Integrated bioinformatics analysis and experimental validation reveals hub genes of rheumatoid arthritis

KUN LUO[1], YUMEI ZHONG[2], YANDING GUO[1], JINGWEI NIE[1], YIMEI XU[1] and HAIYAN ZHOU[1]

[1]Acupuncture and Tuina School, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan 610075; [2]Department of Painology, Chengdu Integrated TCM & Western Medicine Hospital/Chengdu First People's Hospital, Chengdu, Sichuan 610095, P.R. China

Received March 20, 2023; Accepted August 1, 2023

DOI: 10.3892/etm.2023.12179

Abstract. Rheumatoid arthritis (RA) is an autoimmune disease characterized by systemic inflammation, especially synovitis, leading to joint damage. It is important to explore potential biomarkers and therapeutic targets to improve the clinical treatment of RA. However, the potential underlying mechanisms of action of available treatments for RA have not yet been fully elucidated. The present study investigated the potential biomarkers of RA and identified specific targets for therapeutic intervention. A comprehensive analysis was performed using mRNA files downloaded from the Gene Expression Omnibus. Differences in gene expression were analyzed and compared between the normal and RA groups. In addition, Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses were performed on differentially expressed genes (DEGs). A protein-protein interaction network, Molecular Complex Detection and cytoHubba network were evaluated to identify hub genes. Finally, using an experimental RA rat model induced by Freund's complete adjuvant (FCA), the expression of potential biomarkers or target genes in RA were verified through reverse transcription-quantitative PCR. The results of the mRNA dataset processing revealed 195 DEGs in patients with RA when compared with the healthy controls. Moreover, 10 hub genes were identified in patients with RA and four candidate mRNAs were identified, as follows: Discs large homolog-associated protein 5 (*DLGAP5*), kinesin family member 20A (*KIF20A*), maternal embryonic leucine zipper kinase (*MELK*) and nuclear division cycle 80 (*NDC80*). Finally, the bioinformatics analysis results were validated by quantifying the expression of the *DLGAP5*, *KIF20A*, *MELK* and *NDC80* genes in the FCA-induced experimental RA rat model. The findings of the present study suggested that the treatment of RA may be successful through the inhibition of *DLGAP5*, *KIF20A*, *MELK* and *NDC80* expression. Therefore, the targeting of these genes may result in more effective treatments for patients with RA.

Correspondence to: Professor Haiyan Zhou, Acupuncture and Tuina School, Chengdu University of Traditional Chinese Medicine, 37 Shierqiao Road, Chengdu, Sichuan 610075, P.R. China
E-mail: zhouhaiyan@cdutcm.edu.cn

Key words: rheumatoid arthritis, autoimmune disease, bioinformatics, biomarkers, gene expression omnibus

Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterized by high levels of synovitis leading to joint damage, and is associated with notable morbidity and mortality (1). Clinical epidemiological data have confirmed that RA occurs in 0.5-1% of the global population, and that it occurs 2-3 times more frequently in women than in men. In addition, RA can affect people of any age, but the peak onset is between 50 and 59 years (2). If RA occurs owing to a family history of inheritance, then the risk of disease development can increase by 3-5 times (3). Currently, there is no clear understanding about the pathogenesis of RA, but there is a widespread belief it is related to immune abnormalities, probably influenced by genetic and environmental factors (4). Rheumatoid factor (RF) testing is an important indicator of an RA diagnosis and prognosis and indicates disease severity. However, RF testing yields a negative result in the early stages of RA, similar to that obtained for other autoimmune and inflammatory diseases (5). Therefore, more effective and sensitive auxiliary biomarkers of RA are needed. Based on the genetic predisposition, guanylate-binding protein 1 (*GBP1*), C-X-C motif chemokine ligand 10 (*CXCL10*), granulocyte-macrophage colony-stimulating factor (*GM-CSF*) and its receptor colony-stimulating factor 2 receptor β (*CSF2RB*) have been reported to be highly expressed in RA, and therefore may play an important role in the early diagnosis and treatment of RA (6,7). Genetic testing may provide opportunities for the early detection of effective therapeutic responses to RA, even a few months before the preclinical phase (8). Therefore, identifying novel and feasible biomarkers is critical for the early diagnosis and treatment of RA.

In the present study, bioinformatics analysis was performed to determine the potential biomarkers of RA and identify specific targets for therapeutic intervention, with the aim of improving the understanding of the mechanism of RA and the clinical treatment of RA. Initially, raw data of RA from

the NCBI Gene Expression Omnibus (GEO) database were obtained and four hub genes were selected. Furthermore, the results of bioinformatics analysis were verified using a rat model of Freund's complete adjuvant (FCA)-induced RA. These results may provide insights into the underlying molecular mechanisms of RA and potential biomarkers and target genes for treatment.

Materials and methods

Selection of GEO datasets. A flowchart of the overall experimental design is shown in Fig. 1. The mRNA datasets were downloaded from the GEO database (www.ncbi.nlm.nih.gov/geo/). To gain an improved understanding of the underlying mechanism of RA, the independent datasets of GSE55457 (6) and GSE77298 (7) were selected, which included synovial tissues from 13 and 16 patients with RA and 10 and 7 healthy subjects, respectively.

Differential expression analysis. To identify differentially expressed mRNAs (DE-mRNAs) using the GSE55457 and GSE77298 datasets, the mRNA expression levels of the tissue samples from patients with RA and healthy subjects were compared. DE-mRNAs were defined using the 'limma' package (version 3.52.2) (9), based on the criteria of $\log_2$ fold change >1 and P<0.05 (8).

Functional enrichment analysis. Database for Annotation, Visualization and Integrated Discover (DAVID) (https://david.ncifcrf.gov/) analysis was performed on Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) terms (10). In the present study, the 'enrichGO' function of the clusterProfiler package (10) was used to perform GO pathway enrichment analysis based on the differentially expressed genes (DEGs). Based on their function, these genes were divided into three categories: Biological process (BP), molecular function (MF) and cellular component (CC). A significant enrichment of GO terms was observed for DEGs based on the false discovery rate-adjusted P<0.05. The DAVID webtools were used to analyze KEGG enrichment of the integrated DEGs in the present study. Adjusted P<0.05 was considered to indicate a statistically significant difference. The software packages R (version 4.2.2; https://www.R-project.org/) (11), MCODE plug-in (version 2.0.2) (12,13) and CytoHubba plug-in (https://apps.cytoscape.org/apps/cytohubba) in Cytoscape (version 3.9.1) (14,15) were used to obtain the results.

Building a protein-protein interaction (PPI) network and identifying hub genes. Based on the STRING database (version 11.5; using confidence score >0.9) (16), a set of high-confidence PPI networks for humans were constructed. To screen for hub genes, topological algorithms of the cytoHubba plug-in application in Cytoscape (version 3.9.1) were applied to predict and explore important nodes and hubs. The core module was extracted from the PPI network using the MCODE plug-in. The parameters were set as follows: Degree cut-off=2, K-Core=2 and node score cut-off=0.2 (17). It was hypothesized that the potential hub genes could be found in the core module. Thus, the genes selected were considered the hub genes.

Animals. A total of 30 male Sprague Dawley rats (6 weeks old; 180±20 g body weight) were obtained from Chengdu Dashuo Laboratory Animal Co., Ltd. (permit no. SCXK 2015-030). The animals were fed and maintained under laboratory conditions of 22-24˚C and 50-65% humidity, with natural light and dark cycles, and had free access to food and drinking water. The rats were randomized into the control and RA groups (n=15/group). All animal procedures were approved by The Ethics Committee of Chengdu University of Traditional Chinese Medicine (Chengdu, China; approval no. 2018-11).

