Identification of RRM2 in peripheral blood mononuclear cells as a novel biomarker for the diagnosis of rheumatoid arthritis

B. Wu[1,2], R. Peng[1,2], Y. Chen[1,2], J. Zhong[1], S. Zhong[3], B. Wang[1], X. Li[3], N. Jiang[1,2], J. Shang[1,2], D. Wang[1], H. Lu[1]

[1]Department of Orthopaedics, [2]Guangdong Provincial Key Laboratory of Biomedical Imaging, [3]Department of Rheumatology, The Fifth Affiliated Hospital of Sun Yat-Sen University, Zhuhai, Guangdong, China.

Abstract

Objective

Rheumatoid arthritis (RA) is a common autoimmune disease. However, the positive diagnosis value of the current biomarkers is unsatisfactory. Here, we aimed to identify RA-associated susceptibility genes and explore their potential as novel biomarkers for the diagnosis of RA.

Methods

Peripheral blood mononuclear cells (PBMCs) were collected from healthy controls and RA patients. RNA-seq and bioinformatics analyses were performed to identify the hub genes associated with RA. Then, the expression of hub genes was assessed in mRNA expression profiles from GEO datasets. Real time-quantitative PCR (RT-qPCR) was performed to further confirm the expression of the hub genes using the PBMCs that were collected from RA patients (n=47) and healthy controls (n=40). Finally, we evaluated the diagnostic potential of the candidate mRNAs.

Results

RNA-seq analyses revealed 178 dysregulated genes measured by changes in mRNAs between the healthy controls and the RA patients. We identified 3 candidate mRNAs, including ASPM, DTL and RRM2, all of which were highly expressed in RA. RRM2 showed a significant higher expression in remissive RA compared with active RA. Significant correlations were observed between DTL and IL-8, TNF-α which were tested in serum by ELISA, between RRM2 and CDAI, DAS-28, tender and swollen joints, respectively. The expression level of RRM2 was significantly higher in RA patients with the Anti-CCP$^-$ than with the Anti-CCP$^+$. The AUC (RA vs. OA) value of RRM2 was 0.941 ($p<0.0001$; sensitivity=0.867; specificity=0.904).

Conclusion

RRM2 showed high diagnosis efficiency for RA patients. Therefore, the findings provided a novel candidate biomarker for the diagnosis of RA.

Key words

rheumatoid arthritis, susceptibility genes, peripheral blood mononuclear cells, bioinformatics, biomarker

*Biao Wu, MD**
*Rong Peng, MD**
*Yufeng Chen, MD**
Junlong Zhong, MD
Shuping Zhong, MD
Bin Wang, MD
Xuegang Li, MD
Ning Jiang, MD
Jie Shang, MD
Dawei Wang, MD
Huading Lu, MD

**These authors contributed equally.*

Please address correspondence to:
Huading Lu,
Department of Orthopaedics,
The Fifth Affiliated Hospital
of Sun Yat-sen University,
No. 52 Meihua Dong Road,
Zhuhai 519000, China.
E-mail: johnniehuading@163.com

and to:
Dawei Wang: wdw19850226@163.com

Received on July 27, 2021; accepted in revised form on December 9, 2021.

Availability of data and material: the datasets analysed during the current study are available from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/gds) accession number: GSE93272.

Funding: this study was supported by grants from the National Natural Science Foundation of China (no. 81772384), GuangDong Basic and Applied Basic Research Foundation (no. 2021A1515010531), Medical Scientific Research Foundation of Guangdong Province of China (no. A2020283), and Medical Science and Technology planning Project of Zhuhai, China (no. ZH2201200003HJL and no. ZH22036201210053PWC). The financial support had no impact on the outcomes of this study.

Competing interests: none declared.

Introduction

Rheumatoid arthritis (RA) is a common chronic, inflammatory, autoimmune disease which can occur at any age (1). It is characterised by a progressive synovitis which is initiated and maintained by a complex interplay among different immune cells as well as between immune and tissue cells (2, 3). The average prevalence of RA ranges from 0.5% to1.0% globally (4). If treatment is inadequate or delayed, RA can lead to accumulating cartilage destruction, bone erosion, irreversible physical challenge as well as systemic complications (1, 5).

The disease is complex and is known to be prominently associated with some risk factors, including sex, genetic component, epigenetic and environmental factors (6). With the development of Genome-wide association studies and subsequent meta-analyses, plenty of genetic association studies show that RA has a strong genetic element which can account for 60% of the risk of developing RA (1, 6, 7) and more than 100 RA risk genetics loci have been identified, such as HLA-DRB1, CD28, STAT4, PADI4, PTPN22 as well as epigenetic modification (4, 8, 9). However, studies for RA associated susceptibility genes are not robust enough.

In RA treatment strategies, early and precise diagnosis is very important because timely treatment can protect as many as ninety percent patients with early RA from disease progression and subsequent joint damage (2, 3). However, in most RA patients, when the clinical manifestations such as swelling, morning stiffness and tenderness of joints are evident, the pathogenesis has begun many years (1, 10). Moreover, there is no diagnostic criteria for RA (10). Although the current ACR/EULAR 2010 classification criteria which includes joint distribution, symptom duration, serology of anti-citrullinated protein antibodies (ACPAs) and rheumatoid factor (RF), acute phase reactants (such as C-reactive protein and erythrocyte sedimentation rate) has a higher sensitivity for early recognition of RA than the older ACR 1987 classification system (10-12), it has only a specificity of 61% (2). Currently there are some ways for clinicians to evaluate a patient with possible early RA, including inquiry symptoms and medical history, physical examination, imaging test, and laboratory blood tests (13). Despite the fact that imaging tests, such as magnetic resonance imaging (MRI), ultrasonography, or x-ray, are very sensitive for detecting synovitis and pannus before bone erosion and cartilage degradation, their specificity and routine application are limited (3, 10, 14, 15). Additionally, the positive predictive value of the biomarkers (anti-cyclic citrullinated peptide antibody (anti-CCP), RF) is moderate. The sensitivity is 67% and 70% for anti-CCP and RF, respectively (16), which remain challenges facing rheumatologists. Therefore, the development of novel biomarkers for RA remains an unmet need.

In the present study, to identify the novel RA-associated susceptibility genes and test their potential as new biomarkers for the diagnosis of RA, we collected PBMCs from 4 healthy human donors and 4 RA patients. Then RNA-seq analyses were used to identify the differentially expressed genes (DEGs) between RA patients and healthy donors. The expression of hub genes in RA was assessed and screened in GSE93272 dataset which included 23 healthy control subjects and 136 patients with RA. Next, real time-quantitative PCR analyses (RT-qPCR) were performed to further confirm the expression of the hub genes from the PBMCs that were collected from 47 RA patients and 40 healthy volunteers. Furthermore, we further assess the potential of the candidate mRNAs as diagnostic biomarkers for RA (Fig. 1a).

