

# circPTPN22 as a novel biomarker and ceRNA in peripheral blood mononuclear cells of rheumatoid arthritis

ZHUYAN JIANG<sup>1,2</sup>, ZHITING ZHONG<sup>1</sup>, QINGQING MIAO<sup>1</sup>, YUE ZHANG<sup>1,2</sup>,  
BING NI<sup>3</sup>, MENGJIE ZHANG<sup>3</sup> and JUN TANG<sup>1</sup>

<sup>1</sup>Dermatology Department of The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China; <sup>2</sup>Department of Dermatology, The 901st Hospital of The Joint Logistics Support Force of PLA Affiliated to Anhui Medical University, Hefei, Anhui 230001; <sup>3</sup>Department of Pathophysiology, Third Military Medical University, Chongqing 400038, P.R. China

Received December 26, 2020; Accepted May 13, 2021

DOI: 10.3892/mmr.2021.12256

**Abstract.** The role of circular RNAs (circRNAs) in rheumatoid arthritis (RA) remains to be elucidated. To determine the expression of circRNAs in peripheral blood mononuclear cells (PBMCs) and to identify novel biomarkers for RA and explore their potential effects in RA, the present study conducted high-throughput RNA sequencing to analyze circRNA expression profiles in PBMCs from 4 RA patients and 3 healthy controls (HCs). Reverse transcription-quantitative PCR was used to verify the expression of circPTPN22 in 42 RA patients, 44 HCs and 45 systemic lupus erythematosus (SLE) patients. In addition, bioinformatics analysis and Pearson's correlation test were conducted to assess the correlation of the relationships between circPTPN22 and RA progression. A receiver operating characteristic curve was calculated to evaluate the diagnostic value. Multilevel integrated analysis identified 41 upregulated and 30 downregulated circRNAs in RA patients compared with HCs. circPTPN22 was confirmed to be a common differentially expressed gene in RA and SLE compared with HCs. Area under the curve analysis suggested the diagnostic value of circPTPN22 expression to distinguish RA patients from both HCs and SLE patients. In addition, circPTPN22 levels in RA PBMCs were correlated with RA-IgG, RA-IgM, RA-IgA, anti-cyclic citrullinated peptide

(anti-CCP), rheumatoid factor and C reactive protein levels. A total of four putative microRNAs (miRNAs or miRs), namely, hsa-miR-3074-5p, hsa-miR-373-3p, hsa-miR-766-3p and hsa-miR-34c-5p, were screened to be sponged by circPTPN22 via bioinformatics analysis and then experimentally verified to be upregulated in RA PBMCs compared with controls. The data suggested that circPTPN22 might be a novel biomarker for the diagnosis of RA and participate in RA pathogenesis through a sponge mechanism.

## Introduction

Rheumatoid arthritis (RA) is a chronic disease characterized by persistent synovitis involving multiple joints, systemic inflammation and autoantibodies, especially rheumatoid factor (RF) and anti-cyclic citrullinated peptide (anti-CCP), that can eventually lead to cartilage, bone damage and disability (1). A study from the United Kingdom shows that RA is more prevalent in women than in men (2). Studies have identified that genetic factors, environmental factors and lifestyle are high-risk factors for RA, such as smoking, drinking and poor diet (3,4). However, the clear pathogenesis remains to be elucidated. Early diagnosis of RA is critical for the active treatment of irreversible damage to joints or organs and the control of chronic inflammation through the widespread use of conventional drugs and disease-modifying anti-rheumatic drugs (5). However, some RA patients whose RF and anti-CCP are both negative could be misdiagnosed and the belated treatment may cause deterioration of clinical outcome (6). Therefore, it is necessary to discover novel biomarkers to improve the diagnosis and prognosis of RA patients.

Circular RNA (circRNA) was originally found in pathogens and characterized by single-stranded covalent closure and has been considered an aberrant splicing byproduct of endogenous RNA (7,8). circRNAs with unique 'head-to-tail' splice junctions possess the characteristics of wide expression, high conservation and stability and tissue- or cell-specific expression, which determine their important noncoding functions (9,10). A number of biological functions of circRNAs have been demonstrated to act as microRNAs (miRNA) sponges (11), interact with a number of different RBPs to

*Correspondence to:* Dr Mengjie Zhang, Department of Pathophysiology, Third Military Medical University, 30 Gaotanyan Street, Shapingba, Chongqing 400038, P.R. China  
E-mail: zhangmengjie0406@tmmu.edu.cn

Professor Jun Tang, Dermatology Department of The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, 1 Swan Lake Road, Hefei, Anhui 230001, P.R. China  
E-mail: tangjun0212@ahmu.edu.cn

**Key words:** RNA sequencing, circular RNAs, competing endogenous RNAs, rheumatoid arthritis, systemic lupus erythematosus

act as protein sponges to enhance protein function or recruit proteins to specific sites (12,13) and regulate the transcription of maternal genes (14). In addition, circRNAs can perform cap-independent translation (15,16). Based on the above characteristics, circRNA has been verified to be involved in the development and progression of a number of types of cancer, such as cervical and gastric cancer (17,18) or participating in the regulation of autoimmune disease (19,20).

In our previous study, circPTPN22 was screened as a potential marker in the peripheral blood mononuclear cells (PBMCs) of systemic lupus erythematosus (SLE) patients by high-throughput RNA sequencing (21). The parental gene *protein tyrosine phosphatase nonreceptor 22 (PTPN22)* of circPTPN22, located on chromosome 1P13, is a member of the protein tyrosine phosphatase family, encoding lymphoid protein tyrosine phosphatase (22). It is well known that PTPN22 is closely related to the occurrence and development of autoimmune diseases, such as RA, SLE and type 1 diabetes, by modulating signaling pathways through antigen and innate immune receptors. Also, it serves roles in lymphocyte development and activation, establishment of tolerance and innate immune cell mediated host defense and immunoregulation (23,24). Though circPTPN22 was demonstrated to be a potential biomarker for SLE diagnosis in our previous study (21), whether circPTPN22 also presents a diagnostic value in RA patients remains to be elucidated.

To this end, the present study detected the differentially expressed circRNAs in PBMCs of RA patients and HC controls with high-throughput RNA sequencing and found that the differentially expressed circRNAs in RA patients indeed included circPTPN22. Further experiments verified the downregulated circPTPN22 expression in RA patients and the statistical analysis revealed the possibility of circPTPN22 as a diagnostic biomarker to distinguish RA patients from HC controls and even from SLE patients. In addition, the expression level of circPTPN22 was markedly related to the classification of RA and the levels of autoantibodies in RA patients.

## Materials and methods

**Patients.** A total of 46 patients with RA aged 28-76 years old were recruited from the First Affiliated Hospital of Third Military Medical University (Chongqing, China) between March 2019 and November 2020. RA disease activity was defined by the disease activity score 28 (DAS28) (25). In addition, the present study included 47 healthy controls (HCs) aged 27-76 years, who were free from autoimmune or inflammatory diseases and unrelated to the patients, after receiving a medical examination at the First Affiliated Hospital of Third Military Medical University. As an autoimmune disease control, 45 SLE patients diagnosed by the American College of Rheumatology revised criteria (26) were also enrolled from the same hospital at the research stage. The clinical characteristics of the participants in the validation phase are presented in Table I. The present study was approved and supervised by the ethical committee of The First Affiliated Hospital of Third Military Medical University (Hefei, China; approval no. KY2019119). Written informed consent was obtained from all subjects prior to participation.

