory results. For instance, patients taking both methotrexate and MDX1100, an antagonist of CXCL10, achieved reduced disease activity and improved symptom relief (*Yellin et al., 2012*). The application of drugs targeting chemokine receptors, such as CP481,715, showed promising results by reducing the number of macrophages in synovial tissue to ameliorate inflammatory response (*Clucas et al., 2007*).

We obtained 72 miRNA-gene regulatory relationships in our analysis of miRNA-gene interaction, of which seven miRNAs possessed two target genes, including hsa-miR-449a, hsa-miR-297, hsa-miR-203a-3p, hsa-miR-449b-5p, hsa-miR-34c-5p, hsa-miR-573, and hsa-miR-34a-5p. miRNAs belong to non-coding RNAs, which regulate cellular activity by degrading mRNA or inhibiting transcriptional processes (*Esmailzadeh et al., 2017*). It has been reported that about 60% of genes in the human body are regulated by miRNAs, and a single miRNA usually possesses thousands of target genes, which means that abnormal miRNA function may cause pathological consequences such as immune-related diseases and certain types of tumor (*Baltimore et al., 2008*; *O'Connell et al., 2010*). Widely expressed in a variety of immune cells (dendritic cells, monocytes, lymphocytes, etc.), miR-34a plays an important role in regulating cell development, function and survival, which is an important hub in the regulatory network of T cells (*Taheri et al., 2020*). In this study, we found that miR-34a-5p was significantly down-regulated in RA patients compared to health controls. Similarly, it has been shown that miR-34a was not well-expressed in RA synoviocytes, which contributed to their resistance to apoptosis and thus caused persistent inflammation of synovial tissue (*Hou, Wang & Zhang, 2019*; *Niederer et al., 2012*). In addition, *Wu et al. (2021)* found that miR-34a in extracellular vesicles (Evs) was also involved in the development of RA (*Wu et al., 2021*). Elevated expression of miR-34a

in BM-MSC-derived Evs activated the p53 signaling pathway by inhibiting cyclin I (CCNI) expression, which in turn inhibited FLS proliferation and promotes apoptosis, and thus had a positive effect on alleviating the inflammatory state of RA. However, some scholars hold the opposite view. It has been shown that miR-34a expression was up-regulated in PBMCs from RA patients (*Ebrahimiyan et al., 2019*; *Xie et al., 2019*), and silencing miR-34a markedly reduced release of inflammatory mediators (*Kurowska-Stolarska et al., 2017*). In other autoimmune diseases, such as Sjögren Syndrome, the level of miR-34a was also significantly elevated in patients (*Kim et al., 2019*). This controversy suggests that research on miR-34a should continue. miR-449a has recently been identified as an important regulator of RA pathogenesis by suppressing HMGB1 and YY1 expression, thereby attenuating synovial hyperplasia, FLS migration, and the release of inflammatory mediators (*Cai et al., 2019*). *Jin et al. (2018)* found that miR-34c expression was significantly decreased in serum of Treg-depleted mice, an animal model commonly used in studies of autoimmune diseases (*Jin et al., 2018*).

Our study has limitations, including our focus on the most critical turquoise model in the process of co-expression network analysis, without adequate discussion of the black, blue, and purple modules, which were also strongly associated with RA. Our study size was small, and in-depth clinical studies are needed to support our findings.

CONCLUSIONS

In summary, we identified diagnostic genes and vital microRNAs associated with RA, which can help us to better understand the pathogenesis behind this disorder and provide theoretical support for exploring more effective therapies.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests
The authors declare there are no competing interests.

Author Contributions

- Conglin Ren conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Mingshuang Li performed the experiments, prepared figures and/or tables, and approved the final draft.
- Yang Zheng, Fengqing Wu and Weibin Du analyzed the data, prepared figures and/or tables, and approved the final draft.
- Renfu Quan conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

This study was approved by Ethics Committees of Xiaoshan Traditional Chinese Medicine Hospital (Ethical Application Ref: 2020012).

Data Availability

The following information was supplied regarding data availability:

The R codes and raw data are available in the Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.11427#supplemental-information.

REFERENCES

Antonelli A, Ferri C, Ferrari SM, Colaci M, Fallahi P. 2008. Immunopathogenesis of HCV-related endocrine manifestations in chronic hepatitis and mixed cryoglobu-linemia. *Autoimmunity Reviews* **8**:18–23 DOI 10.1016/j.autrev.2008.07.017.

Antonelli A, Ferrari SM, Giuggioli D, Ferrannini E, Ferri C, Fallahi P. 2014. Chemokine (C-X-C motif) ligand (CXCL)10 in autoimmune diseases. *Autoimmun Rev* **13**:272–280 DOI 10.1016/j.autrev.2013.10.010.