Materials and methods

Study population and data resources

Peripheral blood collected from a total of 8 participants (4 RA patients *vs.* 4 healthy controls) was used for RNA-seq analysis. Peripheral blood was collected from a total of 108 participants who were recruited consecutively for this study: 47 patients with RA, 40 healthy controls, and 21 female osteoarthritis (OA) patients. All the RA patients met the classification criteria of ACR/EULAR 2010 and were re-

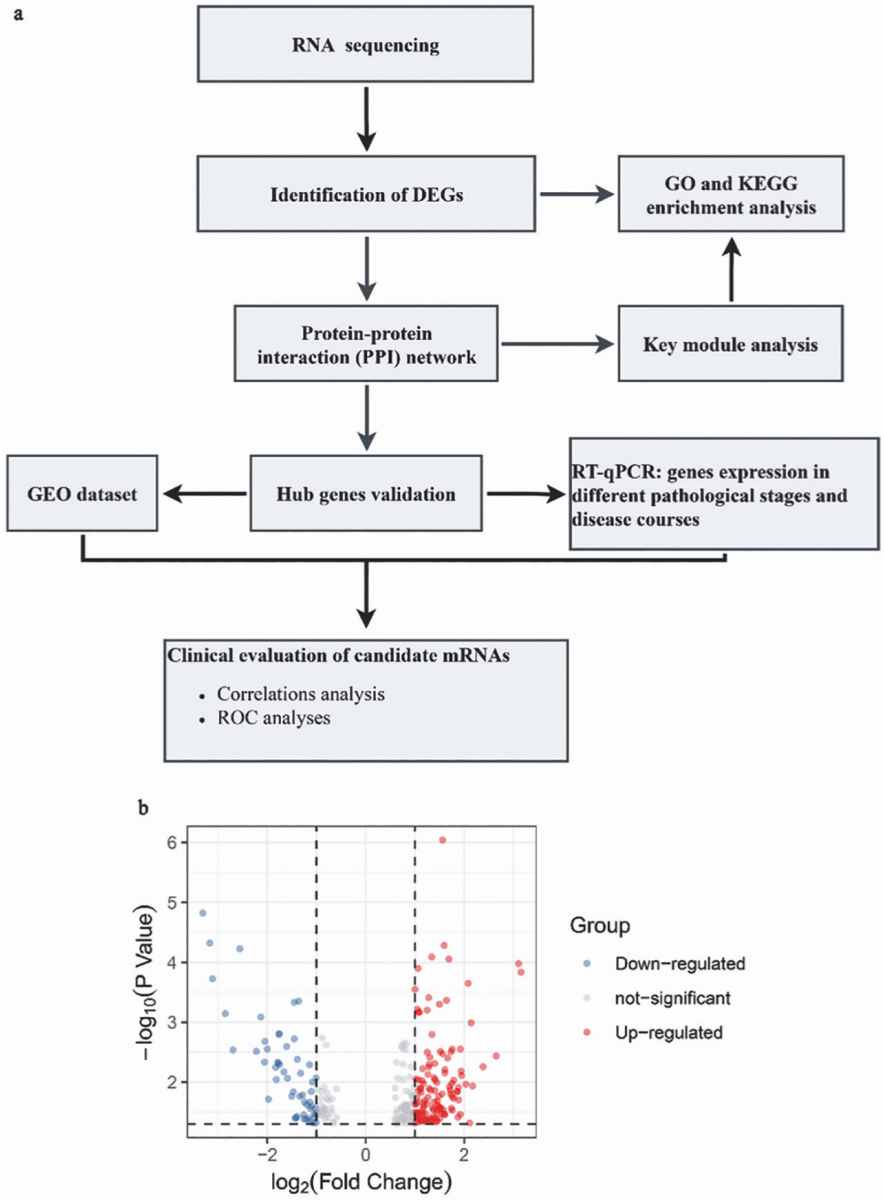

Fig. 1. (**a**) Flow chart of the study. DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein protein interaction; ROC, Receiver operating characteristic curves. CDAI, Crohn's Disease Activity Index; DAS28, Disease Activity Score 28; (**b**) Volcano plots showing the DEGs in the two comparison groups. The red dots represent up-regulated genes (log2FC > 1), and the blue dots represent down-regulated genes (log2FC < -1).

cruited in The Fifth Affiliated Hospital of Sun-Yat-Sen University. The ethics committee of Fifth Affiliated Hospital of Sun Yat-sen University approved the study in accordance with the relevant guidelines and regulations, and informed consent was obtained from all patients.

All the healthy controls (HCs) without osteoarthritis and ankylosing spondylitis were from the same hospital. All the participants underwent examination and evaluation of swollen and tender joints, Clinical Disease Activity Index (CDAI) and Disease Activity Score 28 (DAS28). Laboratory tests included anti-CCP, C-reactive protein (CRP), erythrocyte sedimentation rate (ESR) and blood routine parameters. The characteristics of healthy controls and patients with rheumatoid arthritis are shown in Table I.

To test the expression of hub genes between RA and HCs, we searched the Gene Expression Omnibus (GEO) database) (https://www.ncbi.nlm.nih.gov/geo/) for microarray datasets using the keyword "rheumatoid arthritis". Datasets were included if they met all of the following criteria: (1) were from humans; (2) included expression data from blood mRNA of both RA and HCs; (3) the number of rows in each platform was >50,000; (4) the number of RA samples was ≥20, and the number of HCs samples was ≥20. Finally, one dataset GSE93272 was selected (18).

Preparation of peripheral blood samples and isolation of plasma and RNA

For whole-blood transcriptome analysis, peripheral blood samples were collected from healthy human donors and RA individuals and stored in EDTA-2K -containing vacuum blood collection tubes at -80°C. Peripheral blood mononuclear cells (PBMCs) were extracted by using the Histopaque-1077 (Sigma-Aldrich, UK) from blood samples. And the plasma was isolated and stored at -80°C for Elisa. Follow-up total RNA was extracted from PBMCs using the Total RNA Kit I (Omega Bio-tek, USA) according to manufacturer's instructions. The concentrations of the RNA were quantified using NanoDrop ND-1000 spectrophotometer (ThermoFisher Scientific, USA) and assessed using absorbance ratios of A260/A280 nm >1.8 and RNA Integrity Number >7.

RNA-seq analysis

Before RNA-sequencing, the quality of RNA was tested by Agilent 2100 Bioanalyzer (Agilent Technology). 2 μg of RNA sample was taken for RNA-sequencing. RNase R digested and rRNA depleted RNAs were taken to generate the sequencing libraries by using Total RNA-seq (H/M/R) Library Prep Kit for Illumina (Vazyme Biotech) according to the manufacturer's instructions. The library preparations were sequenced on Hiseq X Ten (Illumina).

Identification of differentially expressed genes

For the identification of the differentially expressed mRNA in all DEGs between RA and healthy control samples, the limma package (v. 3.46.0) was performed with the predefined criterion (|log2FoldChange| >1 and p<0.05) (19).

Table I. Characteristics of healthy controls and patients with rheumatoid arthritis.