Establishment of the RA model. On the first day of the experiment, to establish the RA model, the rats were in the RA group subjected to the induction process, wherein they received an injection of 0.5 ml/kg of FCA (Sigma-Aldrich; Merck KGaA) in the footpad. The control group was injected with the same amount of normal saline. RA progression was evaluated based on the swelling and redness of the foot pad. The overall condition and bipedal lesions of all rats were observed daily. A waterproof marker was used to mark the middle of the footpad of each rat and all future measurements were taken from this point. Calipers were used to measure the footpad thickness before and 1, 2 and 3 weeks after establishment of the RA model (18).

At the end of the experiment, the rats were anesthetized with 7% chloral hydrate (400 mg/kg, intraperitoneally; Sinopharm Chemical Reagent Co., Ltd.). It was confirmed that the rats showed no signs of peritonitis, pain or discomfort following this anesthesia method. Blood was subsequently collected from the abdominal aorta of the live anesthetized rats. After the 5 ml of blood was collected, the rats were subsequently sacrificed by cervical dislocation. Ankle joints, synovial tissue and the spleen were also collected from each animal.

Hematoxylin and eosin (H&E) staining. The ankle joint was fixed with 4% paraformaldehyde (Wuhan Servicebio Technology Co., Ltd.) for >24 h at room temperature. After fixation, the paraffin-embedded tissue sections were sliced into 5-mm sections using a paraffin microtome (Leica Microsystems, Ltd.). The sections were then stained with the H&E staining solution kit (Wuhan Servicebio Technology Co., Ltd.), which included staining with hematoxylin and eosin individually (5 min at room temperature for each staining step). The stained sections were dehydrated with ethanol, cleared with xylene, mounted onto slides with neutral balsam and covered with coverslips. Histopathological examination was performed on the mounted sections and images were captured under a light microscope (Nikon Corporation). The slides were graded according to the degree of inflammatory changes. The rat synovial pathological score included synovial tissue hyperplasia, inflammatory cell infiltration and macrophage hyperplasia. The tissue was scored based on the following: 0, normal; 1, sparse or scattered; 2, diffuse increase; 3, large increase with respect to inflammatory cell infiltration, macrophage hyperplasia and fibroplasia.

Enzyme-linked immunosorbent assay (ELISA). The blood samples collected from each group were centrifuged at 1,200 x g for 20 min at room temperature (Heal Force Bio-meditech Holdings Ltd.) and the supernatant was

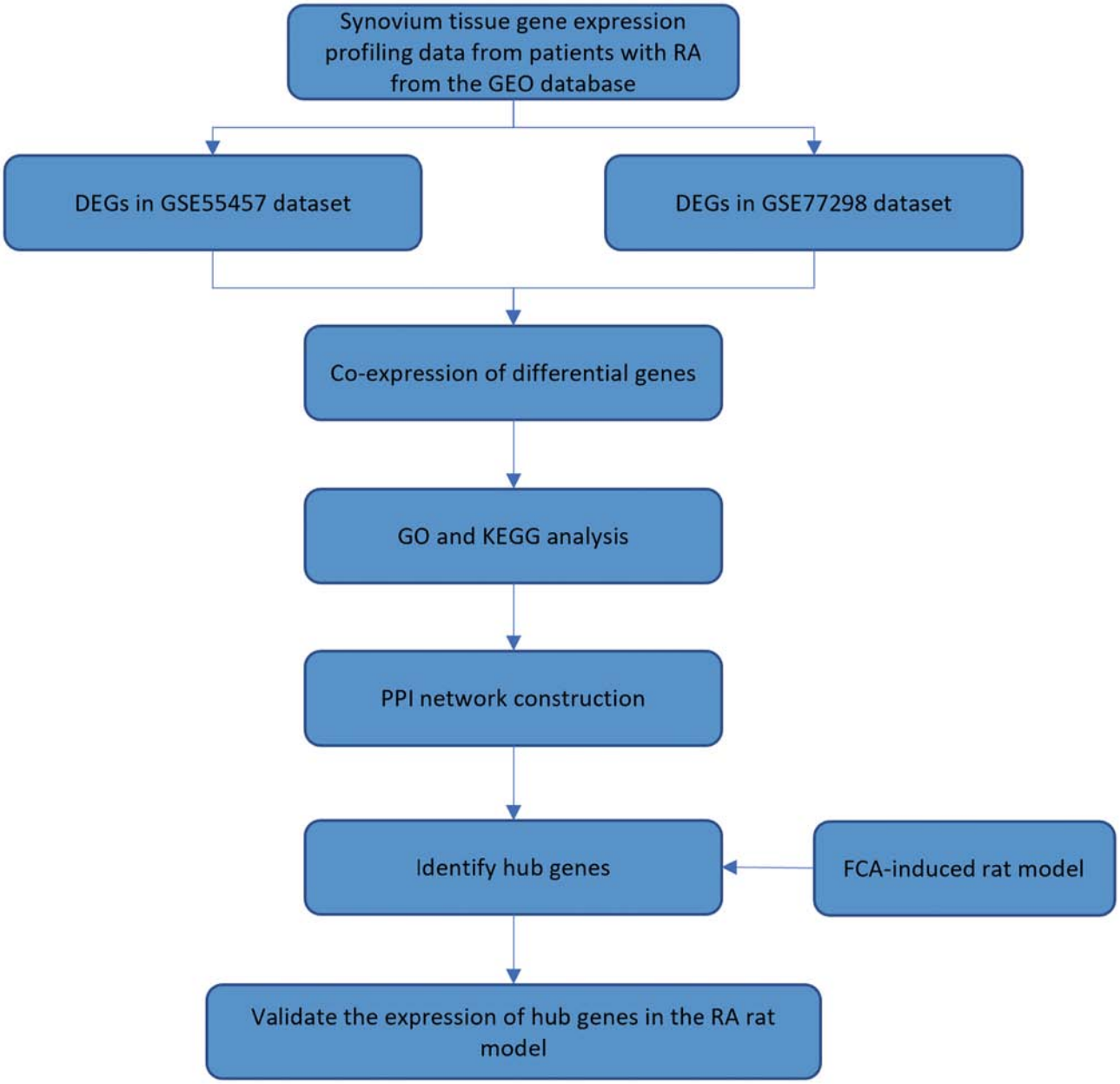

Figure 1. Flowchart of the experimental design. RA, rheumatoid arthritis; DEGs, differentially expressed genes; FCA, Freund's complete adjuvant; GEO, Gene Expression Omnibus; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction.

collected. The expression levels of tumor necrosis factor α (TNF-α) (cat. no. JM-01587R1) and interleukin (IL-) 12 (cat. no. JM10372R1) in the supernatant were detected using an ELISA kit (all Jiangsu Jingmei Biotechnology Co., Ltd.), according to the manufacturer's guidelines. Standards (50 μl) and samples (10 μl of serum and 40 μl of diluent, supplied with the kit) were added to predetermined microplate wells. Horseradish peroxidase (HRP)-labeled conjugate (100 μl; supplied with the kit) was added to the wells, which were then sealed and incubated at 37˚C for 1 h. The plate was washed five times with PBS and substrates A (50 μl) and B (50 μl) were added to each well. After incubation for 15 min at 37˚C, stop solution (50 μl) was added. A microplate reader (BioTek Instruments, Inc.) was used to measure the optical density (OD) values in all wells at a wavelength of 450 nm. TNF-α and IL-12 concentrations were calculated using the obtained OD values of each group of samples, according to the linear regression equation of the standard curve.

