

RESEARCH ARTICLE

Open Access



# RNA sequencing and bioinformatics analysis of differentially expressed genes in the peripheral serum of ankylosing spondylitis patients

Yongchen Bie<sup>1†</sup>, Xiujun Zheng<sup>1†</sup>, Xiaojiong Chen<sup>1</sup>, Xiangyun Liu<sup>1</sup>, Liqin Wang<sup>2</sup>, Yuanliang Sun<sup>1</sup> and Jianqiang Kou<sup>1\*</sup>

## Abstract

**Background** Ankylosing spondylitis (AS) is a chronic progressive autoimmune disease characterized by spinal and sacroiliac arthritis, but its pathogenesis and genetic basis are largely unclear.

**Methods** We randomly selected three serum samples each from an AS and a normal control (NC) group for high-throughput sequencing followed by using edgeR to find differentially expressed genes (DEGs). Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes, Reactome pathway analyses, and Gene Set Enrichment Analysis were used to comprehensively analyze the possible functions and pathways involved with these DEGs. Protein–protein interaction (PPI) networks were constructed using the STRING database and Cytoscape. The modules and hub genes of these DEGs were identified using MCODE and CytoHubba plugins. Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) was used to validate the expression levels of candidate genes in serum samples from AS patients and healthy controls.

**Results** We successfully identified 100 significant DEGs in serum. When we compared them with the NC group, 49 of these genes were upregulated in AS patients and 51 were downregulated. GO function and pathway enrichment analysis indicated that these DEGs were mainly enriched in several signaling pathways associated with endoplasmic reticulum stress, including protein processing in the endoplasmic reticulum, unfolded protein response, and ubiquitin-mediated proteolysis. We also constructed a PPI network and identified the highly connected top 10 hub genes. The expression levels of the candidate hub genes *PPARG*, *MDM2*, *DNA2*, *STUB1*, *UBTF*, and *SLC25A37* were then validated by RT-qPCR analysis. Finally, receiver operating characteristic curve analysis suggested that *PPARG* and *MDM2* may be the potential biomarkers of AS.

**Conclusions** These findings may help to further elucidate the pathogenesis of AS and provide valuable potential gene biomarkers or targets for the diagnosis and treatment of AS.

**Keywords** Differentially expressed genes, RNA sequencing, Serum, Ankylosing spondylitis, Biomarker

<sup>†</sup>Yongchen Bie and Xiujun Zheng contributed equally to this paper.

\*Correspondence:

Jianqiang Kou

[koujianq@qduhospital.cn](mailto:koujianq@qduhospital.cn)

Full list of author information is available at the end of the article



## Introduction

Ankylosing spondylitis (AS) is a chronic inflammatory arthritis and autoimmune disease with an incidence of 0.1–1.4% [1]. This disease is primarily characterized by spinal and sacroiliac arthritis [2] and is frequently associated with extra-articular features including anterior uveitis, psoriasis, or inflammatory bowel disease [3–5]. The stiffness of the affected joints becomes severe with the development of ankylosing spondylitis, leading to back pain, poor quality of life, and in more serious cases, mental illness [6, 7]. Currently, the etiology of AS is multifactorial; both environmental and genetic predispositions have been suggested to be involved in AS pathogenesis [8], including macrophage activation status, infections with particular bacteria, inflammatory cytokines, and autophagy [9–12]. Nevertheless, few genes have been shown to be associated with the disease and the actual cause of AS has remained unclear. Therefore, there is an urgent need to identify new biomarkers that can act as reliable diagnostic or prognostic indicators of AS. Such biomarkers will be invaluable in the prevention, treatment, and control of this disease.

In recent decades, ribonucleic acid (RNA) sequencing has proven to be a novel high-throughput sequencing method that uses deep-sequencing technology [13]. This approach can be used to identify abnormally spliced genes, detect allele-specific expression, and identify differentially expressed genes (DEGs). Bioinformatics analysis can use sequencing data to analyze the genome, transcriptome, and proteome information of organisms and has been used to reveal the mechanisms of disease that occurs due to abnormal biological processes at the molecular level [14, 15]. To date, multiple studies have used microarray expression and RNA sequencing to identify DEGs involved in the pathogenesis of AS [16, 17]. Peripheral blood is often considered a potential resource for the discovery of disease biomarkers [18]. However, the levels of DEGs in peripheral serum from patients with AS have not been explored, nor have these molecular mechanisms been further validated. Altered gene expression profiles that differentiate disease from healthy can be used as a basis for understanding the pathogenesis of AS.