Characteristics	RA	HC
Number	47	40
Early RA vs. established RA	5 *vs*. 42	NA
Disease status (active vs. remission)	39 *vs*. 8	NA
Age (years)[a]	57.13±12.27	61.15±11.03
Sex (M/F)	16/31	11/29
Disease duration (month)[b]	1008.00 (158.40,1872.00)	NA
Anti-CCP(U/ml)[b]	179.40 (57.10,200.00)	NA
DAS28-CRP[a]	4.06±1.14	NA
DAS28-ESR[b]	4.90 (4.10,5.70)	NA
CRP (mg/l)[b]	25.14 (5.10,73.40)	13.35 (1.18,53.06)
ESR (mm/h)[a]	61.17±34.13	36.26±28.36
Swollen joints[b]	6.00 (2.00,10.00)	NA
Tender joints[b]	6.00 (3.00,10.00)	NA
ILD positive (%)	17.02%	NA

In the groups providing PBMCs, there was no significant difference in age and sex between RA patients and HC.
RA: rheumatoid arthritis; HC: healthy controls; M/F: male/female; anti-CCP: anti-cyclic citrullinated peptide antibodies; DAS28: Disease Activity Score in 28 joints; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; ASO: antistreptolysin; C3: complement 3; C4: complement 4; ILD: interstitial lung disease; NA: not available.
[a]Expressed as mean±standard deviation. [b]Expressed as the median (25th to 75th percentile).

Functional exploration for DEGs and key modules

To explore and visualise the potential functions of the identified mRNAs in DEGs and key modules, clusterProfiler package (20) in R was utilised for Gene Ontology (GO) enrichment analysis and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment.

Protein-protein interaction (PPI) network analysis

To establish the interactions of up-or down-regulated RA-related mRNAs in DEGs, we constructed a PPI network by STRING v. 11.0 database (https://string-db.org/) (21), and subsequently visualised it by Cytoscape software v. 3.8.2 (22). Typically, the most widely applied gene centrality measures are degree, closeness and betweenness.

Identification of hub genes and key modules

To screen hub genes, CytoHubba v. 0.1, a Cytoscape plug-in which can predict and explore important nodes/hubs by topological algorithms, was used based on the genes in PPI networks (23). Meanwhile, the Cytoscape plug-in Molecular Complex Detection v. 2.0.0 (MCODE) was employed to analyse the RA associated key modules with MCODE score ≥5 and nodes ≥5 (24).

Real time-quantitative PCR (RT-qPCR) analysis

Firstly, RNA concentration and purity were measured using NanoDrop2000C ultra-micro spectrophotometer, RNA samples A260/280 >2.0 and A260/230 >2.0 were considered to have higher purity. Then we synthesised first strand cDNA from total RNA templates using RevertAid Master Mix (Thermo Scientific, Lithuania). Briefly, the reverse transcription reactions were incubated for 30 min at 42°C, 5 min at 95°C and held at 4°C. Last, Quantitative Real time PCR (qRT-PCR) was performed to measure the relative genes expression on a CFX-96 Touch™ (BIO-RAD, USA) using Forget-Me-Not™ EvaGreen® qPCR Master Mix (Biotium, USA) according to the product's protocol. The primer sequences are listed in Supplementary Table S1. The expression was determined by the threshold cycle (Ct), and relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method with GAPDH as an internal control.

ELISA

To determine the concentration of inflammatory cytokines in all serum samples, such as IL-1α, IL-1β, IL-6, IL-8, TNF-α and IFN-γ, enzyme-linked immunosorbent assay (ELISA) was performed using the corresponding ELISA Kit (Saicheng Bio-tek, China) according to the manufacturer's instructions.

Statistical analysis

Data was tabulated using Microsoft Excel 2019 and analysed using SPSS software v. 24 (IBM, USA) and GraphPad Prism v. 8.0.1 (GraphPad Software, Inc., San Diego, CA). Student's t-test was employed to assess the gene expression between RA cases and controls in GSE93272 dataset which was normally distributed parameters. Mann-Whitney's U-test was applied for skewed distribution including gene expression in RA course and states. Pearson's analysis or Spearman method was performed for testing the correlation between mRNA expression levels and the clinical variables including CDAI, DAS28, counts of swollen/tender joints and inflammatory factors, as appropriate. Receiver operating characteristic (ROC) curves were conducted to evaluate the sensitivity and specificity of candidate mRNAs. $p<0.05$ was considered to be statistically significant threshold.

Results

Identification of DEGs

To identify the different expression genes between control and RA, we collected whole blood samples from 4 healthy individuals and 4 RA patients, and then performed RNA-seq analyses. Analysis of the dataset revealed 178 mRNAs that were differentially expressed between the healthy controls and the rheumatoid arthritis patients (|log2FoldChange| >1 and $p<0.05$) (Suppl. Table S1). Among the 178 mRNAs identified, 124 mRNAs were upregulated and 54 mRNAs were downregulated. Their distribution was presented using a volcano plot (Fig. 1b).

Functional exploration

To further understand the functions of the mRNAs that were differentially expressed and the connections among them, GO and KEGG pathway enrichment analyses were performed on the up-regulated or down-regulated mRNAs by the DAVID functional annotation clustering tool, respectively. The analysis results showed that under the "biological processes" category, up-regulated mRNAs were significantly