**PBMC preparation and total RNA extraction.** Whole blood (~6 ml) was obtained from each subject before the experiment and stored in an ethylenediaminetetraacetic acid anticoagulant tube. Fresh PBMCs from each donor blood were isolated by Ficoll-Hypaque (Beijing Solarbio Science & Technology Co., Ltd.) density-gradient centrifugation (6,500 x g for 20 min) at room temperature. Then, total RNA was extracted by using TRIzol (Thermo Fisher Scientific, Inc.) according to the manufacturer's protocol. Subsequently, RNA concentration and purity were measured by a NanoDrop spectrophotometer (Invitrogen; Thermo Fisher Scientific, Inc.) and 1% agarose gel. RNA was stored in -80°C for further experiments.

**High-throughput RNA sequencing.** RNA with RNA integrity number >8.0 was used for the rRNA depletion library (VAHTSTM Total RNA-seq (H/M/R); Agilent Technologies, Inc.) according to the manufacturer's instructions. Whole transcriptome sequencing was carried out by NovelBio Corp. Laboratory using a HiSeq XTEN Sequencer (Illumina, Inc.). Prior to read mapping, clean reads were obtained from the raw reads by removing the adaptor sequences (reads with >5% ambiguous bases (noted as N) and low-quality reads containing >20% bases with qualities of <13). Then, the clean reads were aligned to the human reference genome sequence GRCh38 (NCBI) using the MapSplice program (netlab.uky.edu/p/bioinfo/MapSplice) V2.1.6 (27). The RNA-seq data have been submitted to the NCBI Gene Expression Omnibus (GEO), under GEO accession number GSE169080 and GSE169082 (ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE169080 and ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE169082).

**Computational bioinformatics analysis.** HTseq (huber.embl.de/HTseq; V2.05) was used to count genes and the reads per kilo base per million mapped reads method was used to normalize the gene expression level (28). To identify differentially expressed genes, the DEseq (29) algorithm was used to identify the differentially expressed genes under the criteria of fold change >2 and false discovery rate <0.05. Gene Ontology (GO; <http://www.geneontology.org/>) was used to clarify the biological significance of differentially expressed genes and Kyoto Encyclopedia of Genes and Genomes (KEGG; [www.genome.jp/kegg/](http://www.genome.jp/kegg/)) enrichment analysis and gene set enrichment analysis (GSEA) were performed to identify important pathways.

**Reverse transcription-quantitative (RT-q) PCR.** Total RNA from PBMCs was reverse-transcribed into cDNA using a Prime-Script RT reagent kit with gDNA Eraser according to the manufacturer's protocol (Takara Bio, Inc.). RT-qPCR was performed using cDNA and SYBR Premix Ex Taq II (Takara Biotechnology Co., Ltd.) using CFX96 Thermal Cycler equipment (Bio-Rad Laboratories, Inc.) following the manufacturer's instructions. The conditions were: 95°C for 30 sec, followed by 39 cycles of 95°C for 15 sec and 60°C for 30 sec. The primers used for quantitative real-time PCR were as follows: circPTPN22: F: 5'-AATTCTCACCAATGTTC CCA-3', R: 5'-AAGGTACATCATGGTCTGGC-3'; GAPDH: F: 5'-GGAGTCCACTGGCGTCTTC-3', R: 5'-GCTGATGAT CTTGAGGCTGTTG-3'. The levels of GAPDH were used to normalize the relative expression levels of circRNA.

Table I. Clinical characteristics of the study population in the validation phase.

Index	RA	Health control	SLE
Sex (male/female)	4/38	5/39	5/40
Age (years)	56.47±12.07	54.57±14.27	53±9.87
Duration (years)	15.10±10.67	N/A	N/A
TNF- $\alpha$ (ng/l)	14.96±6.14	N/A	N/A
IL-6 (pg/ml)	21.21±19.17	N/A	N/A
C3 (g/l)	1.07±0.20	N/A	N/A
C4 (g/l)	0.21±0.07	N/A	N/A
RA-IgG (IU/ml)	201.103±91.77	N/A	N/A
RA-IgM (IU/ml)	289.58±166.40	N/A	N/A
RA-IgA (IU/ml)	327.21±112.02	N/A	N/A
IgA (g/l)	4.59±3.88	N/A	N/A
IgG (g/l)	14.01±4.84	N/A	N/A
IgM (g/l)	2.02±0.71	N/A	N/A
IgE (g/l)	37.515±19.49	N/A	N/A
RF (IU/ml)	530.827±378.19	N/A	N/A
Anti-CCP (U/ml)	477.97±207.23	N/A	N/A
TJC	13.96±8.05	N/A	N/A
SJC	10.12±6.72	N/A	N/A
ESR (mm/h)	68.08±27.91	N/A	N/A
CRP (mg/l)	51.27±29.36	N/A	N/A
DAS28 score	5.59±1.39	N/A	N/A
SLEDAI	N/A	N/A	9.64±4.66

RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; RF, rheumatoid factors; anti-CCP, anti-cyclic citrullinated peptide; TJC, tender joint counts; SJC, swollen joint counts; ESR, erythrocyte sedimentation rate; CRP, C-reactive protein; DAS28, disease activity score 28; SLEDAI, systemic lupus erythematosus disease activity index; 0-4 score (essentially inactive group), 5-9 score (mild activity group) and score >10 (moderate to severe activity group).

The forward primers of microRNA and U6 used in the present study were synthesized by Tiangen Biotech Co., Ltd. and U6 small nuclear RNA was used as an endogenous control for miRNA normalization. The universal reverse primer was provided in the kit (Tiangen Biotech Co., Ltd.). Then, Ct values were obtained and calculated them by the  $2^{-\Delta\Delta C_t}$  method (30) for each sample. Each sample was repeated more than three times independently.

**Competing endogenous (Ce)RNA regulatory network construction and hub gene identification.** RNA pairs were predicted using the databases of RNAhybrid (bibiserv.techfak.uni-bielefeld.de/rnahybrid) (31) and miRanda (<http://www.microrna.org/microrna/home.do>). Based on the predicted binding sites of differentially expressed genes, a coexpression network of circRNA-miRNA-mRNA was established using Cytoscape 3.7.2 (32). By setting the interaction confidence score at the highest level of 0.9, the protein-protein interaction (PPI) network was constructed by the STRING database (<https://string-db.org/>). Cytoscape was used to visualize the PPI network and the hub genes were identified by the 11 topological analysis methods on the CytoHubba Plug-in (Degree  $\geq 36$ ) (33).

**Statistical analysis.** Fisher's exact test and  $\chi^2$  test were applied to identify the significant GO categories and

pathway analysis. One-way ANOVA and nonparametric tests, corrected for multiple comparisons using Tukey's test, were used to evaluate the expression of normally distributed parameters and skewed distribution parameters among the three groups. In correlation analysis, after testing by the one-sample Kolmogorov-Smirnov test, the data were analyzed by Pearson correlation test in case of a normal distribution among groups; otherwise, the data were analyzed by Spearman rank correlation among groups. Receiver operating characteristic (ROC) curves were used to analyze the sensitivity, specificity and area under the curve (AUC) to indicate the diagnostic value with the 95% CI. All the data were analyzed using SPSS statistical software (version 26.0; IBM Corp.).  $P < 0.05$  was considered to indicate a statistically significant difference.