Bader GD, Hogue CW. 2003. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4**:2 DOI 10.1186/1471-2105-4-2.

Baltimore D, Boldin MP, O'Connell RM, Rao DS, Taganov KD. 2008. MicroRNAs: new regulators of immune cell development and function. *Nature Immunology* **9**:839–845 DOI 10.1038/ni.f.209.

Bertolini DR, Nedwin GE, Bringman TS, Smith DD, Mundy GR. 1986. Stimulation of bone resorption and inhibition of bone formation in vitro by human tumour necrosis factors. *Nature* **319**:516–518 DOI 10.1038/319516a0.

Boyle DL, Soma K, Hodge J, Kavanaugh A, Mandel D, Mease P, Shurmur R, Singhal AK, Wei N, Rosengren S, Kaplan I, Krishnaswami S, Luo Z, Bradley J, Firestein

GS. 2015. The JAK inhibitor tofacitinib suppresses synovial JAK1-STAT signalling in rheumatoid arthritis. *Annals of the Rheumatic Diseases* **74**:1311–1316 DOI 10.1136/annrheumdis-2014-206028.

Cai Y, Jiang C, Zhu J, Xu K, Ren X, Xu L, Hu P, Wang B, Yuan Q, Guo Y, Sun J, Xu P, Qiu Y. 2019. miR-449a inhibits cell proliferation, migration, and inflammation by regulating high-mobility group box protein 1 and forms a mutual inhibition loop with Yin Yang 1 in rheumatoid arthritis fibroblast-like synoviocytes. *Arthritis Research & Therapy* **21**:134 DOI 10.1186/s13075-019-1920-0.

Chen J, Wright K, Davis JM, Jeraldo P, Marietta EV, Murray J, Nelson H, Matteson EL, Taneja V. 2016. An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome Medicine* **8**:43 DOI 10.1186/s13073-016-0299-7.

Chen Y, Wang X. 2020. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Research* **48**:D127–D131 DOI 10.1093/nar/gkz757.

Chen Z, Bozec A, Ramming A, Schett G. 2019. Anti-inflammatory and immune-regulatory cytokines in rheumatoid arthritis. *Nature Reviews Rheumatology* **15**:9–17 DOI 10.1038/s41584-018-0109-2.

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vailaya A, Wang P-L, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, Ideker T, Bader GD. 2007. Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols* **2**:2366–2382 DOI 10.1038/nprot.2007.324.

Clucas AT, Shah A, Zhang YD, Chow VF, Gladue RP. 2007. Phase I evaluation of the safety, pharmacokinetics and pharmacodynamics of CP-481,715. *Clinical Pharmacokinetics* **46**:757–766 DOI 10.2165/00003088-200746090-00003.

Ebrahimiyan H, Rezaei N, Vojdanian M, Aslani S, Jamshidi A, Mahmoudi M. 2019. microRNA involvement in the regulation of survivin in peripheral blood mononuclear cells from rheumatoid arthritis patients. *International Journal of Rheumatic Diseases* **22**:1107–1114 DOI 10.1111/1756-185x.13520.

Esmailzadeh S, Mansoori B, Mohammadi A, Baradaran B. 2017. Regulatory roles of micro-RNAs in T cell autoimmunity. *Immunological Investigations* **46**:864–879 DOI 10.1080/08820139.2017.1373901.

Hajishengallis G. 2015. Periodontitis: from microbial immune subversion to systemic inflammation. *Nature Reviews Immunology* **15**:30–44 DOI 10.1038/nri3785.

Hirayama T, Dai S, Abbas S, Yamanaka Y, Abu-Amer Y. 2005. Inhibition of inflammatory bone erosion by constitutively active STAT-6 through blockade of JNK and NF-κB activation. *Arthritis & Rheumatism* **52**:2719–2729 DOI 10.1002/art.21286.

Hou C, Wang D, Zhang L. 2019. MicroRNA-34a-3p inhibits proliferation of rheumatoid arthritis fibroblast-like synoviocytes. *Molecular Medicine Reports* **20**:2563–2570 DOI 10.3892/mmr.2019.10516.

Jin F, Hu H, Xu M, Zhan S, Wang Y, Zhang H, Chen X. 2018. Serum microRNA profiles serve as novel biomarkers for autoimmune diseases. *Frontiers in Immunology* **9**:2381 DOI 10.3389/fimmu.2018.02381.

Kawatkar A, Portugal C, Chu L-H, Iyer R. 2012. Racial/Ethnic trends in incidence and prevalence of rheumatoid arthritis in a large multi-ethnic managed care population. *Arthritis & Rheumatism* **64**:S1061–S1061.