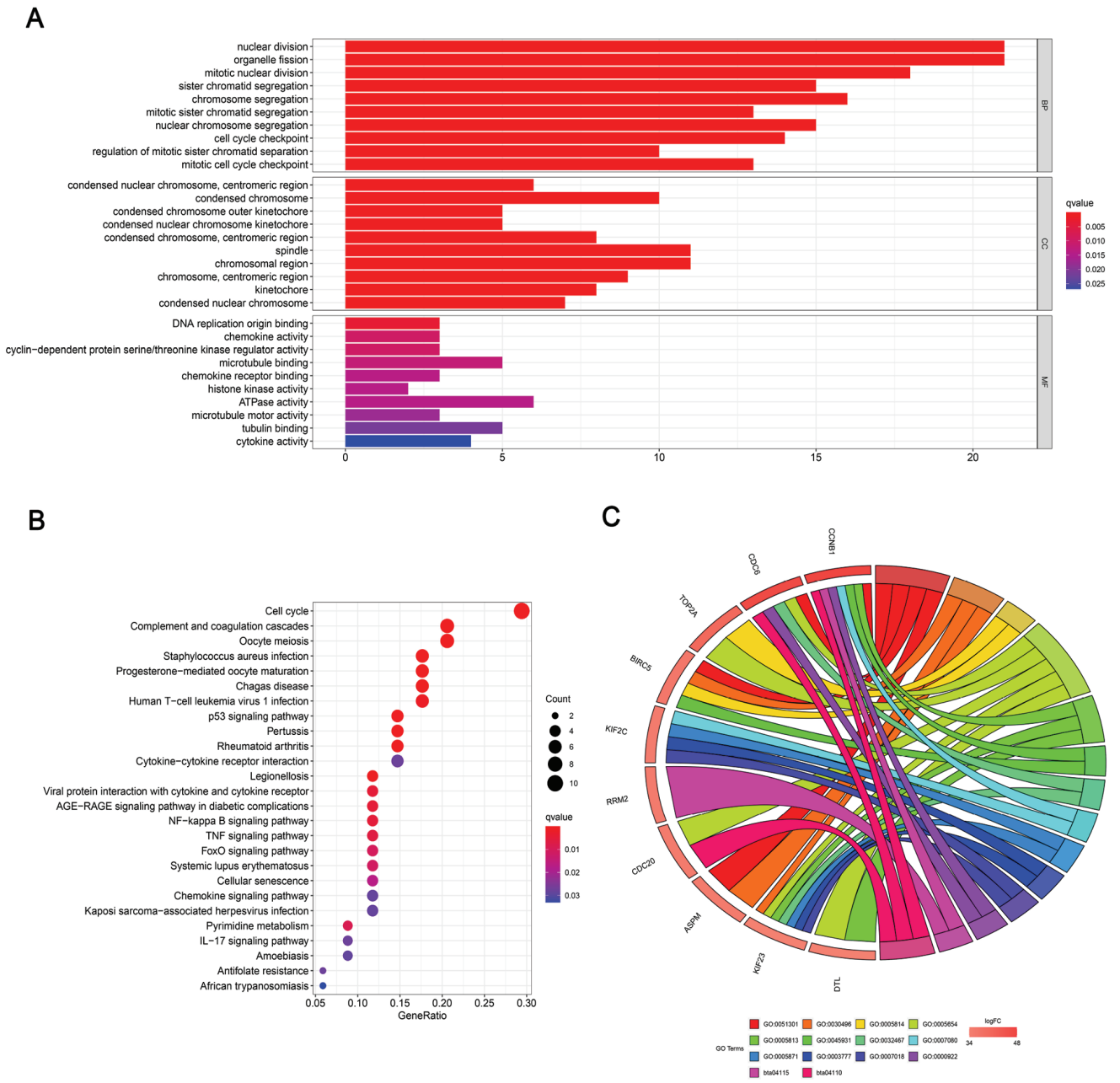

Fig. 2. Functional and pathway enrichment analysis of the mRNAs in each module.
(**a**) GO terms analysis of all mRNAs in 3 functional moduals.
(**b**) KEGG pathway enrichment analyses of all mRNAs in 3 functional moduals.
(**c**) GO terms/KEGG pathways associated with top 10 enriched hub genes.
GO terms: GO:0051301, cell division; GO:0030496, midbody; GO:0005814, centriole; GO:0005654, nucleoplasm; GO:0005813, centrosome; GO:0045931, positive regulation of mitotic cell cycle; GO:0032467, positive regulation of cytokinesis; GO:0007080, mitotic metaphase plate congression; GO:0005871, kinesin complex; GO:0003777, microtubule motor activity; GO:0007018, microtubule-based movement; GO:0000922, spindle pole; bta04115, p53 signaling pathway; KEGG pathway module: bta04110, Cell cycle.

enriched in nucleosome assembly, mitotic nuclear division, cell division, cellular protein metabolic process and gene silencing by RNA. On the other hand, down-regulated mRNAs were enriched in inflammatory response, immune response, cell adhesion, proteolysis, and bicarbonate transport (Suppl. Table S3-1). The top 5 pathways of KEGG enrichment for the mRNAs upregulated in rheumatoid arthritis patients were systemic lupus erythematosus, alcoholism, viral carcinogenesis, cell cycle and transcriptional misregulation in cancer. The top 5 KEGG pathways for the mRNAs downregulated in rheumatoid arthritis patients were complement and coagulation cascades, staphylococcus aureus infection, chagas disease, african trypanosomiasis and rheumatoid arthritis (Suppl. Table S3-2).

Protein-protein interaction (PPI) network construction and key module analysis

To further narrow and target the key regulatory mRNAs, a PPI network

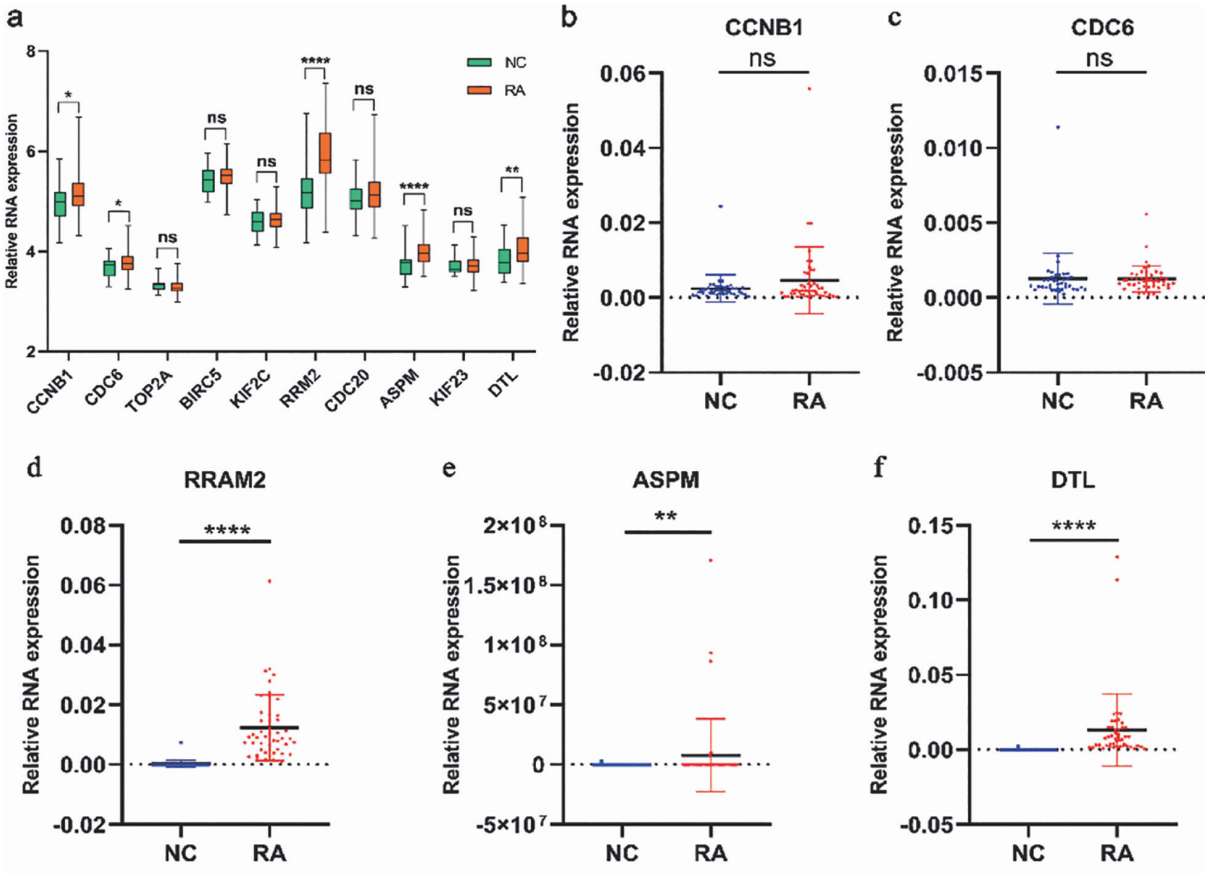

Fig. 3. Verification of the 10 hub genes by GSE93272 dataset and patients' samples. (**a**) Expression of 10 hub genes in GSE93272 dataset. (**b-f**) Expression of the hub genes in RA patients. Only hub genes that had significantly differentially expression between RA and normal control groups were listed. The green box indicated normal control group, and the orange box indicated the RA group. Mann Whitney test was performed to compare the means of two groups.
*p<0.05; **p<0.01; **p<0.001; ****p<0.0001. ns: no significance.

analysis was conducted using STRING (Suppl. Fig. S1). The top 10 hub mRNAs with the highest degrees of interaction were identified by the cytoHubba in Cytoscape. Their degree, closeness and betweenness are shown in Supplementary Table S4.