Reverse transcription-quantitative polymerase chain reaction (RT-qPCR). Spleen tissue (100 mg) from the RA animal model and control group was homogenized using a homogenization tube. Following the manufacturer's instructions, total RNA was extracted using the TRIzol™ method (Thermo Fisher Scientific, Inc.). The concentration and purity of the isolated RNA were determined using a microvolume spectrophotometer (Thermo Fisher Scientific, Inc.). According to the manufacturer's protocol, RNA was subsequently converted into cDNA using a RT kit (Thermo Fisher Scientific, Inc.) and a PCR machine (Eastwin Life Sciences Inc.). For the qPCR step, the reaction mixture constituted the following components: 2 μl cDNA, 2.5 μl forward and reverse primers,

7.5 μl 2X SYBR Green qPCR Master Mix (Wuhan Servicebio Technology Co., Ltd.) and ultrapure water to reach a 15 μl total volume. Amplification was performed using a fluorescent qPCR instrument (Bio-Rad Laboratories, Inc.). The reaction parameters were as follows: Initial denaturation (95˚C for 30 sec), 40 cycles of denaturation (95˚C for 15 sec) and annealing/extension (60˚C for 30 sec). Data were collected and analyzed using SDS software (version 1.4; Applied Biosystems; Thermo Fisher Scientific, Inc.). GAPDH was used as the internal reference gene. Furthermore, $2^{-\Delta\Delta Cq}$ (19) was used to calculate the relative expression level of discs large homolog-associated protein 5 (*DLGAP5*), kinesin family member 20A (*KIF20A*), maternal embryonic leucine zipper kinase (*MELK*) and nuclear division cycle 80 (*NDC80*). The primer sequences used were as follows: GAPDH-forward (F), 5'-CTGGAGAAACCTGCCAAGTATG-3'; GAPDH-reverse (R), 5'-GGTGGAAGAATGGGAGTTGCT-3'; MELK-F, 5'-TCCCATTGAGTGGCAAAGCA-3'; MELK-R, 5'-CCTGCTGTTGCGGTGATGTA-3'; NDC80-F, 5'-CAAAGACCTGGAAGCCGAACAA-3'; NDC80-R, 5'-CTAGCCAACTTGTGGTACTCTGC-3'; DLGAP5-F, 5'-GAAGTGCCAACGCTTTCCTGA-3'; DLGAP5-R, 5'-CGTTCTTTCCTAGAGTCTGTTTGTG-3'; KIF20A-F, 5'-GAACCACCTAGCCATCAGCATA-3'; KIF20A-R, 5'-ACTGTGACTGCGACTAGAGTGCT-3'.

Western blotting. The synovial tissue was washed 2-3 times with cold PBS to remove blood, after which the tissue was cut into small pieces and placed in a homogenizer tube. Protease inhibitors and 10 times the tissue volume (10 ml buffer/g tissue) of radio immunoprecipitation assay lysis buffer (Wuhan Servicebio Technology, Co., Ltd.) were added to the tube and the contents were homogenized thoroughly on ice. The homogenate was transferred to a 1.5 ml centrifuge tube and the contents were shaken well. The homogenate was pipetted repeatedly for 30 min in an ice bath to ensure complete cell lysis, after which the contents were centrifuged at 12,000 x g for 10 min at 4˚C and the supernatant (containing total protein) was collected. Protein concentration was determined using the bicinchoninic acid method. Subsequently, sodium dodecyl sulfate-polyacrylamide gel electrophoresis was performed using 10% gels to separate the isolated proteins (30 μg/lane). The separated proteins were then transferred onto polyvinylidene difluoride membranes. After blocking the membranes with tris-buffered saline containing 5% skimmed milk powder at room temperature for 1 h, the membranes were incubated with the following primary antibodies overnight at 4˚C: KIF20A (cat. no. ab70791; 1:1,000; Abcam), NDC80 (cat. no. ab109496; 1:1,000; Abcam), MELK (cat. no. ab108529; 1:1,000; Abcam), DLGAP5 (cat. no. ab70744; 1:1,000; Abcam) and GAPDH (cat. no. gb15002; 1:2,000; Wuhan Servicebio Technology, Co., Ltd.). After repeated washing, the membranes were incubated with HRP-conjugated goat anti-rabbit (cat. no. gb23303; 1:5,000; Wuhan Servicebio Technology, Co., Ltd.) secondary antibody for 2 h at 37˚C. Images were collected using a chemiluminescent imaging system (Wuhan Servicebio Technology, Co., Ltd.) and grayscale analysis was performed using the system's ChemiScope analysis software (version ChemiScope 6100; Clinx Science Instrument Co., Ltd.).

Statistical analysis. All data had to pass the Shapiro-Wilk normality test and the Levene's test for homogeneity of variance. Then the results were analyzed by pairwise test. For correlation analyses, when the data distribution was normal, the Pearson correlation analysis was used, whereas when the distribution was not normal, the Spearman correlation analysis was used. Data were presented as the mean ± standard deviation. $P<0.05$ was considered to indicate a statistically significant difference. All data were assessed using SPSS 26.0 software (IBM Corp.). The relevant boxplots were compiled using Xiantao Academic (http://www.xiantao.love).

Results

Identification of DE-mRNAs. The mRNA expression levels from the GSE55457 and GSE77298 datasets from the GEO database were investigated and the results were determined through boxplots, volcano plots and heatmaps (Fig. 2). There were 13 samples from patients with RA and 10 samples from healthy individuals in the GSE55457 dataset, whereas there were 16 samples from patients with RA and 7 samples from healthy individuals in the GSE77298 dataset. In the GSE55457 dataset, 664 DE-mRNAs were identified (429 with upregulated expression and 235 with downregulated expression), while in the GSE77298 dataset, 2,150 DE-mRNAs were detected (1,170 with upregulated expression and 980 with downregulated expression). According to Venn diagram analyses, 165 DE-mRNAs that had upregulated expression and 30 DE-mRNAs that had downregulated expression were shared between the two datasets (Fig. 3 and Table SI).

Enrichment analysis of DE-mRNAs. To explore the potential functions of the identified DE-mRNAs, functional enrichment analysis was performed using GO and KEGG analyses. Among them, the top five BP, top five CC and top five MF GO terms are shown in Fig. 4A and Table SII. In BP, 'leukocyte cell-cell adhesion,' 'mononuclear cell differentiation,' 'lymphocyte differentiation,' 'leukocyte migration' and 'lymphocyte mediated immunity' were the top five most enriched terms. Additionally, 'external side of plasma membrane,' 'endocytic vesicle membrane,' 'clathrin-coated endocytic vesicle,' 'clathrin-coated endocytic vesicle membrane' and 'alpha-β T-cell receptor complex' were the top five GO terms on CC. Furthermore, 'antigen binding,' 'immune receptor activity,' 'cytokine receptor activity,' 'chemokine receptor binding' and 'chemokine activity' were the top five GO terms on MF. Based on the KEGG analysis, significant enrichments were observed in nine KEGG pathways for the apparent DE-mRNAs, such as 'hematopoietic cell lineage,' 'cytokine-cytokine receptor interaction,' 'viral protein interaction with cytokine and cytokine receptor,' 'chemokine signaling pathway' and 'Th17 cell differentiation' (Fig. 4B and Table SII).