In the current study, by expression profiling of high-throughput sequencing and experimental analysis, we identified DEGs in the serum of AS patients and normal controls. Then, a molecular mechanism of AS was proposed after analyzing these pathways and functional enrichments. Finally, with the use of these DEGs, we established the protein–protein interaction (PPI) network to identify hub genes for targeting AS.

## Materials and methods

### Patients and samples

A total of 18 patients with AS and 18 healthy age-matched controls were selected from the Affiliated Hospital of Qingdao University (Qingdao, China) between December 2020 and September 2021. All patients met the modified New York 1984 criteria [19] and were initially diagnosed with AS, drug-naive patients with short disease durations. None of the patients or controls had any previous history of cardiovascular disease, diabetes, hepatitis, malignancy, or other autoimmune and inflammatory illnesses.

Serum samples were collected using standard phlebotomy procedures and centrifuged at 3000 g for 10 min. The separated sera were stored in RNase-free centrifuge tubes at 80 °C until further processing. This research was approved and reviewed by the Medical Ethics Review Committee of the Affiliated Hospital of Qingdao University (approval number: QYFY WZLL 27251). All participants provided written informed consent in accordance with policies of the hospital ethics committee.

### RNA extraction and sequencing

We randomly selected three patients with AS and three normal control (NC) for high-throughput sequencing [20]. Total RNA was extracted from serum using TRIzol reagent (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA) according to the manufacturer's instructions. Subsequently, the concentration and integrity of the isolated RNA were determined using the Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) and an Agilent 2100 Bioanalyzer (Applied Biosystems, Carlsbad, CA, USA), respectively. RNA-seq libraries were prepared using the SMARTer Stranded Total RNA-Seq kit v.2 (Takara Bio USA, Mountain View, CA, USA) as previously [21]. The RNA samples were fragmented and reversely transcribed into first-strand cDNA, followed by second-strand synthesis. After cDNA synthesis, a tailing and adapter ligation was performed, and then the cDNA was amplified by PCR. Subsequently, the cDNA library quality and concentration were evaluated using the Agilent 2100 Bioanalyzer (Applied Biosystems, Carlsbad, CA, USA). The qPCR-based KAPA Biosystems Library Quantification kit (Kapa Biosystems, Inc.) was used for the quantification of the cDNA library. Ribosomal RNA depletion was performed during library construction according to the manufacturer's protocol. Sequencing was carried out in a 150-bp paired-end run (PE150) using the NovaSeq 6000 system (Illumina, San Diego, CA, USA).

**Table 1** Primers used in the present study

| Gene     | Forward primer            | Reverse primer            |
|----------|---------------------------|---------------------------|
| PPARG    | GCCCTTCACTACTGTTGACTTCTCC | CAGGCTCCACTTTGATTGCACTTTG |
| MDM2     | AGGCAGGGGAGAGTGATACAGATTC | CAGGAAGCCAATTCTCACGAAGGG  |
| DNA2     | GAAACCCAGCATCTGAAGCAAACAC | TCTCCATTTCGGAAGCAGGCATTAG |
| STUB1    | TCCTACCTCTCCAGGCTCATTGC   | ATGTCCGCCATGTACTTGTCTGTG  |
| UBTF     | ATCTCCCAGAGCCAGAAGGAG     | GGGAGACAGGCTCTTAACCCA     |
| SLC25A37 | CCTTCTACCGAGCTACACCA      | CCTGAGATGATGTGGGACTGC     |
| GAPDH    | GCACCGTCAAGGCTGAGAAC      | TGGTGAAGACGCCAGTGGGA      |

### Analysis of DEGs

Reads were first mapped to the latest UCSC transcript set using Bowtie2 version 2.1.0 [22] and gene expression levels were estimated using RSEM v1.2.15 [23]. The trimmed mean of M-values was used to standardize gene expression. DEGs were then identified using edgeR software [24]. Genes showing altered expression with  $p < 0.05$  and more than 1.5-fold changes were considered to be differentially expressed.