Subsequently, the top 3 significant functional modules were screened from the PPI network by the Molecular Complex Detection (MCODE) plug-in with MCODE score ≥5 and nodes ≥5 (Suppl. Table S5). All genes in the three modules were utilised to perform GO and KEGG enrichment analyses (Fig. 2a-b). We also performed GO and KEGG analyses for the 10 hub genes. The result was visualised in Fig. 2c.

Validation of the expression of hub genes

All of these 10 hub genes underwent expression validation in the GSE93272 dataset. The results, as shown in Fig. 3a, indicated that CCNB1, CDC6, RRM2, ASPM and DTL were significantly up-regulated in RA samples compared to healthy control samples. However, the five genes have been little studied on their associations with RA. To further confirm their roles in rheumatoid arthritis, we collected peripheral blood from 47 patients with RA and 40 healthy donors. Then RT-PCR analysis was performed to detect the expression of the five mRNAs (Fig. 3b-f). Eventually, we found that ASPM, DLT and RMM2 were significantly up-regulated in RA group compared to healthy controls.

Clinical evaluation for the 3 candidate mRNAs

Next, we assessed the expression of examined mRNAs in early RA patients and established RA patients, and in RA patients that were in remissive stages and in active stages. Consistent with the validation data above, the expression of all the three mRNAs showed great differences between RA and HCs. However, the expression of ASPM, DTL and RRM2 have no significant difference between early RA and established RA (Fig. 4a), and only RRM2 showed a significant high expression in remissive RA compared with active RA (Fig. 4b).

Further, we performed correlation analyses between the 3 candidate mRNAs and inflammatory factors, blood cells, acute phase reactants, RA disease activity scores, disease duration, and clinical symptoms (Suppl. Table S6). As shown in Fig. 4c-h, significant correlations were observed between DTL and IL-8, TNF-α, between RRM2 and CDAI, DAS-28, tender joints, swollen joints, respectively.

The above results taken together showed a great power of RRM2 in discriminating between RA patients and healthy subjects. Moreover, RRM2

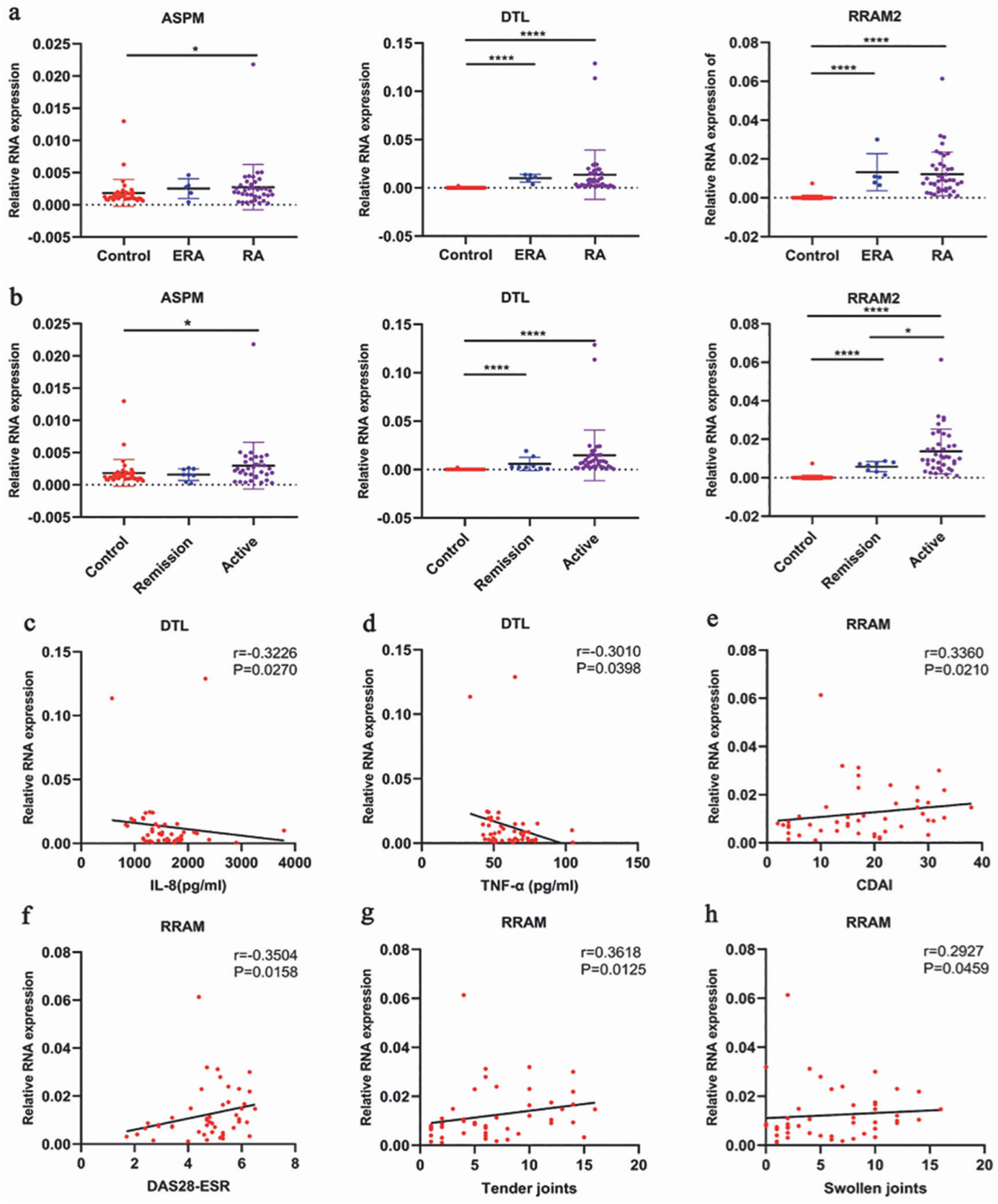

Fig. 4. Clinical evaluation for the 3 candidate mRNAs.
(**a**) levels of the 3 candidate mRNAs expression in PBMCs from healthy volunteers (control) (n=40), early RA patients (eRA) (n=5) and established RA patients (RA) (n=42). Bars show the mean with standard deviation. *p*-values were determined by Mann-Whitney test.
(**b**) levels of the 3 candidate mRNAs expression in PBMCs from RA patients in an active disease state (n=39) and those with RA in remission (n=8), and healthy volunteer s(n=40). Bars show the mean with standard deviation. *p*-values were determined by Mann-Whitney test.
(**c-h**) Spearman correlation analyses for the relationships between the 3 candidate mRNAs and inflammatory factors (c, d), CDAI (e), DAS28 (f), tender (g) and swollen joints (h).