Analysis of DE-mRNAs and screening of hub mRNAs. To further explore the functions of these mRNAs, the STRING database was analyzed using a PPI network analysis and mRNAs with the highest confidence levels were selected. Overall, 195 DE-mRNAs were filtered into a PPI network complex containing 181 nodes and 781 edges. Then, MCODE was used to identify the most connected hub mRNAs in the PPI

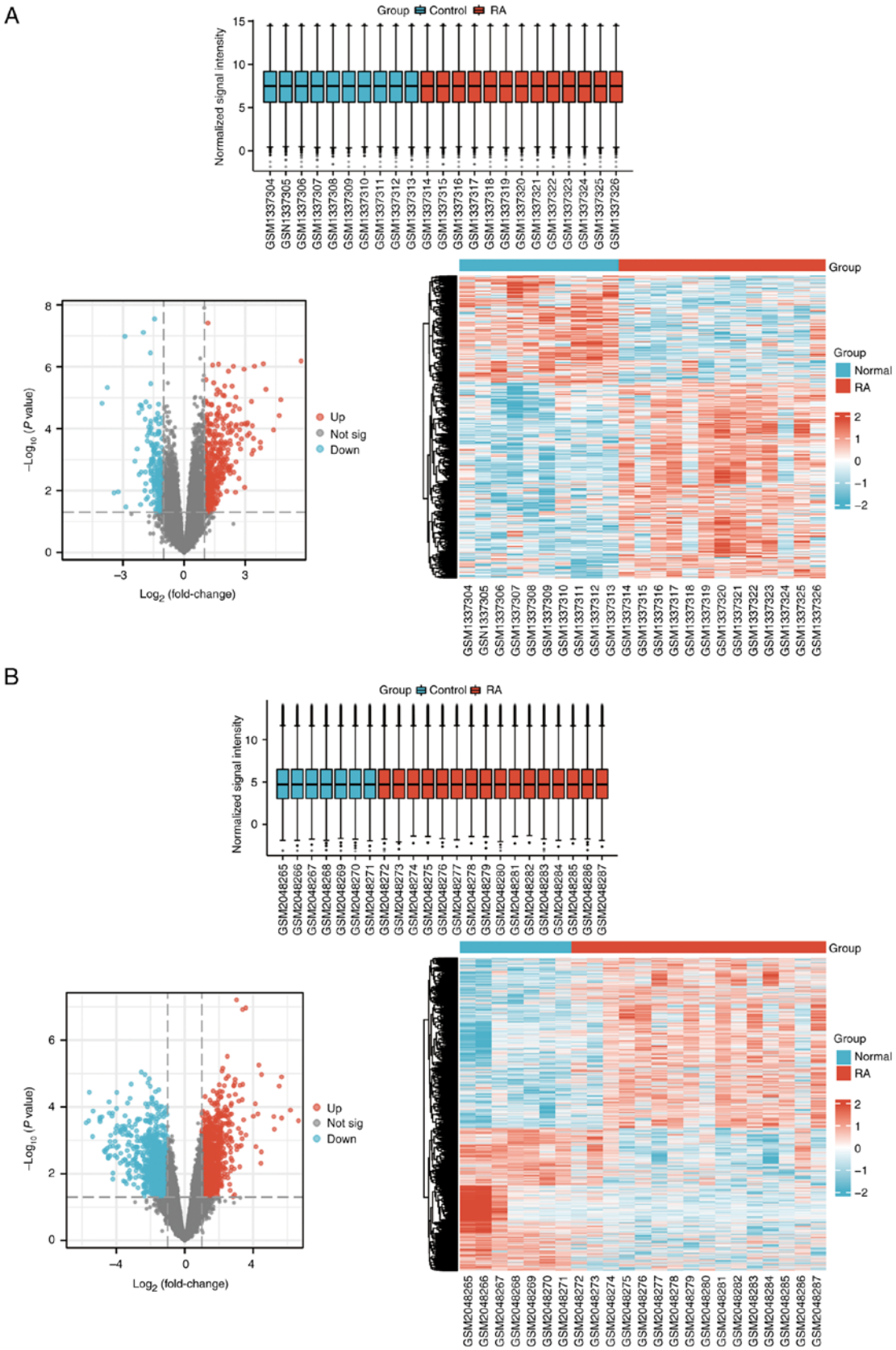

Figure 2. Screening of DE-mRNAs from samples obtained from patients with RA and healthy controls from the GSE55457 and GSE77298 datasets. Boxplots, volcano plots and heatmaps of the DE-mRNAs from the (A) GSE55457 and (B) GSE77298 datasets. DE-mRNA, differentially expressed mRNA; RA, rheumatoid arthritis.

network (Fig. 5A and B). Then, cytoHubba was used to analyze the top 10 most interacting central mRNAs: Assembly factor for spindle microtubules (AS*PM)*, *DLGAP5*, DNA topoisomerase IIα *(TOP2A)*, *KIF20A*, *BUB1*, *KIF11*, *CDC20*, *MELK*,

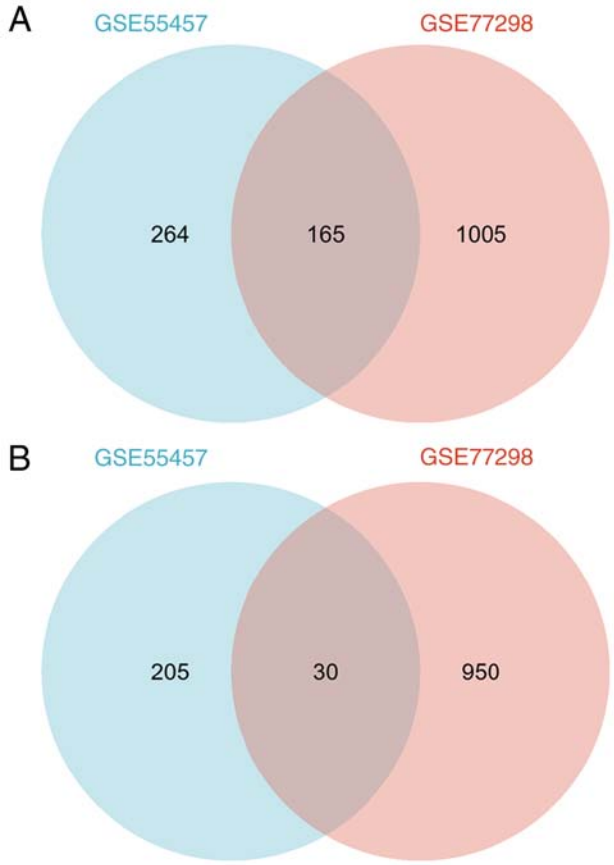

Figure 3. Identification of shared DEGs in the GSE55457 and GSE77298 datasets. (A) DEGs with upregulated expression in both the GSE55457 and GSE77298 datasets. (B) DEGs with downregulated expression in both the GSE55457 and GSE77298 datasets. DEGs, differentially expressed genes; RA, rheumatoid arthritis.

NDC80 and ribonucleotide reductase regulatory subunit M2 *(RRM2)* (Fig. 5C).

Validation of the FCA-induced arthritis rat model. A schematic diagram of the flowchart after the establishment of the FCA-induced arthritis rat model is shown in Fig. S1A. The rat footpad thickness was measured and the thickness in the model group was significantly increased compared with the control group (P<0.01; Fig. S1B). H&E staining of the joint tissue demonstrated that, compared with the control group, a greater amount of inflammatory cell infiltration, macrophage hyperplasia and fibrous tissue hyperplasia was present in the model group (Fig. S1C). The groups were then scored blindly with all criteria tested by at least two investigators and the average scores were obtained. The scores represent the degree of tissue damage with a higher score indicating a higher degree of tissue damage. Compared with the control group, the joints of the rats in the model group showed apparent inflammatory cell infiltration, fibrous tissue hyperplasia and macrophage proliferation (P<0.01; Fig. S1D-F). Therefore, the results demonstrated that, compared with the control group, the RA group had a higher degree of bone tissue destruction and inflammation, indicating that the FCA-induced RA rat model was successfully established.