## Results

**Expression profiling of circRNA and mRNA in PBMCs of patients with RA.** To understand the expression profiles of circRNAs and mRNAs in RA, RNA sequence analysis was applied to the PBMCs of 4 patients with RA and 3 healthy controls. A total of 25,646 circRNAs were detected by RNA-Seq and were widely distributed on all chromosomes, including sex chromosomes X and Y (Fig. 1A). There were

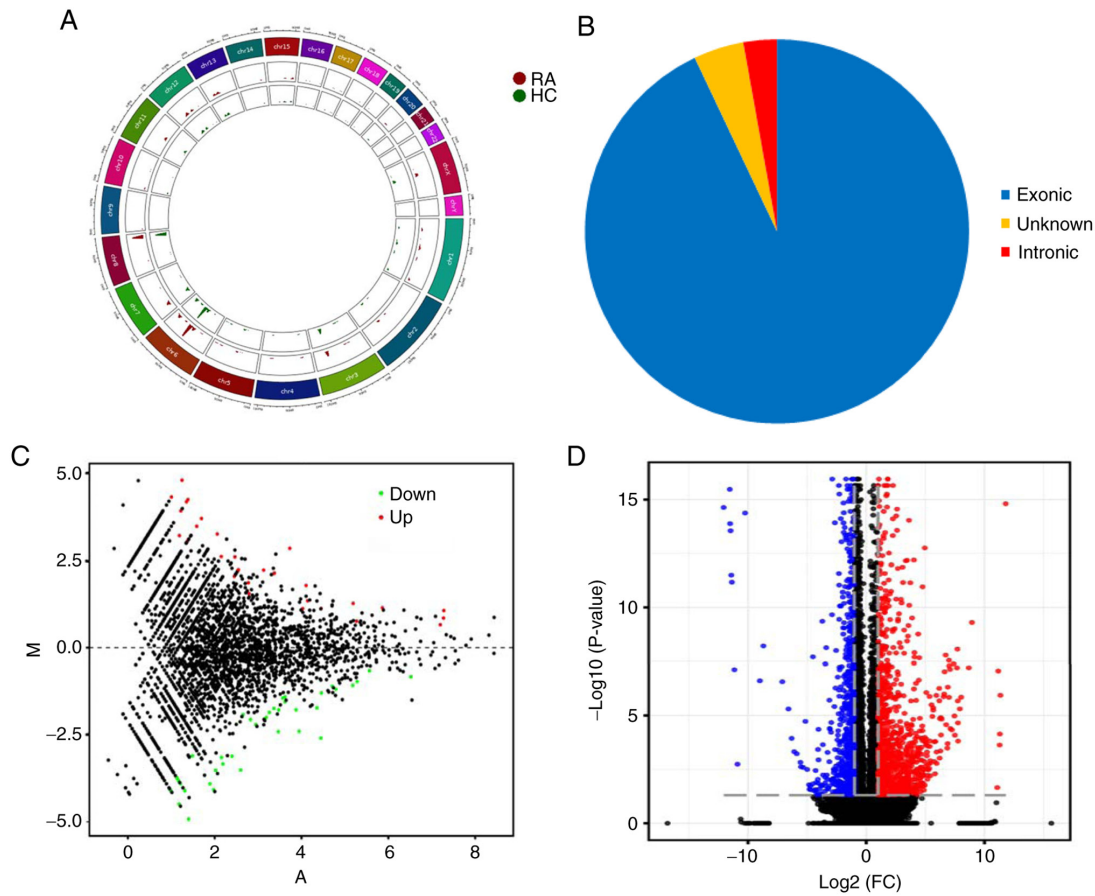


Figure 1. Expression profiling of circRNA and mRNA in PBMCs of patients with RA. (A) Circos plot showing the locations of circRNAs on human chromosomes. The outermost layer was a chromosome map of the human genome. The inner 2 circles represent all circRNAs of RA and HC samples detected by RNA-seq. The bar chart presents the expression levels of circRNA. (B) Classification of the significantly dysregulated circRNAs based on genomic origin. (C) MA plot showing significantly differentially expressed genes across the range of mean values between RA and HC. The x-axis is A (mean average) scales and the y-axis is M (log<sub>2</sub>-fold change). The red points represent significantly upregulated circRNAs and the green points represent significantly downregulated circRNAs. (D) Volcano plot representing the dysregulated mRNAs between RA and healthy controls. The two vertical grey dotted lines correspond to two-fold up- and downregulation (log<sub>2</sub> scaled), while the horizontal grey dotted line represents a P=0.05 (log<sub>10</sub> scaling). The red points represent significantly upregulated mRNAs and the blue points represent significantly downregulated mRNAs. circRNA, circular RNA; PBMCs, peripheral blood mononuclear cells; RA, rheumatoid arthritis; RNA-seq, RNA sequencing; HC, healthy controls; FC, fold change.

71 differentially expressed circRNAs between RA patients and HCs, of which 66 circRNAs were derived from exons (Fig. 1B). A violin plot demonstrated approximately equivalent distributions of normalized intensities for all the datasets in the tested subjects (Fig. S1). Compared with healthy controls, 41 and 30 circRNAs were upregulated and downregulated in RA patients, respectively. MA plots fully demonstrate the relationship between gene abundance and the significantly differentially expressed circRNAs with cutoff criteria of fold change >2.0 and P<0.05 in PBMCs of RA (Fig. 1C). In addition, 924 mRNAs were differentially expressed using the same cutoff criteria, including 597 upregulated and 327 downregulated mRNAs. A volcano plot demonstrated the variance among differentially expressed genes between RA and healthy controls (Fig. 1D).

*Network construction of circRNA-miRNA interactions and GO and KEGG pathway analysis for the differentially expressed genes.* To investigate the function of circRNAs in RA, the present study predicted the differentially expressed miRNAs of their potential role through specific base-paired and constructed circRNA-miRNA networks of the top five

up- and downregulated circRNAs (Fig. 2A). Additionally, the functions of differentially expressed mRNAs were investigated using GO analyses, sequencing the top 20 GO terms according to the P-value of enrichment in the biological process (BP), cellular component (CC) and molecular function (MF) categories (Fig. 2B). In the BP category, enriched terms included cytoplasmic mRNA processing body assembly, regulation of NF-κB transcription factor activity and nuclear transcription. The degree of MF was significantly enriched in poly(A)-specific ribonuclease activity and syntaxin-1 binding. The genes in the CC group were significantly enriched in the cytoplasm and cytoplasmic side of the mitochondrial outer membrane. The specific term information is shown in Table S1. Furthermore, gene set enrichment analysis (GSEA) was conducted. These functional analyses therefore identified relevant metabolic pathways that are important in the development of autoimmune disease. The analysis results demonstrated that the Toll-like receptor signaling pathway and MAPK signaling pathway were the significantly enriched pathways (Fig. 2C and D), which contained the core enrichment genes for the related pathways, such as JUN, NF-κB, TLR2 and NFκBIA.