### Gene ontology (GO), kyoto encyclopedia of genes and genomes (KEGG), and reactome analyses

To better understand the function and pathways of DEGs in AS, we performed GO functional annotation, KEGG enrichment [25], and Reactome analyses using the R package clusterProfiler [26]. GO analysis was used to investigate biological functions based on differentially expressed coding genes. This analysis classifies functions according to the three following aspects: biological process (BP), cellular component (CC), and molecular function (MF). A lower  $p$ -value indicated a higher significance of a GO term ( $p$ -value  $< 0.05$ ). The KEGG and Reactome enrichment analyses were used to predict the related pathways of each DEG. A  $p$ -value of  $< 0.05$  reflected significant enrichment.

### Gene set enrichment analysis (GSEA)

GSEA was done as described in Subramanian et al. [27]. We used the R package fgsea to analyze the expression of filtered genes against MSigDB, a well-known molecular feature database. In addition, only two typical gene sets from MSigDB, H (hallmark gene sets) and C2: CP (curated gene sets, canonical pathways), were analyzed by GSEA here. Finally, we retained results with statistical significance  $p$ -values  $< 0.05$ .

### Construction of a protein–protein interaction (PPI) network and identification of hub genes and key modules

To gain insights into the correlation among DEGs at the protein level, the Search Tool for the Retrieval of

Interacting Genes (STRING, <https://stringdb.org>) database was used to construct a PPI network of these DEGs [28]. The minimum required interaction score used to construct this PPI network was 0.4, and the isolated nodes were abandoned. Cytoscape [29] was used to visualize this PPI network and we used the plug-in Cytohubba to explore the hub genes of this PPI network [30]. Based on the centrality score, the key nodes in this PPI network were determined, and then the hub genes were deduced. Simultaneously, we used the molecular complex detection (MCODE) plugin for clustering analysis of gene networks to select the key subnetwork modules.