Validation of four DE-mRNAs in the FCA-induced arthritis rat model. The mRNA expression levels of four previously unreported genes (*DLGAP5*, *KIF20A*, *MELK* and *NDC80*) among the top 10 central DE-mRNAs (or hub genes) from the PPI network analysis were examined in RA rat spleen tissues using qPCR, to validate the reliability of the dataset. Compared with the control group, the expression levels of *DLGAP5*, *KIF20A*, *MELK* and *NDC80* in the RA group were significantly increased (P<0.01; Fig. 6A-D). In addition, western blotting was performed to detect the protein expression levels of DLGAP5, KIF20A, MELK and NDC80. Compared with the control group, the protein expression levels of DLGAP5, KIF20A, MELK and NDC80 in the RA group were significantly increased (P<0.01; Fig. S2). Furthermore, ELISA results demonstrated that, compared with the control group, IL-12 and TNF-α levels in the RA group were significantly increased (P<0.01; Fig. 6E and F). In addition, correlation analysis was performed to analyze the correlation between inflammatory factors and hub genes, and the results demonstrated that the inflammatory factors were highly correlated with the hub genes (Fig. S3).

Discussion

RA is a chronic and systemic inflammatory disease that particularly affects the joints, leading to decreased physical function and even disability. Since the complex pathogenesis of RA is still unclear, it has become one of the important frontier research topics in the field of immunity. Previously, studies have focused on the role of inflammation in the occurrence of RA or the regulation of RA development and notable progress has been made in this research area (20-23). Moxibustion treatment. can effectively reduce synovial inflammation in RA rats by regulating the macrophage migration inhibitory factor/glucocorticoid signaling pathway (24). In addition, it can also improve RA symptoms by regulating the polarization of macrophage signaling pathways and exert a potentially bone protective effect (25,26). Chen *et al* (27) reported that receptor activator of NF-κB ligand (RANKL) plays an essential role in the orchestration of osteoclast maturation, activation, migration and survival. However, osteoprotegerin can inhibit RANKL activity to achieve bone homeostasis by inhibiting osteoclastogenesis and subsequent bone resorption. Thus, cartilage degeneration and bone destruction can be alleviated through the osteoprotegerin/RANKL signaling pathway. These findings suggested that pharmacological treatment or moxibustion treatment can effectively improve the joint swelling, inflammation level and bone destruction symptoms of RA. Targeted treatment of key genes is becoming increasingly important in RA. It has been reported that using (+)-JQ-1 as an inhibitor can displace bromodomain-containing protein 2 (*BRD2*) and bromodomain-containing protein 4 (*BRD4*) from upstream regions of key super-enhancers, a unique group of cis-regulatory elements, associated with key genes controlling intracellular state, cell identity and cell-type-specific functions. Loss of BRD2 and BRD4 leads to downregulation of the expression of genes that are critically involved in the pathogenesis and progression of RA (28). Upon activation or high expression of four hub genes identified in the present study (*DLGAP5*, *KIF20A*, *MELK* and *NDC80*), growth, proliferation and migration of cells, including pro-inflammatory cells and macrophages, is promoted (29-39). Under the action

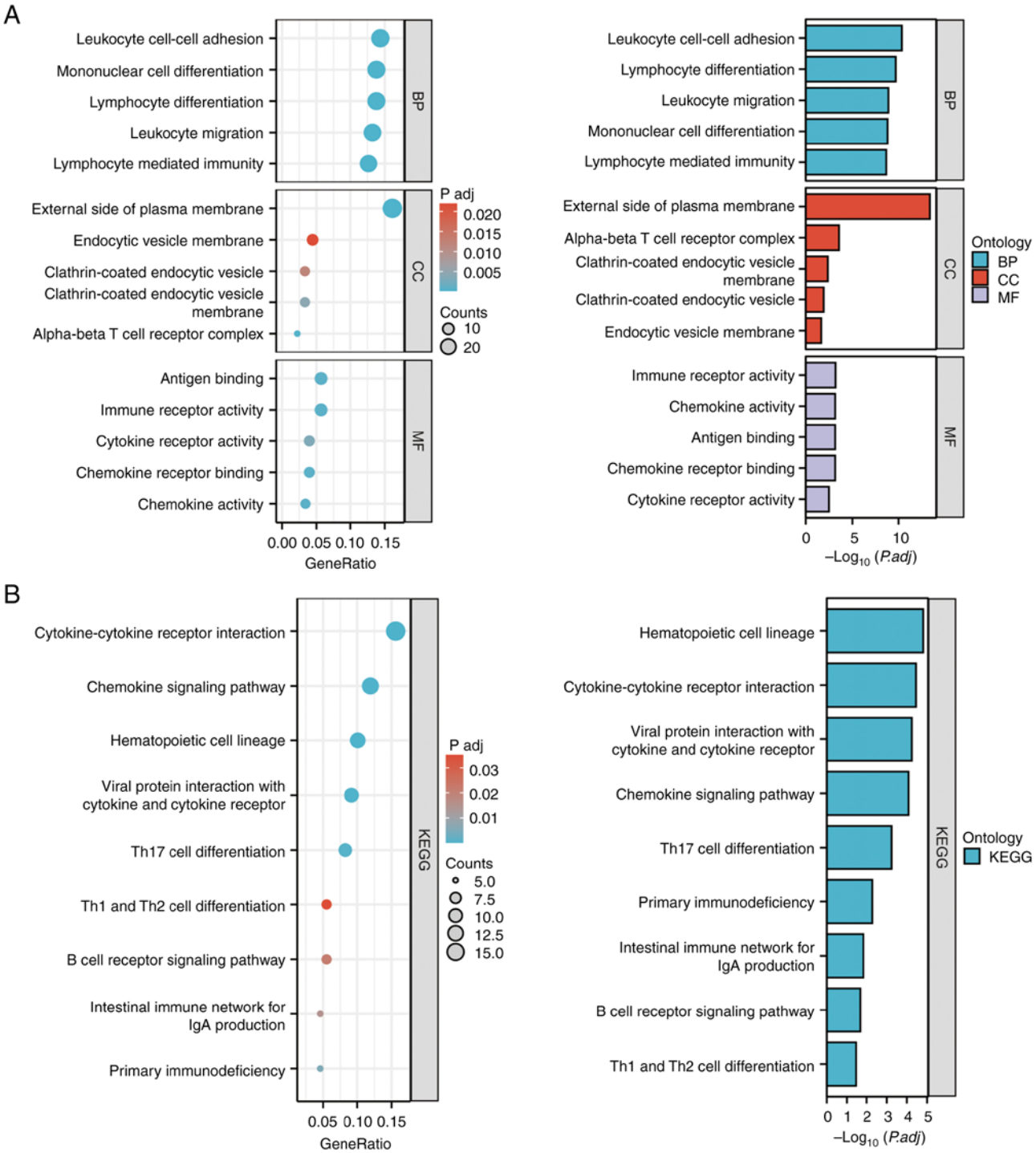

Figure 4. Screening of the DE-mRNAs and enrichment analysis. (A) Top five significant terms for BP, CC and MF in the GO enrichment analysis. (B) Top nine significance terms in the KEGG enrichment analysis. BP, biological process; CC, cellular component; DE-mRNA, differentially expressed messenger RNA; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function; Th, T helper.

of these genes, pro-inflammatory cells and macrophages then migrate to the lesion site and the inflammatory cells begin to release a number of inflammatory factors (such as IL-12 and TNF-α), which causes systemic inflammation and macrophage infiltration (29,40,41). The balance between M1 and M2 macrophages is crucially involved in the pathogenesis of RA. M1 macrophages act as pro-inflammatory mediators in the synovium, whereas M2 macrophages suppress inflammation and promote tissue repair. Notably, an imbalance in the ratio of M1/M2 macrophages can induce the activation of pro-inflammatory cytokines, the production and release of chemokines, and promote the proliferation and survival of macrophages, eventually leading to the occurrence of RA (42-44).