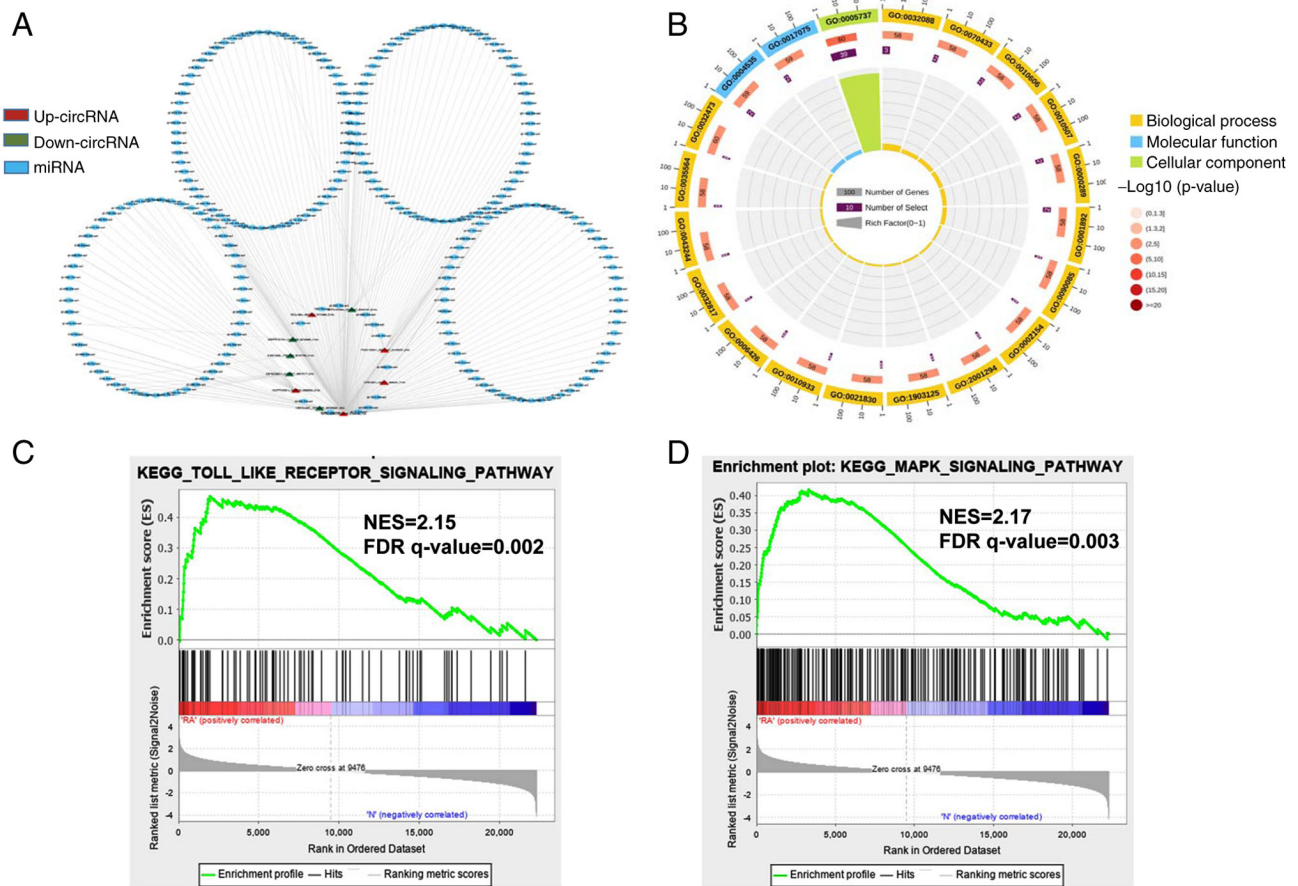


Figure 2. Network construction of circRNA-miRNA interactions and GO and KEGG pathway analysis for the differentially expressed genes. (A) Construction of the circRNA-miRNA regulatory network. The circRNA-miRNA regulatory network included 10 circRNAs and 251 miRNAs (red, upregulated circRNAs; green, downregulated circRNAs; blue, miRNA). (B) The top 20 significantly enriched GO annotations in BP, MF and CC placed in three columns in the circle diagram. The outermost layer was a GO ID and specific GO term information is shown in Table S1. The inner circles from outside to inside corresponded to the number of genes sequenced by their enrichment scores ( $-\log_{10}$  of the P-values), number of selected genes and rich factor, respectively. The coloured dots represent the enriched P-values, yellow represents BP, blue represents MF and green represents CC. Gene set enrichment analysis was performed to determine the differential expression of mRNAs for the (C) Toll-like receptor and (D) MAPK signaling pathways. circRNA, circular RNA; miRNA, microRNA; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, biological process; MF, molecular function; CC, cellular component; NES, Normalized enrichment scores; FDR, false discovery rate.

*circPTPN22* is a common differentially expressed circRNA among RA and SLE patients and acts as a potential diagnostic biomarker for RA. To screen out circRNAs simultaneously associated with both RA and SLE, clustering was used to identify common differentially expressed circRNAs by referring to our previous study that identified 128 aberrantly expressed circRNAs in SLE patients compared with HCs (21). Calculating the intersection of differentially expressed circRNAs between RA patients and SLE patients compared with healthy controls, 21 common differentially expressed circRNAs were obtained. A Venn diagram of this clustering demonstrated that these circRNAs consisted of 8 significantly upregulated and 13 downregulated circRNAs overlapping between the foregoing comparisons (Fig. 3A). Notably, *circPTPN22* was included in the 13 downregulated circRNAs among RA and SLE patients (detailed in Table SII).

To verify whether the expression level of *circPTPN22* in RA was consistent with the sequencing data, RT-qPCR was performed using an independent set of samples from 42 RA patients, 44 healthy controls and 45 SLE patients as an autoimmune disease control. The results demonstrated

that the average expression level of *circPTPN22* was significantly downregulated in the PBMCs of RA and SLE patients compared to healthy controls; In addition, the *circPTPN22* level was even lower in RA patients than in SLE patients (Fig. 3B). Similarly, the expression of the parental gene *PTPN22* of *circPTPN22* was also examined in RA patients, HCs and SLE patients. The results demonstrated that compared with the HC group, the expression level of *PTPN22* in PBMCs in RA and SLE patients was significantly downregulated and the expression level of *PTPN22* in PBMCs in RA patients was lower than that in SLE patients (Fig. S2A). In addition, Pearson's correlation test demonstrated that the expression levels of *circPTPN22* and *PTPN22* in the PBMCs of RA patients were positively correlated (Fig. S2B). Furthermore, receiver operating characteristic (ROC) curve analysis was used to validate the potential diagnostic biomarker of *circPTPN22* in RA. The AUC values indicated that *circPTPN22* in PBMCs could distinguish RA patients from HCs (Fig. 3C) and even from SLE patients (Fig. 3D), with the ability to discriminate SLE patients from HCs (Fig. 3E), as expected.