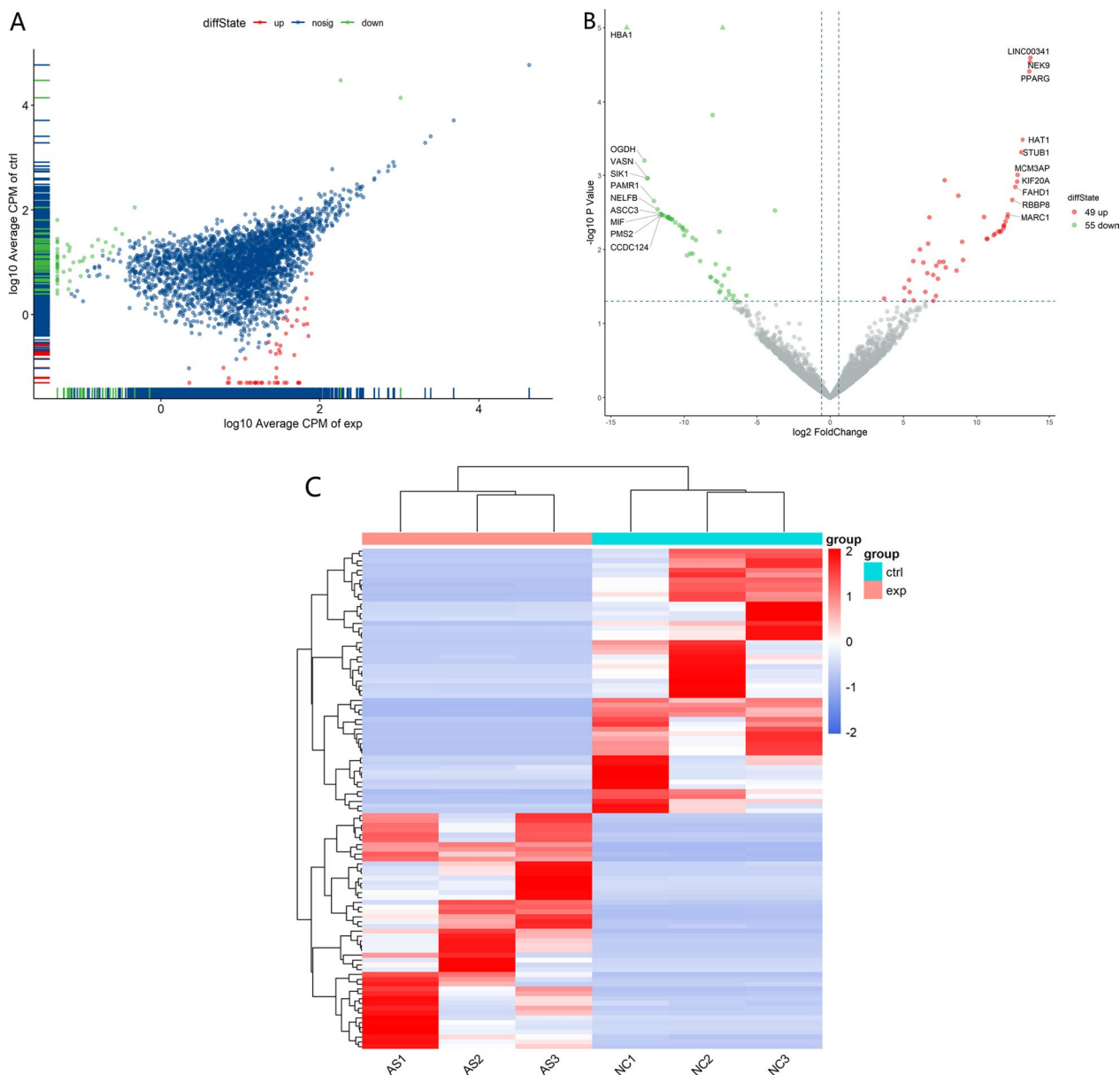
### Validation of DEGs

Reverse transcription-quantitative polymerase chain reaction (qRT-PCR) experiments were conducted to validate the DEGs identified using high-throughput sequencing. cDNA was synthesized using the Prime Script RT reagent kit with genomic DNA eraser (TaKaRa, Tokyo, Japan), and qRT-PCR was performed on a Light-Cycler 480 (Roche, Indianapolis, IN, USA) using SYBR

**Table 2** Characteristics of AS patients and NC

| Items                     | NC (n = 18) | AS (n = 18) |
|---------------------------|-------------|-------------|
| Age (years)               | 31.6 ± 7.2  | 32.1 ± 7.7  |
| Gender, No. (%)           |             |             |
| Female                    | 3 (16.7)    | 2 (11.1)    |
| Male                      | 15 (83.3)   | 16 (88.9)   |
| CRP (mg/L)                | –           | 15.4 ± 8.7  |
| ESR (mm/H)                | –           | 32.2 ± 18.9 |
| HLA-B27 positive, No. (%) |             |             |
| No                        | –           | 1 (6.6)     |
| Yes                       | –           | 17 (94.4)   |
| BASDAI                    | –           | 4.1 ± 1.3   |
| BASFI                     | –           | 3.7 ± 1.2   |
| VAS                       | –           | 5.8 ± 1.8   |

CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; HLA-B27, human leukocyte antigen-B27; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; VAS, Visual Analogue Scale



**Fig. 1** **A** scatter plot and **B** volcano plot. The red dots represent upregulated DEGs, and the green dots represent downregulated DEGs. **C** Hierarchical clustering of differentially expressed genes between the AS and control groups

Green Master Mix (TaKaRa, Tokyo, Japan). The primer sequences are listed in Table 1. Validation experiments were performed using serum from 15 AS samples and 15 NC samples. mRNA levels of the selected new target genes were quantified by the  $2^{-\Delta\Delta C_t}$  method after normalization to the housekeeping gene *GAPDH*.

**Receiver operating characteristic (ROC) analyses**

To assess the diagnostic value of DEGs in AS, we performed ROC analyses using the pROC R package. We

calculated the area under the curve (AUC) under the binomial exact confidence interval and Ggplot2 was applied for further visualization.

**Statistical analysis**

Statistical analysis was performed using SPSS 26.0 software (SPSS Inc., Chicago, IL, USA). All data are expressed as the mean  $\pm$  SD. Statistical significance was determined by Student’s *t* test, and *P* values < 0.05 were considered to indicate a statistically significant difference.