Abnormal gene expression is closely associated with a variety of pathological conditions, such as RA. However, it remains unclear which genes play a key role in the onset and progression of this disease. In the present study, 429 RA-related DEGs with upregulated expression and 235 with

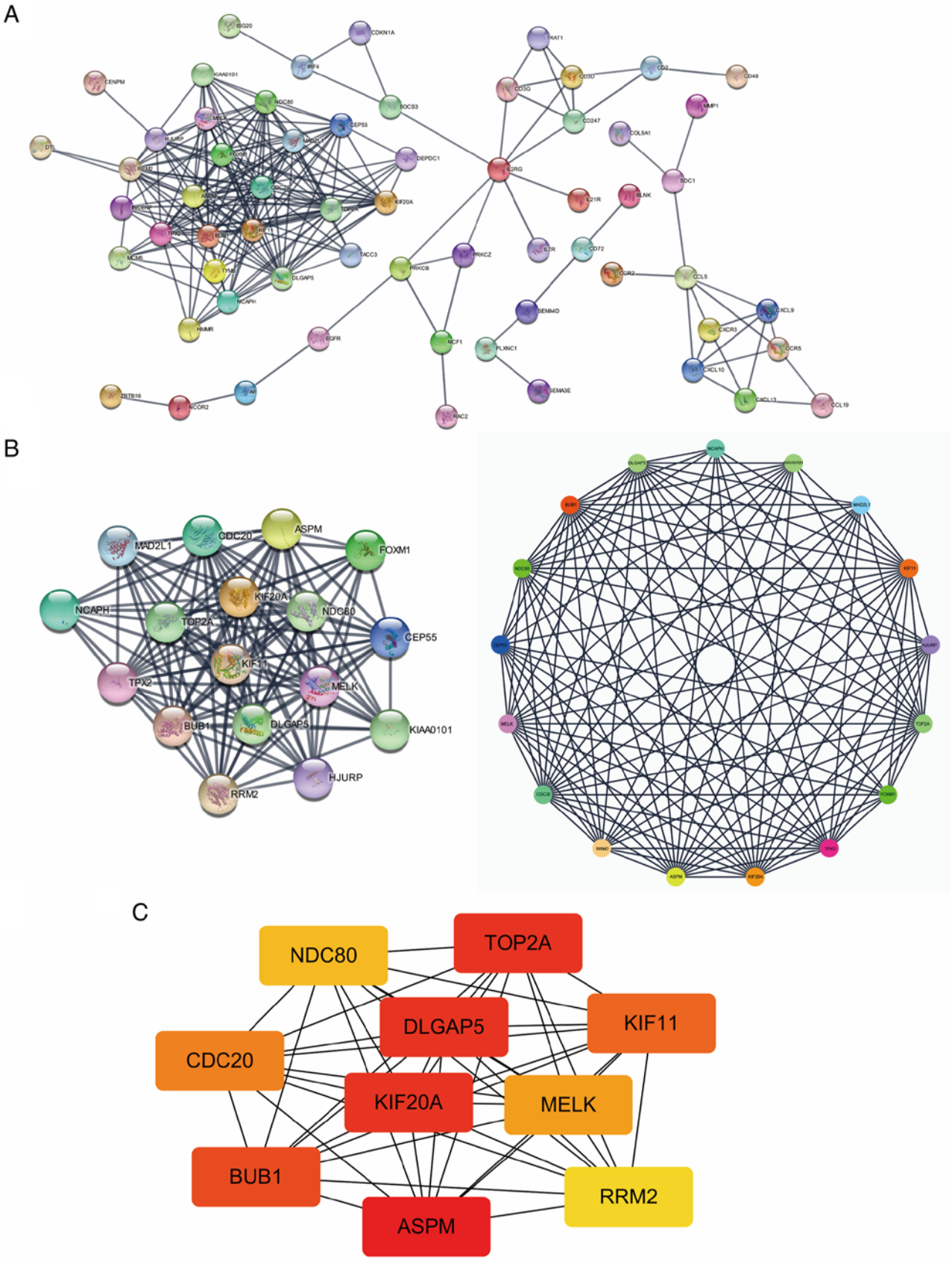

Figure 5. Identification of candidate overlapping mRNAs. (A) Protein-protein interaction network of overlapping mRNA. (B) Correlation diagram of the first module screened using MCODE. (C) Top 10 hub genes identified in overlapping mRNAs using cytoHubba.

downregulated expression were identified in the GSE55457 dataset and 1,170 RA-related DEGs with upregulated expression and 980 with downregulated expression were identified in the GSE77298 dataset. Meanwhile, 165 RA-related DEGs with

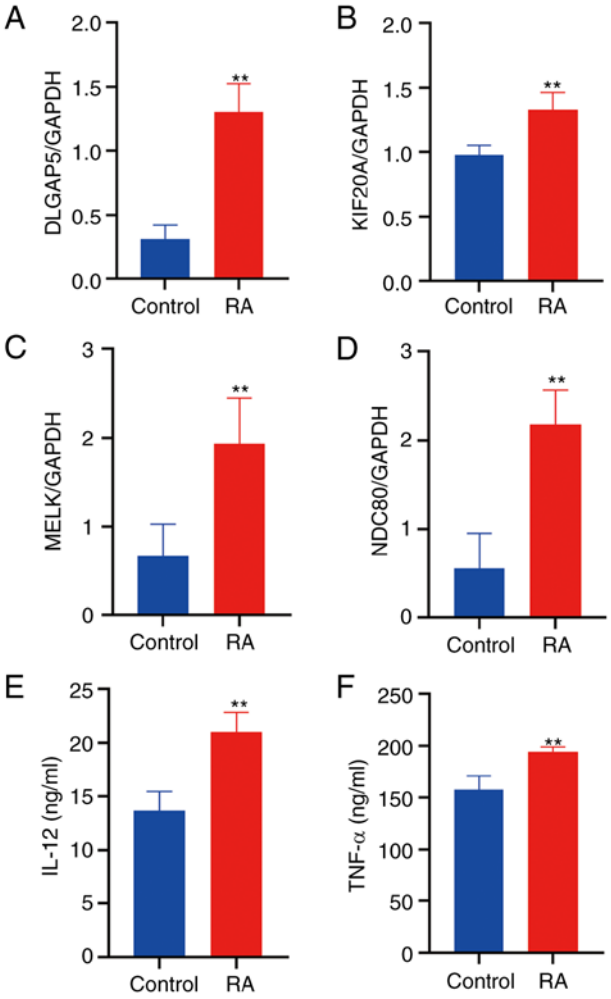

Figure 6. Validation of relative gene expression in the RA model using qPCR and ELISA. Relative expression of (A) *DLGAP5*, (B) *KIF20A*, (C) *MELK* and (D) *NDC80* in the FCA-induced and control rats was determined using qPCR. Levels of the inflammatory cytokines, (E) IL-12 and (F) TNF-α, in the FCA-induced and control rats were determined using ELISA. Data are presented as the mean ± standard deviation (n=6). **P<0.01 vs. the control group. FCA, Freund's complete adjuvant; RA, rheumatoid arthritis; IL-12, interleukin 12; TNF-α, tumor necrosis factor-α; DLGAP5, discs large homolog-associated protein 5; KIF20A, kinesin family member 20A; MELK, maternal embryonic leucine zipper kinase; NDC80, nuclear division cycle 80; qPCR, quantitative PCR; ELISA, enzyme-linked immunosorbent assay.

upregulated expression and 30 with downregulated expression were confirmed in the intersection of DEGs between the two datasets. These genes were associated with T helper (Th)1 and Th2 cell differentiation, chemokine signaling pathways, Th17 cell differentiation and the B-cell receptor signaling pathway. Among the DEGs, 10 hub genes associated with RA (*ASPM*, *DLGAP5*, *TOP2A*, *KIF20A*, *BUB1*, *KIF11*, *CDC20*, *MELK*, *NDC80* and *RRM2*) were identified. To the best of our knowledge, there is no clear understanding of the specific target gene or molecular biomarker responsible for the changes in the expression of the hub genes in the tissues of patients with RA. Biomarkers for RA and associated diseases include changes in the mRNA expression levels of genes, such as GBP1, CXCL10, GM-CSF and its receptor CSF2RB (6,7). In the present study, four new genes were selected as candidates from the identified hub genes, namely, *DLGAP5*, *KIF20A*, *MELK* and *NDC80*.