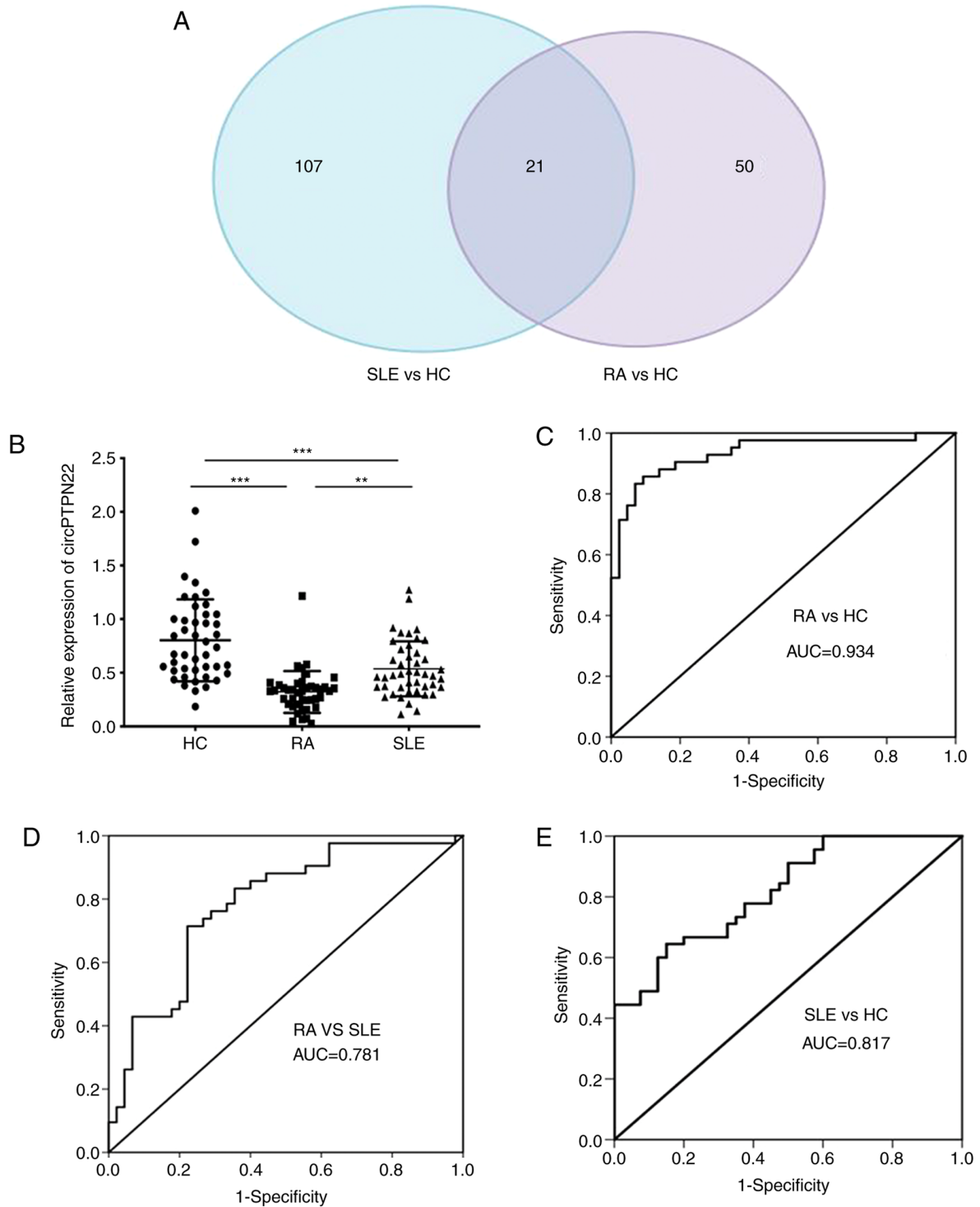


Figure 3. circPTPN22 expression as a potential biomarker for RA diagnosis. (A) Venn diagram showing the overlap of 21 significantly differentially expressed circRNAs in RA and SLE compared with HCs. There were 128 significantly dysregulated circRNAs in PBMCs of SLE patients vs. HCs (blue area) and 71 significantly dysregulated circRNAs in PBMCs of RA patients vs. HCs (purple area). (B) The expression levels of circPTPN22 in PBMCs of 42 RA patients, 45 SLE patients and 44 HC were determined by reverse transcription-quantitative PCR. ROC curve analysis of the sensitivity and specificity of (C) circPTPN22 in RA patients and HCs, (D) circPTPN22 in PBMCs of RA patients and SLE and (E) circPTPN22 in PBMCs of SLE patients and HCs. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . circRNA, circular RNA; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; HC, healthy controls; ROC, receiver operating characteristic; PBMCs, peripheral blood mononuclear cells; AUC, area under the curve.

**Correlation between circPTPN22 expression and RA clinical characteristics.** To determine whether circPTPN22 in PBMCs of RA patients was a potentially relevant biomarker of RA severity, the correlations between circPTPN22 levels and RA clinical features were examined, including duration, TNF- $\alpha$ ,

IL-6, C3, C4, RA-IgG, RA-IgM, RA-IgA, IgA, IgG, IgM, IgE, anti-CCP, TJC, ESR, CRP and DAS28. The results demonstrated that the circPTPN22 level was negatively associated with RA-IgG, RA-IgA, RA-IgM, anti-CCP, RF and CRP levels in RA (Fig. 4); however, the circPTPN22 level did not

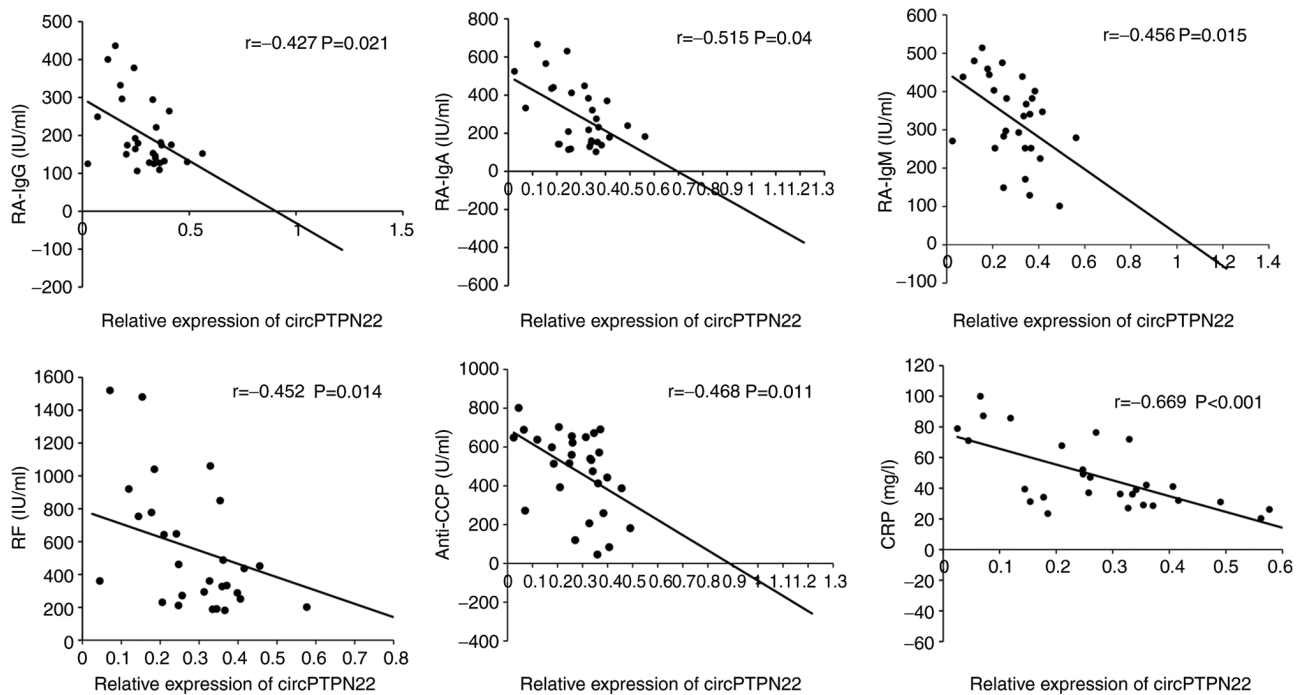


Figure 4. Pearson's correlation test of clinical disease activities and circPTPN22 levels in PBMCs of RA patients. circRNA, circular RNA; PBMCs, peripheral blood mononuclear cells; RA, rheumatoid arthritis; RF, rheumatoid factor; anti-CCP, anti-cyclic citrullinated peptide; CRP, C-reactive protein.