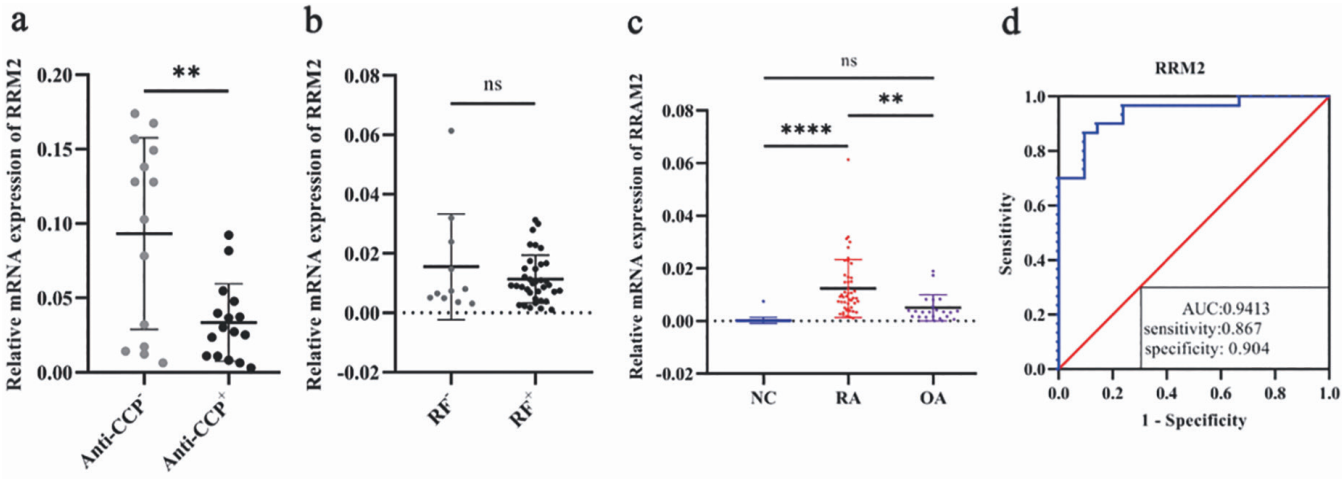

Fig. 5. Validation of the diagnostic values of RRM2 in RA patients.
(**a**) the mRNA expression level of RRM2 in RA patients with Anti-CCP$^-$ (n=14) and Anti-CCP$^+$(n=16).
(**b**) the mRNA expression level of RRM2 in RA patients with RF$^-$ (n=11) and RF$^+$(n=36).
(**c**) the mRNA expression level of RRM2 in RA patients (n=47), OA patients (n=21) and Healthy Controls (n=40).
(**d**) ROC curve analysis of RRM2 in patients with RA compared with OA.
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. ns: no significance.

could not only distinguish between RA patients with remission and active stage, but also had positive correlations with inflammatory factors and clinical manifestation.

RRM2 might serve as a novel diagnostic biomarker for RA

Based on the results showing that RRM2 was highly expressed in the PBMC of RA patients, and had obviously correlations with disease stage, inflammatory factors and clinical manifestation, we further explored whether RRM2 is a novel diagnostic biomarker for RA. We divided the RA cohorts into the Anti-CCP positive group (Anti-CCP$^+$) and the Anti-CCP negative group (Anti-CCP$^-$) by the cut-off value of Anti-CCP. Importantly, we found that the expression level of RRM2 was significantly higher in the Anti-CCP$^-$ group than that in the Anti-CCP$^+$ group (Fig. 5a), which indicated that RRM2 has a better diagnostic efficiency for RA patients with Anti-CCP$^-$. Additionally, we divided the RA cohorts into the RF negative group (RF$^-$) and the RF positive group (RF$^+$), then we analysed the mRNA expression level of RRM2 in both of them. The results showed that the expression level of RRM2 was no significant difference between the RF$^-$ group and the RF$^+$ group (Fig. 5b).

Next, we analysed the potential value of RRM2 expression in PBMC in the differential diagnosis of RA and osteoarthritis (OA) (n=21). The RT-qPCR analysis demonstrated that the RRM2 expression of RA was significantly higher than that of OA and HC, but there was no difference in the expression of RRM2 in NC and OA (Fig. 5c). Subsequently, to verify whether the RRM2 could discriminate RA from OA patients, we performed Receiver-operating characteristic (ROC) curve analysis and measured the area under the curve (AUC) to assess the diagnostic potential. The AUC (RA *vs*. OA) value of RRM2 was 0.941 (p<0.0001; sensitivity = 0.867; specificity = 0.904) (Fig. 5d). The above results suggested that RRM2 might serve as a novel diagnostic biomarker for RA, and RRM2 could improve the diagnostic efficiency for RA patients with Anti-CCP$^-$

.

Discussion

Here, we performed RNA sequencing analyses of cohorts of patients with RA and HCs. Using bioinformatics analysis, this study comprehensively identified PBMC expressed mRNAs associated with RA and provided a novel candidate mRNA biomarker, RRM2, for RA diagnosis, as supported by both GEO dataset and clinical samples validations. Importantly, RRM2 has a high discriminative power with extremely high levels of sensitivity = 0.867 and specificity = 0.904. Moreover, we observed that PBMC-mRNA-RRM2 levels in RA patients raised following exacerbation of clinical symptoms and disease activity, including CDAI, DAS28-ESR, tender joints, swollen joints and a transformation of the disease state from remission to active. These results suggested that the potential role of RRM2 as a novel biomarker for diagnosing RA. As we know, the predictive value of the existing biomarkers (anti-CCP and RF) is moderate (16). The PBMC-RRM2 level could improve the diagnostic efficiency of RA.

Our results showed that the expression of RRM2 in active RA was obviously higher than that in remissive RA, this may imply that RRM2 may have the potential to assess the prognosis of RA. However, there is a great deal of work that needs further investigation, whether RA patients with high mRNA expression of RRM2 indeed have a worse prognosis, *e.g*. with more destructive disease or treatment failure.

Early diagnosis of RA is an important step towards a more effective prevention of the RA progression and subsequent damage (25). The RF and anti-CCP are the well-known serological biomarkers for RA diagnosis. Raised serum titre of RF is associated with disease activity, longer disease duration and extraarticular manifestations (26-28). And circulating anti-CCP can

be detected even 10 years before the first symptoms onset (29, 30). Therefore, the presence of RF and anti-CCP are commonly used as a diagnostic biomarker as well as a prognostic biomarker (31). However, the sensitivity of anti-CCP and RF is just 67% and 70%, respectively (16), which means that negative results do not exclude RA (10), which remains challenges facing rheumatologists. With the development of sequencing technology and bioinformatics analysis, many scientists are trying to find a novel highly sensitive and specific biomarker of the disease. Mounting data has suggested that non-coding RNAs (ncRNAs) in RA play a pivotal role in regulating inflammation and autoimmunity and can be regarded as candidate biomarkers for RA (32). For example, miR-146a, miR-150 and miR-223 have been found to be highly expressed in peripheral blood and joint tissues, and can serve as promising biomarkers for RA diagnosis (33). Another candidate long non-coding RNA biomarker is GAPLINC, which participated in the proliferation, invasion and migration of FLS, and whose expression was increased in the peripheral blood, T cell, and synovial tissues of RA patients (34). Intriguingly, the ncRNAs participate in the pathogenesis of RA through the lncRNA/circRNA-miRNA-mRNA network (32). And only a few of them have been employed in clinical diagnosis. Yet it is worth noting that many of these biomarkers only can be detected from tissue, such as synovium tissues, obtained by biopsies. By contrast, those that can be detected in peripheral blood through non-invasion are more useful (35).