DLGAP5 was initially identified as a cell cycle-regulating protein. However, it is now regarded as a microtubule-associated protein, associated with the regulation of KIF18A located at the plus end of centromere microtubules, to stabilize K fibers and promote chromosome aggregation (32). Previous studies analyzed the structure and function of DLGAP5 in different species from physiological and clinicopathological perspectives, and reported that DLGAP5 plays an important role in promoting cell growth, proliferation and migration (30,31). The *KIF20A* gene encodes a protein with 890 amino acid residues, which belongs to the KIF superfamily and is mainly distributed in the central region of the mitotic spindle. Therefore, it participates in the process of driving mitosis and promoting cell proliferation (32,33). *MELK* is involved in various physiological and pathological processes of cells, such as cell proliferation, apoptosis and tumor development. In the last few decades, numerous studies have shown that *MELK* is significantly upregulated in a number of cancer types and plays an important role in promoting cell proliferation and maintaining cancer cell stemness (34,35). Silencing the *MELK* gene can slow down or even inhibit the proliferation, invasion and tumorigenicity of cancer cells (36,37). NDC80/Hec1, a subunit of the kinetochore complex (also called the NDC80 complex), constitutes and stabilizes microtubule-kinetochore attachment during the segregation of mitotic chromosomes, and plays an important role in mitosis and cell proliferation (38). NDC80 upregulation is also an important biomarker for cancer; high levels of NDC80 have been reported to indicate a poor prognosis in various types of cancer, such as gastric cancer, ovarian cancer and blood cancer (39). Although *DLGAP5*, *KIF20A*, *MELK* and *NDC80* have been studied separately in different tumors, to the best of our knowledge, there is still a lack of evidence for the correlation between *DLGAP5*, *KIF20A*, *MELK* and *NDC80* and RA development.

Previous studies have confirmed that RA can be treated in a number of different ways, such as by targeting signaling pathways, key proteins and circadian rhythms (18,22,24,26). Therefore, to further explore the potential biomarkers and target genes in RA, validation with FCA-induced RA rats was performed in the present study. As FCA-induced RA resembles RA in humans in terms of pathogenesis and clinical symptoms, it is widely used for research owing to its simplicity, reliability and reproducibility for RA modeling. As an immune organ, the spleen serves a regulatory role in systemic immunity. Therefore, in the present study, an RA rat model was established and the expression of *DLGAP5*, *KIF20A*, *MELK* and *NDC80* mRNA in the spleen of this model was examined. As expected, the results demonstrated that the expression levels of these genes were significantly increased in the RA model, consistent with the results obtained by bioinformatics analysis. In addition, the expression of representative inflammatory factors of RA (IL-12 and TNF-α) were verified in this model and these factors were correlated with the expression of the key genes (*DLGAP5*, *KIF20A*, *MELK* and *NDC80*). Thus, the present study demonstrated that these genes may play a crucial role in eliciting immune responses in RA.

In conclusion, four possible RA-related hub genes were identified in the present study, namely *DLGAP5*, *KIF20A*, *MELK* and *NDC80*, using the GEO database, and the expression of these four genes in an RA model were confirmed in

validation experiments. These findings provide new insights into the mechanism of RA development, and the upregulated expression of these four hub genes is closely related to inflammation in RA. The present study provides an improved understanding of RA mechanisms and may lead to the identification of potential biomarkers in the future.

Acknowledgements

The authors wish to thank the Animal Experiment Center of Chengdu University of Traditional Chinese Medicine (Chengdu, China) for the housing of the animals in this study.

Funding

This work was supported by The National Natural Science Foundation of China (grant no. 81973959), The National Key R&D Program of China (grant no. 2019YFC1709001), The Science and Technology Innovation Seedling Project of Sichuan Province (grant no. 2022037) and The Foundation of Sichuan Provincial Administration of Traditional Chinese Medicine (grant no. 2018JC007).

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Authors' contributions

KL drafted the manuscript, performed most of the experiments and analyzed the data. YDG, YMX and JWN conducted part of the animal experiments. HYZ and YMZ contributed to the design of the study, and were involved in proofreading and editing. HYZ and YMZ confirm the authenticity of all the raw data. All authors contributed considerably to this study. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

All animal procedures were approved by The Ethics Committee of Chengdu University of Traditional Chinese Medicine (Chengdu, China; approval no. 2018-11). The authors observed the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Giraud EL, Jessurun NT, van Hunsel FPAM, van Puijenbroek EP, van Tubergen A, Ten Klooster PM and Vonkeman HE: Frequency of real-world reported adverse drug reactions in rheumatoid arthritis patients. Expert Opin Drug Saf 19: 1617-1624, 2020.
2. Smith MH and Berman JR: What is rheumatoid arthritis? JAMA 327: 1194, 2022.
3. Smolen JS, Aletaha D and McInnes IB: Rheumatoid arthritis. Lancet 388: 2023-2038, 2016.
4. Gravallese EM and Firestein GS: Rheumatoid arthritis-common origins, divergent mechanisms. N Engl J Med 388: 529-542, 2023.
5. Aletaha D, Smolen JS: Diagnosis and management of rheumatoid arthritis: A review. JAMA 320: 1360-1372, 2018.
6. Woetzel D, Huber R, Kupfer P, Pohlers D, Pfaff M, Driesch D, Häupl 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 16: R84, 2014.
7. Broeren MG, de Vries M, Bennink MB, Arntz OJ, Blom AB, Koenders MI, van Lent PL, van der Kraan PM, van den Berg WB and van de Loo FA: Disease-regulated gene therapy with anti-inflammatory interleukin-10 under the control of the CXCL10 promoter for the treatment of rheumatoid arthritis. Hum Gene Ther 27: 244-254, 2016.
8. Davis S and Meltzer PS: GEOquery: A bridge between the gene expression omnibus (GEO) and BioConductor. Bioinformatics 23: 1846-1847, 2007.
9. Smyth GK: Limma: Linear models for microarray data. In: Bioinformatics and Computational Biology Solutions using R and Bioconductor. Springer, New York, NY, 397-420, 2005.
10. Yu G, Wang LG, Han Y and He QY: clusterProfiler: An R package for comparing biological themes among gene clusters. OMICS 16: 284-287, 2012.
11. R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2023. https://www.R-project.org/
12. Lu S, Yadav AK and Qiao X: Identification of potential miRNA-mRNA interaction network in bone marrow T cells of acquired aplastic anemia. Hematology 25: 168-175, 2020.
13. Liu S, Chen H, Ma W, Zhong Y, Liang Y, Gu L, Lu X and Li J: Non-coding RNAs and related molecules associated with form-deprivation myopia in mice. J Cell Mol Med 26: 186-194, 2022.
14. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY: cytoHubba: Identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 8 (Suppl 4): S11, 2014.
15. Otasek D, Morris JH, Bouças J, Pico AR and Demchak B: Cytoscape automation: Empowering workflow-based network analysis. Genome Biol 20: 185, 2019.
16. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, *et al*: Correction to 'The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets.' Nucleic Acids Res 49: 10800, 2021.
17. Li A, Zhang Z, Ru X, Yi Y, Li X, Qian J, Wang J, Yang X and Yao Y: Identification of SLAMF1 as an immune-related key gene associated with rheumatoid arthritis and verified in mice collagen-induced arthritis model. Front Immunol 13: 961129, 2022.
18. Zhong YM, Zhang LL, Lu WT, Shang YN and Zhou HY: Moxibustion regulates the polarization of macrophages through the IL-4/STAT6 pathway in rheumatoid arthritis. Cytokine 152: 155835, 2022.
19. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408, 2001.
20. Chen Y, Liu K, Qin Y, Chen S, Guan G, Huang Y, Chen Y and Mo Z: Effects of Pereskia aculeate Miller petroleum ether extract on complete Freund's adjuvant-induced rheumatoid arthritis in rats and its potential molecular mechanisms. Front Pharmacol 13: 869810, 2022.
21. Pan H, Guo R, Ju Y, Wang Q, Zhu J, Xie Y, Zheng Y, Li T, Liu Z, Lu L, *et al*: A single bacterium restores the microbiome dysbiosis to protect bones from destruction in a rat model of rheumatoid arthritis. Microbiome 7: 107, 2019.
22. Han C, Yang Y, Sheng Y, Wang J, Zhou X, Li W, Guo L, Zhang C and Ye Q: Glaucocalyxin B inhibits cartilage inflammatory injury in rheumatoid arthritis by regulating M1 polarization of synovial macrophages through NF-κB pathway. Aging (Albany NY) 13: 22544-22555, 2021.
23. Kumar H and Bot A: In this issue: Role of immune cells and molecules in rheumatoid arthritis pathogenesis and cancer immunotherapy. Int Rev Immunol 37: 127-128, 2018.
24. Zhang L, Zhong Y, Lu W, Shang Y, Guo Y, Luo X, Chen Y, Luo K, Hu D, Yu H, *et al*: Moxibustion of Zusanli (ST36) and Shenshu (BL23) alleviates the inflammation of rheumatoid arthritis in rats through regulating macrophage migration inhibitory factor/glucocorticoids signaling. J Tradit Chin Med (In Chinese).