correlate with other indicators of RA severity (data not shown). Therefore, circPTPN22 could at least partly reflect the severity of the disease.

**RT-qPCR validation of the circPTPN22-miRNA-mRNA network.** Based on the prediction software, differentially expressed microRNAs matched with circPTPN22 were predicted to construct the interaction network (Fig. 5A). The predicted target mRNAs selected from the RNA-seq data were associated with inflammation. Using Miranda software, it was found that the sequence of circPTPN22 had potential binding sites with the seed regions of the four putative target miRNAs (Fig. 5B).

To confirm whether circPTPN22 is involved in the development and progression of RA disease through a sponge mechanism, RT-qPCR was conducted to validate the expression levels of miRNA in the PBMCs of RA and HC. The results demonstrated that hsa-miR-3074-5p, hsa-miR-373-3p, hsa-miR-766-3p and hsa-miR-34c-5p were upregulated in the PBMCs of the RA group compared to the HC group (Fig. 6), which was consistent with the RNA-seq results.

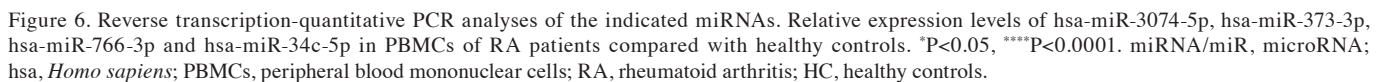
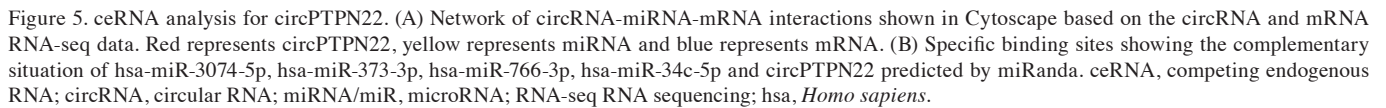
**Construction of the PPI network and identification of hub genes.** First, the differentially expressed mRNA genes were constructed into a PPI network using the STRING database, with the maximum interaction confidence set to 0.9. A total of 246 nodes and 998 edges were involved in the PPI network (Fig. 7A). To find the key nodes in the network, the 11 topological analysis methods available on the Cytoscape plug in CytoHubba were used to rank all the nodes by their network features (33). Then, the top 10 genes were identified (Fig. 7B). The results demonstrated that IL6 was the most prominent gene with connectivity degree=76, followed by IL-8 (CXCL8;

degree=63), transcription factor AP-1 (JUN; degree=56), IL-10 (degree=53), IL-1 $\beta$  (degree=51), MMP9 (degree=48), growth-regulated  $\alpha$  protein (CXCL1; degree=48), prostaglandin G/H synthase 2 (PTGS2; degree=42), intercellular adhesion molecule 1 (ICAM1; degree=41) and C-C chemokine receptor type 5 (CCR5; degree=36). Since gene products are the core of the PPI network, these hub genes were considered potential therapeutic targets and some genes have been proven to serve a role in RA (34).

## Discussion

RA is a common complex autoimmune disease involving multiple joints and systemic inflammation. Early diagnosis is the key event for RA treatment; nevertheless, a number of patients with negative RF and/or anti-CCP are easily misdiagnosed at an early stage (1,6). It has been found that circRNAs are important members of the ncRNA family. Compared with microRNAs and lncRNAs, circRNAs are more widely expressed and more stable in mammalian cells (8,9), which indicates that circRNAs may be potential diagnostic markers for a number of diseases. The development of high-throughput sequencing and bioinformatics analysis has increased our understanding of the important role of circRNAs in the pathogenesis and development of human disease.

To explore new biomarkers for RA, the present study investigated the expression profile of circRNAs in PBMCs in RA patients through high-throughput RNA sequencing. A total of 71 circRNAs were found to be significantly dysregulated (41 upregulated and 30 downregulated) in RA patients compared with healthy controls. Based on a previous study (21), 21 differentially expressed circRNAs shared in both RA and SLE PBMCs were found, which notably included circPTPN22,



The pathogenesis of RA is commonly attributed to the infiltration of T cells into the synovial membrane and T cells are the main culprits in inducing inflammation (35). In addition,

the loss of tolerance to autoantigens may be another cause of RA (36). Despite these biological factors, other predisposing factors, such as genetic components, have been reported as high-risk factors for RA (3). Due to the complex clinical manifestations of RA, a single experimental indication cannot



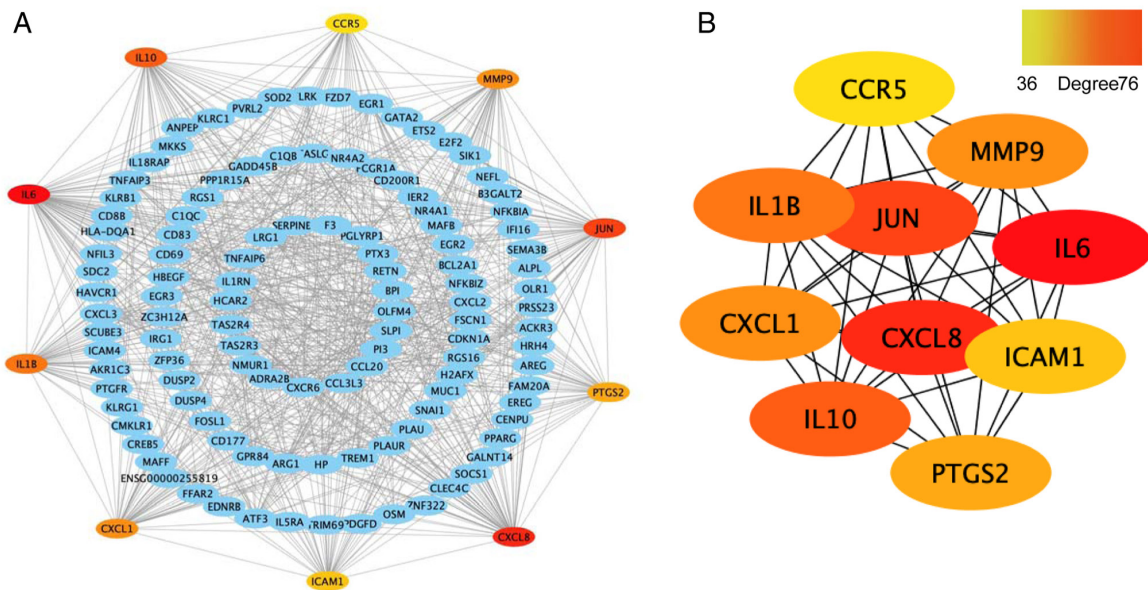


Figure 7. Construction of the PPI network and identification of hub genes. (A) PPI network constructed with the differentially expressed mRNA genes. Each node represents a protein and edges represent interactions between proteins. (B) Top 10 genes with a high degree calculated by cytoHubba (degree >35). The darker the red colour, the higher the degree. PPI, protein-protein interaction.