**Table 3** Top ten upregulated and ten downregulated DEGs in AS

| Gene name | logFC        | Fold change  | P value     | Regulation |
|-----------|--------------|--------------|-------------|------------|
| SYNE3     | 13.70024767  | 13,310.22803 | 2.53347E-05 | Up         |
| NEK9      | 13.66669431  | 13,004.23858 | 2.90169E-05 | Up         |
| PPARG     | 13.62684124  | 12,649.92523 | 3.87972E-05 | Up         |
| HAT1      | 13.15463256  | 9118.827157  | 0.000323846 | Up         |
| STUB      | 13.0556646   | 8514.255069  | 0.000478438 | Up         |
| MCM3AP    | 12.78714097  | 7068.267708  | 0.000980806 | Up         |
| KIF20A    | 12.77003424  | 6984.950792  | 0.001209986 | Up         |
| FAHD1     | 12.64685533  | 6413.318437  | 0.001424853 | Up         |
| RBBP8     | 12.43477239  | 5536.553204  | 0.002137449 | Up         |
| MARC1     | 12.1303543   | 4483.327884  | 0.003358303 | Up         |
| MIF       | -11.57772609 | 0.000327158  | 0.003321108 | Down       |
| ASCC3     | -11.5932764  | 0.00032365   | 0.003319781 | Down       |
| CCDC124   | -11.6024601  | 0.000321597  | 0.003347325 | Down       |
| PMS2      | -11.6352721  | 0.000314365  | 0.003485706 | Down       |
| NELFB     | -11.82545438 | 0.000275539  | 0.002867456 | Down       |
| PAMR1     | -12.07196817 | 0.000232261  | 0.002207181 | Down       |
| VASN      | -12.48849677 | 0.000174015  | 0.001089923 | Down       |
| SIK1      | -12.54756591 | 0.000167035  | 0.001085811 | Down       |
| OGDH      | -12.71646833 | 0.00014858   | 0.000627557 | Down       |
| HBA1      | -13.92287772 | 6.43867E-05  | 3.65243E-06 | Down       |

## Results

### Characteristics of AS and NC patients

Table 2 presents the demographic and clinical characteristics of the 18 patients with AS and 18 NC in the present study. Patients with AS and NC were matched in terms of age and sex.

### Identification of DEGs in AS

In the AS Peripheral Blood Serum samples, a total of 100 DEGs were differentially expressed, including 49 upregulated DEGs and 51 downregulated DEGs, respectively. These DEGs are represented by scatter and volcano plots in Fig. 1A and B, respectively. Figure 1C shows the hierarchical clustering of these DEGs. The ten most upregulated and downregulated DEGs are shown in Table 3.

### GO and pathway enrichment analyses

GO function enrichment analysis was conducted on 100 DEGs using the R package clusterProfiler, and 167 GO items with significant differences were identified, including 130 BP items, 16 CC items, and 21 MF items. It can be seen from the graph that the biological processes of BP mediated by DEGs were mainly concentrated in response to toxic substance, response to the antibiotic, and ER unfolded protein response. The results for CC were mainly concentrated in cytoplasmic vesicle lumen,

vesicle lumen, and mitochondrial matrix. The results for MF showed that fatty acid binding, DNA helicase activity, and histone acetyltransferase activity were significantly enriched items (Fig. 2A).

KEGG pathway enrichment analysis of these DEGs yielded numerous signaling pathways that were significantly altered in the serum of AS. A total of six KEGG pathways were enriched for these DEGs (Fig. 2B), and protein processing in the endoplasmic reticulum (ER) was the most significantly enriched KEGG pathway associated with AS. For validating the biofunctions related to such DEGs, Reactome is another well-known signaling pathway database. We also identified six pathways based on Reactome analysis in the present study (Fig. 2C). The most enriched Reactome terms of AS were neutrophil degranulation and unfolded protein response (UPR).