Ribonucleotide reductase M2 (RRM2) acts as a subunit of ribonucleotide reductase (RR) which is important for DNA replication and damage repair via providing deoxy- ribonucleoside triphosphates (dNTPs) (36, 37). Many studied have showed that RRM2 plays a key role in tumour cell proliferation, apoptosis, DNA damage and epithelial mesenchymal (38-40). Therefore, RRM2 is regarded as a key gene in tumour metastasis and progression, as well as a promising biomarker for a variety of cancers (41, 42). Meanwhile, previous studies indicated that both mRNA and protein of RRM2 were responsible for chemotherapy sensitivity and resistance (43), such as imatinib-based therapy resistance in chronic myeloid leukaemia (44), adriamycin resistance in breast cancer (45), gemcitabine resistance in advanced lung squamous cell carcinoma (46) and so on. Interestingly, only one report about RRM2 was associated with RA, using liposome-polycation-dNA (LPd) complex loaded with RRM2 small interfering RNA, suggested that suppressed RRM2 gene may cause the downregulation of the levels of proinflammatory cytokines TNF-α and IL-6 in RA-fibroblast-like synoviocytes (RA-FLSs) via increasing the levels of apoptosis and inhibiting the proliferation of RA-FLSs (47). Evidence from this study elucidated that RRM2 could play a critical role in the pathogenesis of RA. However, we could not find positive correlations between RRM2 and inflammatory factors including TNF-α and IL-6 in plasma. These inconsistent results indicated that RRM2 has a cell and tissue-specific expression pattern in RA. In the light of the complex biological role of RRM2, more studies are needed to explore the mechanism and function of RRM2 in RA.

Our study has several potential limitations. The first weakness was that the population of cohorts used for high-throughput sequencing analyses was small. Secondly, the individual samples we collected and those in GSE93272 dataset used to validate the hub genes and evaluate their correlation with clinical characters were all Japanese and Chinese; more samples from other countries are needed to verify these results. Thirdly, the number of early RA patients is too low to provide secure data regarding the eventual discriminating role of RRM2 between early and established RA. Similarly, the number of patients with active RA is not well balanced by the number of patients with remitting disease, which may lead to a high sensitivity. Finally, more research is needed to investigate the utility of RRM2 in the differential diagnosis of other autoimmune diseases such as ankylosing spondylitis.

Conclusion

In summary, this study identified 3 newly discovered mRNAs associated with RA and their expression levels significantly increased in PBMC in RA. With the validations *in vivo* and *vitro*, RRM2 showed high diagnosis power with extremely high levels of sensitivity (86.7.%) and specificity (90.4%). Moreover, RRM2 could improve the diagnostic efficiency for RA patients with anti-CCP$^-$. Therefore, the findings provide a novel candidate biomarker for the diagnosis of RA.

Acknowledgement

We are grateful to Xianghe Xu for his suggestions and conceptualisation of the research design.

References

1. SMOLEN JS, ALETAHA D, BARTON A *et al*.: Rheumatoid arthritis. *Nat Rev Dis Prim* 2018; 4: 1-23. https://doi.org/10.1038/nrdp. 2018.1
2. ALETAHA D, SMOLEN JS: Diagnosis and management of rheumatoid arthritis: a review. *JAMA* 2018; 320: 1360-72.
3. LIN Y-J, ANZAGHE M, SCHÜLKE S: Update on the pathomechanism, diagnosis, and treatment options for rheumatoid arthritis. *Cells* 2020; 9: 880. https://doi.org/10.3390/cells9040880
4. OKADA Y, EYRE S, SUZUKI A, KOCHI Y, YAMAMOTO K: Genetics of rheumatoid arthritis: 2018 status. *Ann Rheum Dis* 2019; 78: 446-53. https://doi.org/10.1136/annrheumdis-2018-213678
5. SMOLEN JS, ALETAHA D, MCINNES IB: Rheumatoid arthritis. *Lancet* 2016; 388: 2023-38. https://doi.org/10.1016/s0140-6736 (16)30173-8
6. GIANNINI D, ANTONUCCI M, PETRELLI F, BILIA S, ALUNNO A, PUXEDDU I: One year in review 2020: pathogenesis of rheumatoid arthritis. *Clin Exp Rheumatol* 2020; 38: 387-97.
7. VIATTE S, BARTON A: Genetics of rheumatoid arthritis susceptibility, severity, and treatment response. *Semin Immunopathol* 2017; 39: 395-408. https://doi.org/10.1007/s00281-017-0630-4
8. SUZUKI A, YAMADA R, CHANG X *et al*.: Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* 2003; 34: 395-402. https://doi.org/10.1038/ng1206
9. AI R, LARAGIONE T, HAMMAKER D *et al*.: Comprehensive epigenetic landscape of rheumatoid arthritis fibroblast-like synoviocytes. *Nat Commun* 2018; 9: 1921. https://doi.org/10.1038/s41467-018-04310-9
10. SPARKS JA: Rheumatoid arthritis. *Ann Intern Med* 2019; 170: ITC1-16 https://doi.org/10.7326/aitc201901010