Kharlamova N, Jiang X, Sherina N, Potempa B, Israelsson L, Quirke AM, Eriksson K, Yucel-Lindberg T, Venables PJ, Potempa J, Alfredsson L, Lundberg K. 2016. Antibodies to porphyromonas gingivalis indicate interaction between oral infection, smoking, and risk genes in rheumatoid arthritis etiology. *Arthritis Rheumatol* **68**:604–613 DOI 10.1002/art.39491.

Kim YJ, Yeon Y, Lee WJ, Shin YU, Cho H, Sung YK, Kim DR, Lim HW, Kang MH. 2019. Comparison of MicroRNA expression in tears of normal subjects and Sjögren syndrome patients. *Investigative Ophthalmology and Visual Science* **60**:4889–4895 DOI 10.1167/iovs.19-27062.

Kurowska-Stolarska M, Alivernini S, Melchor EG, Elmesmari A, Tolusso B, Tange C, Petricca L, Gilchrist DS, Di Sante G, Keijzer C, Stewart L, Di Mario C, Morrison V, Brewer JM, Porter D, Milling S, Baxter RD, McCarey D, Gremese E, Lemke G, Ferraccioli G, McSharry C, McInnes IB. 2017. MicroRNA-34a dependent regulation of AXL controls the activation of dendritic cells in inflammatory arthritis. *Nature Communications* **8**:15877 DOI 10.1038/ncomms15877.

Lackie J. 2010. *A dictionary of biomedicine.* Oxford: Oxford University Press.

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**:559 DOI 10.1186/1471-2105-9-559.

Langfelder P, Horvath S. 2012. Fast R functions for robust correlations and hierarchical clustering. *Journal of Statistical Software* **46**:1–17.

Lee EY, Lee ZH, Song YW. 2009. CXCL10 and autoimmune diseases. *Autoimmun Rev* **8**:379–383 DOI 10.1016/j.autrev.2008.12.002.

Leng RX, Pan HF, Qin WZ, Chen GM, Ye DQ. 2011. Role of microRNA-155 in autoimmunity. *Cytokine and Growth Factor Reviews* **22**:141–147 DOI 10.1016/j.cytogfr.2011.05.002.

Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Systems* **1**:417–425 DOI 10.1016/j.cels.2015.12.004.

Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**:1739–1740 DOI 10.1093/bioinformatics/btr260.

Liu W, Wang X. 2019. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biology* **20**:18 DOI 10.1186/s13059-019-1629-z.

MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, Silman AJ. 2000. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthtitis and Rheumatism* **43**:30–37 DOI 10.1002/1529-0131(200001)43:1<30::Aid-anr5>3.0.Co;2-b.

Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC. 2003. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics* **34**:267–273 DOI 10.1038/ng1180.

Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. 2015. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* **12**:453–457 DOI 10.1038/nmeth.3337.

Niederer F, Trenkmann M, Ospelt C, Karouzakis E, Neidhart M, Stanczyk J, Kolling C, Gay RE, Detmar M, Gay S, Jüngel A, Kyburz D. 2012. Down-regulation of microRNA-34a* in rheumatoid arthritis synovial fibroblasts promotes apoptosis resistance. *Arthtitis and Rheumatism* **64**:1771–1779 DOI 10.1002/art.34334.

Norii M, Yamamura M, Iwahashi M, Ueno A, Yamana J, Makino H. 2006. Selective recruitment of CXCR3+ and CCR5+ CCR4+ T cells into synovial tissue in patients with rheumatoid arthritis. *Acta Medica Okayama* **60**:149–157.

O'connell RM, Rao DS, Chaudhuri AA, Baltimore D. 2010. Physiological and pathological roles for microRNAs in the immune system. *Nature Reviews Immunology* **10**:111–122 DOI 10.1038/nri2708.

Ohshima S, Saeki Y, Mima T, Sasai M, Nishioka K, Nomura S, Kopf M, Katada Y, Tanaka T, Suemura M. 1998. Interleukin 6 plays a key role in the development of antigen-induced arthritis. *Proceedings of the National Academy of Sciences of the United States of America* **95**:8222–8226 DOI 10.1073/pnas.95.14.8222.

Otasek D, Morris JH, Boucas J, Pico AR, Demchak B. 2019. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biology* **20**:185 DOI 10.1186/s13059-019-1758-4.

Pan W, Zhu S, Yuan M, Cui H, Wang L, Luo X, Li J, Zhou H, Tang Y, Shen N. 2010. MicroRNA-21 and microRNA-148a contribute to DNA hypomethylation in lupus CD4+ T cells by directly and indirectly targeting DNA methyltransferase 1. *Journal of Immunology* **184**:6773–6781 DOI 10.4049/jimmunol.0904060.