### GSEA-enriched pathways

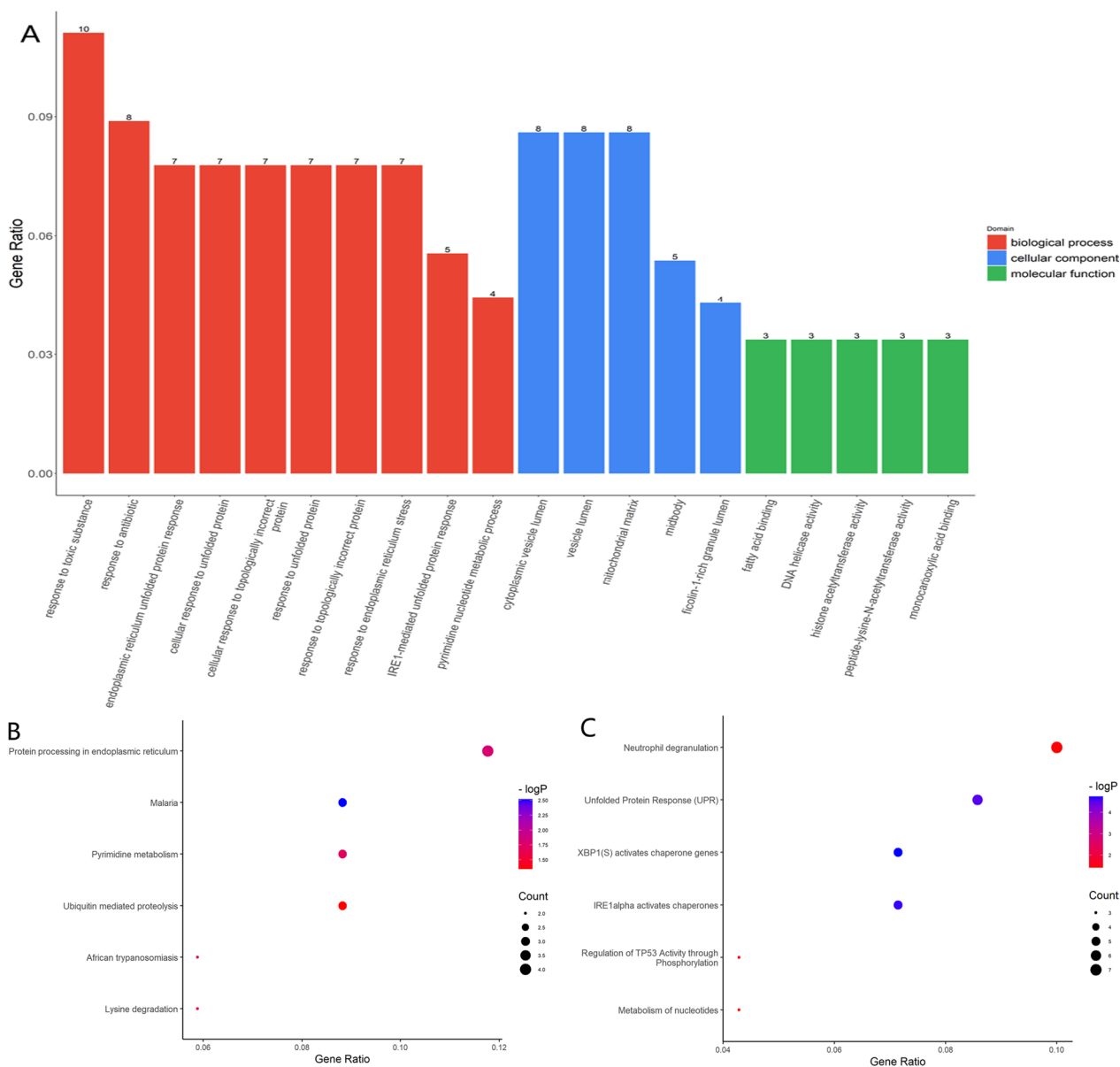
Simultaneously, we utilized GSEA to identify the potential mechanism underlying AS. We analyzed 100 top genes and identified 17 pathways that were enriched significantly. Among the pathways with the highest enrichment scores were the nr1h2 and nr1h3 mediated signaling pathways and the hnf3b pathway. We visualized the top ten pathways ( $\leq 10$ ) with the most significant activation and inhibition by GSEA (Fig. 2D).

### PPI network analysis

A total of 55 proteins and 58 edges were obtained with a cutoff value of credibility  $> 0.4$ , as shown in Fig. 3A. Isolated nodes from the PPI network were abandoned. The hub genes were selected from the PPI network of AS-related genes using the cytoHubba plugin. The results demonstrated that ten hub genes could be identified using the Maximal Clique Centrality (MCC) algorithm, including *PPARG*, *MDM2*, *DNA2*, *STUB1*, *UBTF*, *SLC25A37*, *TICRR*, *RBBP8*, *DDX11*, and *NME2*, as shown in Fig. 3B. A summary of the degree values for the ten genes is provided in Table 4. Of these, *PPARG* (degree=6) and *MDM2* (degree=6) were identified as the most significant genes. The key subnetwork modules were identified using the Cytoscape plug-in MCODE. According to the degree of importance, two significant modules were selected from this PPI network. The results revealed two modules as shown in Fig. 3C and D. Module 1 was composed of seven nodes and nine edges (score: 3.000), followed by module 2, which was composed of three nodes and three edges (score: 3.000).