11. BERGLIN E, DAHLQVIST SR: Comparison of the 1987 ACR and 2010 ACR/EULAR classification criteria for rheumatoid arthritis in clinical practice: A prospective cohort study. *Scand J Rheumatol* 2013; 42: 362-8. https://doi.org/10.3109/03009742.2013.776103
12. KOURILOVITCH M, GALARZA-MALDONADO C, ORTIZ-PRADO E: Diagnosis and classification of rheumatoid arthritis. *J Autoimmun* 2014; 48-49: 26-30. https://doi.org/10.1016/j.jaut.2014.01.027
13. ALLEN A, CARVILLE S, McKENNA F: Diagnosis and management of rheumatoid arthritis in adults: Summary of updated NICE guidance. *BMJ* 2018; 362: k3015. https://doi.org/10.1136/bmj.k3015
14. LITTLEJOHN EA, MONRAD SU: Early diagnosis and treatment of rheumatoid arthritis. *Prim Care* 2018; 45: 237-55. https://doi.org/10.1016/j.pop.2018.02.010
15. TAKASE-MINEGISHI K, HORITA N, KOBAYASHI K *et al*.: Diagnostic test accuracy of ultrasound for synovitis in rheumatoid arthritis: systematic review and meta-analysis. *Rheumatology* (Oxford) 2018; 57: 49-58. https://doi.org/10.1093/rheumatology/kex036
16. WHITING PF, SMIDT N, STERNE JAC *et al*.: Systematic review: Accuracy of anti-citrullinated peptide antibodies for diagnosing rheumatoid arthritis. *Ann Intern Med* 2010; 152: 456-64. https://doi.org/10.7326/0003-4819-152-7-201004060-00010
17. BARRETT T, WILHITE SE, LEDOUX P *et al*.: NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Res* 2013; 41: D991. https://doi.org/10.1093/nar/gks1193
18. TASAKI S, SUZUKI K, KASSAI Y *et al*.: Multi-omics monitoring of drug response in rheumatoid arthritis in pursuit of molecular remission. *Nat Commun* 2018; 9: 2755. https://doi.org/10.1038/s41467-018-05044-4
19. RITCHIE ME, PHIPSON B, WU D *et al*.: Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47. https://doi.org/10.1093/nar/gkv007
20. YU G, WANG LG, HAN Y, HE QY: Cluster Profiler: An R package for comparing biological themes among gene clusters. *Omi A J Integr Biol* 2012; 16: 284-7. https://doi.org/10.1089/omi.2011.0118
21. SZKLARCZYK D, GABLE AL, LYON D *et al*.: STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019; 47: D607-13. https://doi.org/10.1093/nar/gky 1131
22. SHANNON P, MARKIEL A, OZIER O *et al*.: Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498-504. https://doi.org/10.1101/gr.1239303
23. CHIN CH, CHEN SH, WU HH, HO CW, KO MT, LIN CY: cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014; 8 (Suppl. 4): S11. https://doi.org/10.1186/1752-0509-8-s4-s11
24. BADER GD, HOGUE CWV: An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 2003; 4: 2. https://doi.org/10.1186/1471-2105-4-2
25. FONSECA LJS DA, NUNES-SOUZA V, GOULART MOF, RABELO LA: Oxidative stress in rheumatoid arthritis: what the future might hold regarding novel biomarkers and add-on therapies. *Oxid Med Cell Longev* 2019; 2019: 7536805. https://doi.org/10.1155/2019/7536805
26. ALETAHA D, NEOGI T, SILMAN AJ *et al*.: 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010; 69: 1580-8. https://doi.org/10.1136/ard.2010.138461
27. ALETAHA D, ALASTI F, SMOLEN JS: Rheumatoid factor, not antibodies against citrullinated proteins, is associated with baseline disease activity in rheumatoid arthritis clinical trials. *Arthritis Res Ther* 2015; 17: 229. https://doi.org/10.1186/s13075-015-0736-9
28. SOKOLOVE J, JOHNSON DS, LAHEY LJ *et al*.: Rheumatoid factor as a potentiator of anti-citrullinated protein antibody-mediated inflammation in rheumatoid arthritis. *Arthritis Rheumatol* 2014; 66: 813-21. https://doi.org/10.1002/art.38307
29. MATUSZEWSKA A, MADEJ M, WILAND P: Immunological markers of rheumatoid arthritis. *Postepy Hig Med Dosw* 2016; 70: 251-7. https://doi.org/10.5604/17322693.1198270
30. NIELEN MMJ, VAN SCHAARDENBURG D, REESINK HW *et al*.: Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum* 2004; 50: 380-6. https://doi.org/10.1002/art.20018
31. SMOLEN JS, ALETAHA D, BARTON A: Rheumatoid arthritis. *Nat Rev Dis Prim* 2018; 4: 18001. https://doi.org/10.1038/nrdp.2018.1
32. WANG J, YAN S, YANG J, LU H, XU D, WANG Z: Non-coding RNAs in rheumatoid arthritis: from bench to bedside. *Front Immunol* 2020; 10: 3129. https://doi.org/10.3389/fimmu.2019.03129
33. CHUROV AV, OLEINIK EK, KNIP M: MicroRNAs in rheumatoid arthritis: Altered expression and diagnostic potential. *Autoimmun Rev* 2015; 14: 1029-37. https://doi.org/10.1016/j.autrev.2015.07.005
34. MO BY, GUO XH, YANG MR *et al*.: Long noncoding RNA GAPLINC promotes tumor-like biologic behaviors of fibroblast-like synoviocytes as microRNA sponging in rheumatoid arthritis patients. *Front Immunol* 2018; 9: 702. https://doi.org/10.3389/fimmu.2018.00702
35. ATZENI F, TALOTTA R, MASALA IF, BONGIOVANNI S, BOCCASSINI L, SARZI-PUTTINI P: Biomarkers in rheumatoid arthritis. *Isr Med Assoc J* 2017; 19: 512-6.
36. DUXBURY MS, ITO H, ZINNER MJ, ASHLEY SW, WHANG EE: RNA interference targeting the M2 subunit of ribonucleotide reductase enhances pancreatic adenocarcinoma chemosensitivity to gemcitabine. *Oncogene* 2004; 23: 1539-48. https://doi.org/10.1038/sj.onc.1207272
37. NORD-UND P, REICHARD P: Ribonucleotide reductases. *Ann Rev Biochem* 2006; 75: 681-706. https://doi.org/10.1146/annurev.biochem.75.103004.142443
38. MAZZU YZ, ARMENIA J, CHAKRABORTY G *et al*.: A novel mechanism driving poor-prognosis prostate cancer: Overexpression of the DNA repair gene, ribonucleotide reductase small subunit M2 (RRM2). *Clin Cancer Res* 2019; 25: 4480-92. https://doi.org/10.1158/1078-0432.ccr-18-4046
39. GANDHI M, GROß M, HOLLER JM *et al*.: The lncRNA lincNMR regulates nucleotide metabolism via a YBX1 - RRM2 axis in cancer. *Nat Commun* 2020; 11: 3214. https://doi.org/10.1038/s41467-020-17007-940.
RASMUSSEN RD, GAJJAR MK, TUCKOVA L *et al*.: BRCA1-regulated RRM2 expression protects glioblastoma cells from endogenous replication stress and promotes tumorigenicity. *Nat Commun* 2016; 7: 13398. https://doi.org/10.1038/ncomms13398
41. GAO J, CHEN H, YU Y *et al*.: Inhibition of hepatocellular carcinoma growth using immunoliposomes for co-delivery of adriamycin and ribonucleotide reductase M2 siRNA. *Biomaterials* 2013; 34: 10084-98. https://doi.org/10.1016/j.biomaterials.2013.08.088
42. LIU X, ZHANG H, LAI L *et al*.: Ribonucleotide reductase small subunit M2 serves as a prognostic biomarker and predicts poor survival of colorectal cancers. *Clin Sci* 2013; 124: 567-78. https://doi.org/10.1042/cs20120240
43. ZHAN Y, JIANG L, JIN X *et al*.: Inhibiting RRM2 to enhance the anticancer activity of chemotherapy. *Biomed Pharmacother* 2021; 133: 110996. https://doi.org/10.1016/j.biopha.2020.110996
44. LIU C, LI Y, HU R, HAN W, GAO S: Knockdown of ribonucleotide reductase regulatory subunit M2 increases the drug sensitivity of chronic myeloid leukemia to imatinib-based therapy. *Oncol Rep* 2019; 42: 571-80. https://doi.org/10.3892/or.2019.7194
45. ZHOU S, LI J, XU H *et al*.: Liposomal curcumin alters chemosensitivity of breast cancer cells to Adriamycin via regulating microRNA expression. *Gene* 2017; 622: 1-12. https://doi.org/10.1016/j.gene.2017.04.026
46. CHEN P, WU JN, SHU Y *et al*.: Gemcitabine resistance mediated by ribonucleotide reductase M2 in lung squamous cell carcinoma is reversed by GW8510 through autophagy induction. *Clin Sci* 2018; 132: 1417-33. https://doi.org/10.1042/cs20180010
47. WANG X, WANG X, SUN J, FU S: An enhanced RRM2 siRNA delivery to rheumatoid arthritis fibroblast-like synoviocytes through a liposome-protamine-DNA-siRNA complex with cell permeable peptides. *Int J Mol Med* 2018; 42: 2393-402. https://doi.org/10.3892/ijmm.2018.3815