provide an accurate diagnosis of RA. Therefore, the present study evaluated the clinical diagnostic value of circPTPN22 in RA by using ROC curves and found the potential role of a biomarker for the diagnosis of RA. The results also demonstrated that the circPTPN22 expression levels were negatively correlated with RA-IgG, RA-IgA, RA-IgM, anti-CCP, RF and CRP levels in RA PBMCs but not correlated with TNF- $\alpha$ , IL-6, C3, C4, IgA, IgG, IgM, IgE, TJC, ESR or DAS28. These results further indicate that circPTPN22 could be a promising novel diagnostic biomarker for RA.

circRNAs have been shown to regulate the level of target miRNAs through molecular sponge mechanisms and then regulate the expression of corresponding target genes of miRNAs (11). To explore the molecular mechanism of circPTPN22 in RA, a ceRNA interaction network was constructed through Cytoscape. A total of four putative targets of miRNAs were predicted and their roles in RA are unclear. GO analysis demonstrated that the differentially expressed genes for 'transcription factor NF- $\kappa$ B activity', 'nuclear-transcribe', 'poly(A)-specific ribonuclease activity' and 'syntaxin-1 binding' are key checkpoints of effector function in RA. In addition, gene set enrichment analysis demonstrated that they were involved in 'TLR signaling pathway' and 'MAPK signaling pathway'. PTPN22 has been reported to be involved in a variety of autoimmune diseases and modulates signaling through antigens and innate immune receptors (23). Additionally, it serves roles in lymphocyte development and activation, establishment of tolerance and innate immune cell-mediated host defence and immunoregulation (23). However, whether circPTPN22 also participates in autoimmune diseases through the same or similar mechanism has not yet been reported.

Few circRNA disorders have previously been reported in RA patients. Due to different sample and database comparison methods, the results of these data may be inconsistent. Ouyang *et al* (37) identified 12 circRNAs that were differentially

expressed in patients with RA and healthy controls using microarray profiling of circRNAs. Their RT-qPCR analysis demonstrated increased expression of circRNA\_104871, circRNA\_003524, circRNA\_101873 and circRNA\_103047 in PBMCs from RA. The present study suggested that the five increased circRNAs had a significant value in RA diagnosis by ROC curve analysis. In addition, Zheng *et al* (38) screened 584 circRNAs differentially expressed in PBMCs of RA patients compared with HCs by circRNA microarray and they demonstrated that the expression levels of hsa\_circRNA\_104194, hsa\_circRNA\_104593, hsa\_circRNA\_103334, hsa\_circRNA\_101407 and hsa\_circRNA\_102594 were consistent with the results from the microarray analysis by RT-qPCR validation. Yang *et al* (39) found 71 markedly dysregulated circRNAs using RNA-seq. Their results indicated that hsa\_circ\_0000396 and hsa\_circ\_0130438 were downregulated in an RA group compared with a healthy group. The ROC curve indicated the diagnostic value of both circRNAs for RA. In Wen *et al* (40), a total of 165 circRNAs and 63 miRNAs were differentially expressed between RA patients and healthy individuals according to RNA-seq. They suggested that the expression of hsa\_circ\_0001200, hsa\_circ\_0001566, hsa\_circ\_0003972 and hsa\_circ\_0008360 in PBMCs from RA patients might serve as potential biomarkers for the diagnosis of RA. Although the above studies investigated the RNA expression profile in RA PBMC by vary sequencing methods, most did not provide raw RNA-seq dataset, or some papers only present top 10 circRNAs, or only randomly selected six circRNAs for further RT-qPCR validation. Therefore, the authors of the present study could not find how a number of circRNAs are the same and how a number of circRNAs differ among these studies. The present study demonstrated significant differences in circPTPN22 expression between RA patients, SLE patients and HCs in three validation cohorts. Although circPTPN22 expression in RA is lower compared with that in SLE patients, circPTPN22 expression in both RA

and SLE patients was significantly lower compared with that in HCs, suggesting that circPTPN22 could be a common diagnostic circRNA for autoimmune diseases such as RA, SLE and other possible immune diseases. To verify that circPTPN22 can reliably distinguish RA patients from SLE patients and HCs, a larger sample size is needed.

In conclusion, circPTPN22 may be associated with the classification of RA and involved in the pathogenesis of the RA process through ceRNA mechanisms. In our previous study, circPTPN22 was verified as a potential biomarker of SLE and was predicted to translate a truncated form of PTPN22 (21) and it is possible that the truncated PTPN22 would interact with the full-length PTPN22 protein directly or indirectly. Since the *PTPN22* gene codes for a tyrosine phosphatase, with a potential function in the regulation of T-cell and B-cell activation, it is of clear interest for the study of the etiology of RA to know whether circPTPN22 also participates in autoimmune diseases through regulating its parental gene *PTPN22*. However, the detailed and specific mechanisms of circPTPN22 potentially involved in RA pathogenesis remain to be clarified in the future.

### Acknowledgements

The authors would like to thank Dr Shifei Li and Dr Qijun Wu (Department of Dermatology, The First Affiliated Hospital of Third Military Medical University, Chongqing, China) for helping them discuss the revision of the manuscript.

### Funding

The present study was supported by grants from the National Key Research and Development Project (grant no. 2016YFA0502203), the Natural Science Foundation of China (grant no. 81773316) and the Key Research and Development Project of Anhui Province (grant no. 202004j07020002).

### Availability of data and materials

The datasets used and/or analyzed during the present study are available from the corresponding author upon reasonable request.

### Authors' contributions

ZJ performed the main experiments in this study. ZZ, QM and YZ performed the clinical data collection and statistical analysis. ZJ and ZZ confirm the authenticity of all the raw data. BN interpreted the clinical data of patients. MZ and JT designed the study and drafted the manuscript. All authors read and approved the final manuscript.

### Ethics approval and consent to participate

The present study was approved by the ethics committee of the First Affiliated Hospital of Third Military Medical University (Chongqing, China) and the reference ethical number is KY2019119. Written informed consent was obtained from all subjects prior to participation.