Pandya JM, Lundell AC, Andersson K, Nordstrom I, Theander E, Rudin A. 2017. Blood chemokine profile in untreated early rheumatoid arthritis: CXCL10 as a disease activity marker. *Arthritis Research & Therapy* **19**:20 DOI 10.1186/s13075-017-1224-1.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**:e47 DOI 10.1093/nar/gkv007.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**:77 DOI 10.1186/1471-2105-12-77.

Romagnani P, Rotondi M, Lazzeri E, Lasagni L, Francalanci M, Buonamano A, Milani S, Vitti P, Chiovato L, Tonacchera M, Bellastella A, Serio M. 2002. Expression of IP-10/CXCL10 and MIG/CXCL9 in the thyroid and increased levels of IP-10/CXCL10 in the serum of patients with recent-onset Graves' disease. *American Journal of Pathology* **161**:195–206 DOI 10.1016/s0002-9440(10)64171-5.

Schett G, Elewaut D, McInnes IB, Dayer J-M, Neurath MF. 2013. How cytokine networks fuel inflammation: toward a cytokine-based disease taxonomy. *Nature Medicine* **19**:822–824 DOI 10.1038/nm.3260.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**:2498–2504 DOI 10.1101/gr.1239303.

Simpson LJ, Ansel KM. 2015. MicroRNA regulation of lymphocyte tolerance and autoimmunity. *The Journal of Clinical Investigation* **125**:2242–2249 DOI 10.1172/JCI78090.

Smolen JS, Aletaha D, McInnes IB. 2016. Rheumatoid arthritis. *Lancet* **388**:2023–2038 DOI 10.1016/s0140-6736(16)30173-8.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**:15545–15550 DOI 10.1073/pnas.0506580102.

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, von Mering C. 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research* **43**:D447–452 DOI 10.1093/nar/gku1003.

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research* **47**:D607–D613 DOI 10.1093/nar/gky1131.

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, Von Mering C. 2017. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Research* **45**:D362–d368 DOI 10.1093/nar/gkw937.

Taheri F, Ebrahimi SO, Shareef S, Reiisi S. 2020. Regulatory and immunomodulatory role of miR-34a in T cell immunity. *Life Sciences* **262**:118209 DOI 10.1016/j.lfs.2020.118209.

Vergunst CE, Tak PP. 2005. Chemokines: their role in rheumatoid arthritis. *Current Rheumatology Reports* **7**:382–388 DOI 10.1007/s11926-005-0026-7.

Wu H, Zhou X, Wang X, Cheng W, Hu X, Wang Y, Luo B, Huang W, Gu J. 2021. miR-34a in extracellular vesicles from bone marrow mesenchymal stem cells reduces

rheumatoid arthritis inflammation via the cyclin I/ATM/ATR/p53 axis. *Journal of Cellular and Molecular Medicine* **25**:1896–1910 DOI 10.1111/jcmm.15857.

Xie M, Wang J, Gong W, Xu H, Pan X, Chen Y, Ru S, Wang H, Chen X, Zhao Y, Li J, Yin Q, Xia S, Zhou X, Liu X, Shao Q. 2019. NF-kappaB-driven miR-34a impairs Treg/Th17 balance via targeting Foxp3. *Journal of Autoimmunity* **102**:96–113 DOI 10.1016/j.jaut.2019.04.018.

Yang H, Li H. 2019. CD36 identified by weighted gene co-expression network analysis as a hub candidate gene in lupus nephritis. *PeerJ* **7**:e7722 DOI 10.7717/peerj.7722.

Yao Q, Song Z, Wang B, Qin Q, Zhang JA. 2019. Identifying key genes and functionally enriched pathways in sjogren's syndrome by weighted gene co-expression network analysis. *Front Genet* **10**:1142 DOI 10.3389/fgene.2019.01142.

Yellin M, Paliienko I, Balanescu A, Ter-Vartanian S, Tseluyko V, Xu LA, Tao X, Cardarelli PM, LeBlanc H, Nichol G. 2012. A phase II, randomized, double-blind, placebo-controlled study evaluating the efficacy and safety of MDX-1100, a fully human anti-CXCL10 monoclonal antibody, in combination with methotrexate in patients with rheumatoid arthritis. *Arthritis & Rheumatism* **64**:1730–1739.

Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* **16**:284–287 DOI 10.1089/omi.2011.0118.

Zhang J, Wang X, Xu L, Zhang Z, Wang F, Tang X. 2020. Investigation of potential genetic biomarkers and molecular mechanism of ulcerative colitis utilizing bioinformatics analysis. *BioMed Research International* **2020**:4921387 DOI 10.1155/2020/4921387.

Zhong D, Wu C, Bai J, Xu D, Zeng X, Wang Q. 2020. Co-expression network analysis reveals the pivotal role of mitochondrial dysfunction and interferon signature in juvenile dermatomyositis. *PeerJ* **8**:e8611 DOI 10.7717/peerj.8611.