25. Su H, Su SY, Yang P, Guo YJ and Li J: Progress of research on mechanism of acupuncture and moxibustion in the treatment of rheumatoid arthritis. Zhen Ci Yan Jiu 48: 500-507, 2023 (In Chinese).
26. Zhong YM, Wu F, Luo XC, Chen Y, Ren JG, Yang X, Ma WB and Zhou HY: Mechanism on moxibustion for rheumatoid arthritis based on PD-1/PD-L1 signaling pathway. Zhongguo Zhen Jiu 40: 976-982, 2020 (In Chinese).
27. Chen Y, Li H, Luo X, Liu H, Zhong Y, Wu X and Liu X: Moxibustion of Zusanli (ST36) and Shenshu (BL23) alleviates cartilage degradation through RANKL/OPG signaling in a rabbit model of rheumatoid arthritis. Evid Based Complement Alternat Med 2019: 6436420, 2019.
28. Krishna V, Yin X, Song Q, Walsh A, Pocalyko D, Bachman K, Anderson I, Madakamutil L and Nagpal S: Integration of the transcriptome and genome-wide landscape of BRD2 and BRD4 binding motifs identifies key superenhancer genes and reveals the mechanism of bet inhibitor action in rheumatoid arthritis synovial fibroblasts. J Immunol 206: 422-431, 2021.
29. Tang N, Dou X, You X, Shi Q, Ke M and Liu G: Pan-cancer analysis of the oncogenic role of discs large homolog associated protein 5 (DLGAP5) in human tumors. Cancer Cell Int 21: 457, 2021.
30. Li K, Fu X, Wu P, Zhaxi B, Luo H and Li Q: DLG7/DLGAP5 as a potential therapeutic target in gastric cancer. Chin Med J 135: 1616-1618, 2022.
31. Zhang H, Liu Y, Tang S, Qin X, Li L, Zhou J, Zhang J and Liu B: Knockdown of DLGAP5 suppresses cell proliferation, induces G2/M phase arrest and apoptosis in ovarian cancer. Exp Ther Med 22: 1245, 2021.
32. Ren X, Chen X, Ji Y, Li L, Li Y, Qin C and Fang K: Upregulation of KIF20A promotes tumor proliferation and invasion in renal clear cell carcinoma and is associated with adverse clinical outcome. Aging (Albany, NY) 12: 25878-25894, 2020.
33. Wu C, Qi X, Qiu Z, Deng G and Zhong L: Low expression of KIF20A suppresses cell proliferation, promotes chemosensitivity and is associated with better prognosis in HCC. Aging (Albany, NY) 13: 22148-22163, 2021.
34. Zhang X, Wang J, Wang Y, Liu G, Li H, Yu J, Wu R, Liang J, Yu R and Liu X: MELK inhibition effectively suppresses growth of glioblastoma and cancer stem-like cells by blocking AKT and FOXM1 pathways. Front Oncol 10: 608082, 2020.
35. Ye J, Deng W, Zhong Y, Liu H, Guo B, Qin Z, Li P, Zhong X and Wang L: MELK predicts poor prognosis and promotes metastasis in esophageal squamous cell carcinoma via activating the NF-κB pathway. Int J Oncol 61: 94, 2022.
36. Tang B, Zhu J, Liu F, Ding J, Wang Y, Fang S, Zheng L, Qiu R, Chen M, Shu G, *et al*: xCT contributes to colorectal cancer tumorigenesis through upregulation of the MELK oncogene and activation of the AKT/mTOR cascade. Cell Death Dis 13: 373, 2022.
37. Xu Q, Ge Q, Zhou Y, Yang B, Yang Q, Jiang S, Jiang R, Ai Z, Zhang Z and Teng Y: MELK promotes endometrial carcinoma progression via activating mTOR signaling pathway. EBioMedicine 51: 102609, 2020.
38. Chen X, He Q, Zeng S and Xu Z: Upregulation of nuclear division cycle 80 contributes to therapeutic resistance via the promotion of autophagy-related protein-7-dependent autophagy in lung cancer. Front Pharmacol 13: 985601, 2022.
39. Chen J and Ünal E: Meiotic regulation of the Ndc80 complex composition and function. Curr Genet 67: 511-518, 2021.
40. Huang J, Zheng M, Li Y, Xu D and Tian D: DLGAP5 promotes gallbladder cancer migration and tumor-associated macrophage M2 polarization by activating cAMP. Cancer Immunol Immunother: Jul 8, 2023 (Epub ahead of print).
41. Zhang Z, Sun C, Li C, Jiao X, Griffin BB, Dongol S, Wu H, Zhang C, Cao W, Dong R, *et al*: Upregulated MELK leads to doxorubicin chemoresistance and M2 macrophage polarization via the miR-34a/JAK2/STAT3 pathway in uterine leiomyosarcoma. Front Oncol 10: 453, 2020.
42. Chung SJ, Yoon HJ, Youn H, Kim MJ, Lee YS, Jeong JM, Chung JK, Kang KW, Xie L, Zhang MR and Cheon GJ: 18F-FEDAC as a targeting agent for activated macrophages in DBA/1 mice with collagen-induced arthritis: Comparison with ^{18}F-FDG. J Nucl Med 59: 839-845, 2018.
43. Park SY, Lee SW, Lee SY, Hong KW, Bae SS, Kim K and Kim CD: SIRT1/adenosine monophosphate-activated protein kinase α signaling enhances macrophage polarization to an anti-inflammatory phenotype in rheumatoid arthritis. Front Immunol 8: 1135, 2017.
44. Cutolo M, Campitiello R, Gotelli E and Soldano S: The role of M1/M2 macrophage polarization in rheumatoid arthritis synovitis. Front Immunol 13: 867260, 2022.