### Patient consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### References

- Smolen JS, Aletaha D and McInnes IB: Rheumatoid arthritis. *Lancet* 388: 2023-2038, 2016.
- Symmons D, Turner G, Webb R, Astin P, Barrett E, Lunt M, Scott D and Silman A: The prevalence of rheumatoid arthritis in the United Kingdom: New estimates for a new century. *Rheumatology (Oxford)* 41: 793-800, 2002.
- Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, Yoshida S, *et al*: Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 506: 376-381, 2014.
- Sugiyama D, Nishimura K, Tamaki K, Tsuji G, Nakazawa T, Morinobu A and Kumagai S: Impact of smoking as a risk factor for developing rheumatoid arthritis: A meta-analysis of observational studies. *Ann Rheum Dis* 69: 70-81, 2010.
- Nisar MK and Ostor AJ: Disease remission the goal of therapy in rheumatoid arthritis. *Practitioner* 254: 17-21, 2010.
- van Venrooij WJ, van Beers JJ and Pruijn GJ: Anti-CCP antibody, a marker for the early detection of rheumatoid arthritis. *Ann N Y Acad Sci* 1143: 268-285, 2008.
- Cocquerelle C, Mascres B, Hétiuin D and Bailleul B: Mis-splicing yields circular RNA molecules. *FASEB J* 7: 155-160, 1993.
- Sanger HL, Klotz G, Riesner D, Gross HJ and Kleinschmidt AK: Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc Natl Acad Sci USA* 73: 3852-3856, 1976.
- Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, *et al*: Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495: 333-338, 2013.
- Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF and Sharpless NE: Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 19: 141-157, 2013.
- Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK and Kjems J: Natural RNA circles function as efficient microRNA sponges. *Nature* 495: 384-388, 2013.
- Abdelmohsen K, Panda AC, Munk R, Grammatikakis I, Dudekula DB, De S, Kim J, Noh JH, Kim KM, Martindale JL and Gorospe M: Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by circPABPN1. *RNA Biol* 14: 361-369, 2017.
- Chen N, Zhao G, Yan X, Lv Z, Yin H, Zhang S, Song W, Li X, Li L, Du Z, *et al*: A novel FLI1 exonic circular RNA promotes metastasis in breast cancer by coordinately regulating TET1 and DNMT1. *Genome Biol* 19: 218, 2018.
- Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, *et al*: Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol* 22: 256-264, 2015.
- Zhang M, Huang N, Yang X, Luo J, Yan S, Xiao F, Chen W, Gao X, Zhao K, Zhou H, *et al*: A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. *Oncogene* 37: 1805-1814, 2018.
- Chen L, Kong R, Wu C, Wang S, Liu Z, Liu S, Li S, Chen T, Mao C and Liu S: Circ-MALAT1 functions as both an mRNA translation brake and a microRNA sponge to promote Self-Renewal of hepatocellular cancer stem cells. *Adv Sci (Weinh)* 7: 1900949, 2020.
- Mao Y, Zhang L and Li Y: circEIF4G2 modulates the malignant features of cervical cancer via the miR218/HOXA1 pathway. *Mol Med Rep* 19: 3714-3722, 2019.
- Zhang H, Zhu L, Bai M, Liu Y, Zhan Y, Deng T, Yang H, Sun W, Wang X, Zhu K, *et al*: Exosomal circRNA derived from gastric tumor promotes white adipose browning by targeting the miR-133/PRDM16 pathway. *Int J Cancer* 144: 2501-2515, 2019.

19. Guo G, Wang H, Ye L, Shi X, Yan K, Lin K, Huang Q, Li B, Lin Q, Zhu L, *et al*: Hsa\_circ\_0000479 as a novel diagnostic biomarker of systemic lupus erythematosus. *Front Immunol* 10: 2281, 2019.
20. Zhong S, Ouyang Q, Zhu D, Huang Q, Zhao J, Fan M, Cai Y and Yang M: Hsa\_circ\_0088036 promotes the proliferation and migration of fibroblast-like synoviocytes by sponging miR-140-3p and upregulating SIRT 1 expression in rheumatoid arthritis. *Mol Immunol* 125: 131-139, 2020.
21. Miao Q, Zhong Z, Jiang Z, Lin Y, Ni B, Yang W and Tang J: RNA-seq of circular RNAs identified circPTPN22 as a potential new activity indicator in systemic lupus erythematosus. *Lupus* 28: 520-528, 2019.
22. Cohen S, Dadi H, Shaoul E, Sharfe N and Roifman CM: Cloning and characterization of a lymphoid-specific, inducible human protein tyrosine phosphatase, Lyp. *Blood* 93: 2013-2024, 1999.
23. Bottini N and Peterson EJ: Tyrosine phosphatase PTPN22: Multifunctional regulator of immune signaling, development, and disease. *Annu Rev Immunol* 32: 83-119, 2014.
24. Svensson MN, Doody KM, Schmiedel BJ, Bhattacharyya S, Panwar B, Wiede F, Yang S, Santelli E, Wu DJ, Sacchetti C, *et al*: Reduced expression of phosphatase PTPN2 promotes pathogenic conversion of Tregs in autoimmunity. *J Clin Invest* 129: 1193-1210, 2019.
25. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB and van Riel PL: Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 38: 44-48, 1995.
26. Hochberg MC: Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 40: 1725, 1997.
27. Wang K, Singh D, Zeng Z, Coleman SJ, Huang Y, Savich GL, He X, Mieczkowski P, Grimm SA, Perou CM, *et al*: MapSplice: Accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res* 38: e178, 2010.
28. Anders S, Pyl PT and Huber W: HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166-169, 2015.
29. Anders S and Huber W: Differential expression analysis for sequence count data. *Genome Biol* 11: R106, 2010.
30. 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.
31. Rehmsmeier M, Steffen P, Hochsmann M and Giegerich R: Fast and effective prediction of microRNA/target duplexes. *RNA* 10: 1507-1517, 2004.
32. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T: Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498-2504, 2003.
33. 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.
34. Bie Y, Ge W, Yang Z, Cheng X, Zhao Z, Li S, Wang W, Wang Y, Zhao X, Yin Z and Li Y: The crucial role of CXCL8 and Its receptors in colorectal liver metastasis. *Dis Markers* 2019: 8023460, 2019.
35. Smolen JS and Steiner G: Therapeutic strategies for rheumatoid arthritis. *Nat Rev Drug Discov* 2: 473-488, 2003.
36. Choy E: Understanding the dynamics: Pathways involved in the pathogenesis of rheumatoid arthritis. *Rheumatology (Oxford)* 51 (Suppl 5): v3-v11, 2012.
37. Ouyang Q, Wu J, Jiang Z, Zhao J, Wang R, Lou A, Zhu D, Shi GP and Yang M: Microarray expression profile of circular rnas in peripheral blood mononuclear cells from rheumatoid arthritis patients. *Cell Physiol Biochem* 42: 651-659, 2017.
38. Zheng F, Yu X, Huang J and Dai Y: Circular RNA expression profiles of peripheral blood mononuclear cells in rheumatoid arthritis patients, based on microarray chip technology. *Mol Med Rep* 16: 8029-8036, 2017.
39. Yang X, Li J, Wu Y, Ni B and Zhang B: Aberrant dysregulated circular RNAs in the peripheral blood mononuclear cells of patients with rheumatoid arthritis revealed by RNA sequencing: Novel diagnostic markers for RA. *Scand J Clin Lab Invest* 79: 551-559, 2019.
40. Wen J, Liu J, Zhang P, Jiang H, Xin L, Wan L, Sun Y, Huang D, Sun Y, Long Y, *et al*: RNA-seq reveals the circular RNA and miRNA expression profile of peripheral blood mononuclear cells in patients with rheumatoid arthritis. *Biosci Rep* 40: BSR20193160, 2020.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.