### Verification of DEGs by qRT-PCR

To validate these major results, we selectively performed qRT-PCR analysis of six hub genes including *PPARG*



**Fig. 2** GO, KEGG, Reactome, and GSEA analyses of the differentially expressed genes in AS patient serum. **A** GO analysis. BP, CC, and MF are represented in red, blue, and green, respectively. **B** KEGG analysis. The six KEGG pathways are shown. **C** Reactome analysis. The six Reactome pathways. **D** GSEA enriched pathway. The top ten pathways ( $\leq 10$ ) are shown

, *MDM2*, *DNA2*, *STUB1*, *UBTF*, and *SLC25A37* using RNA from 15 AS patients and 15 healthy controls. The relative gene expression levels of *PPARG*, *MDM2*, *DNA2*, and *STUB1* were significantly higher in blood samples from AS patients when compared with those from healthy controls. Moreover, the expression levels of *SLC25A37* were significantly lower in the AS group relative to the healthy control group (Fig. 4).

**Receiver operating characteristics curve (ROC) analysis of confirmed DEGs in serum**

To explore the possibility of these six genes as diagnostic biomarkers of AS patients, our qRT-PCR data were subjected to ROC analysis to evaluate their diagnostic ability. As shown in Fig. 5, except *UBTF*, the areas under the ROCs of the other five genes were all  $>0.7$ . This finding indicates that these genes may be sensitive biomarkers that can distinguish AS patients from individuals without AS.

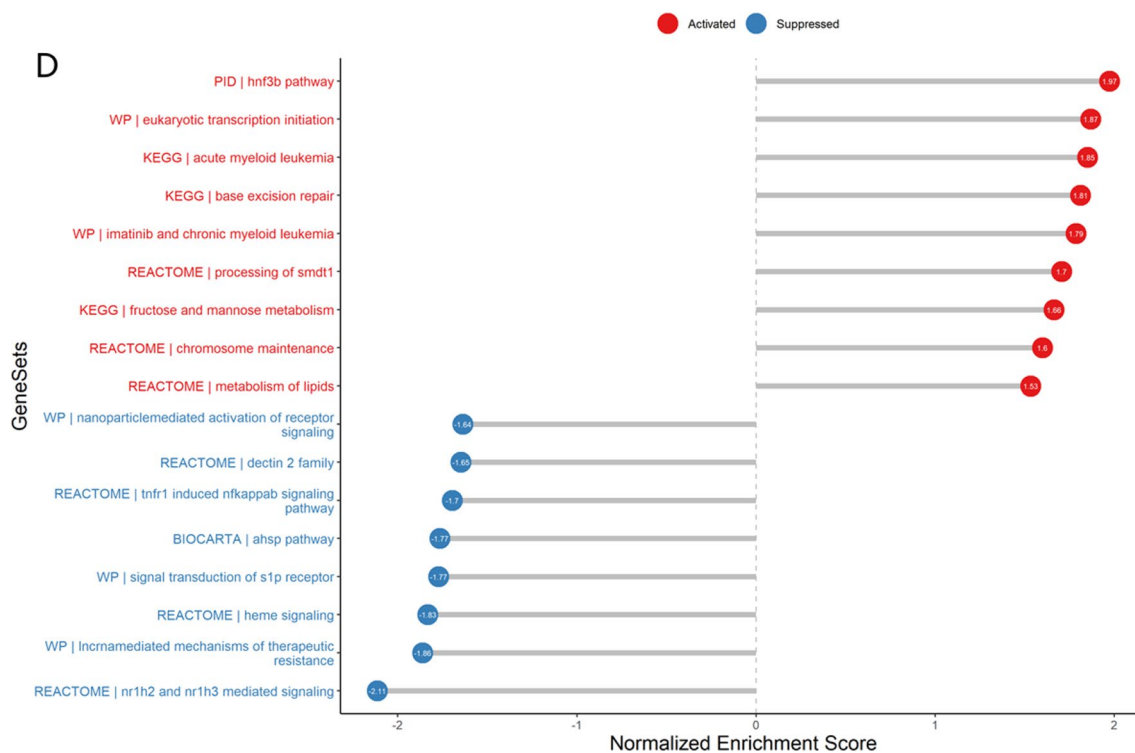


Fig. 2 continued

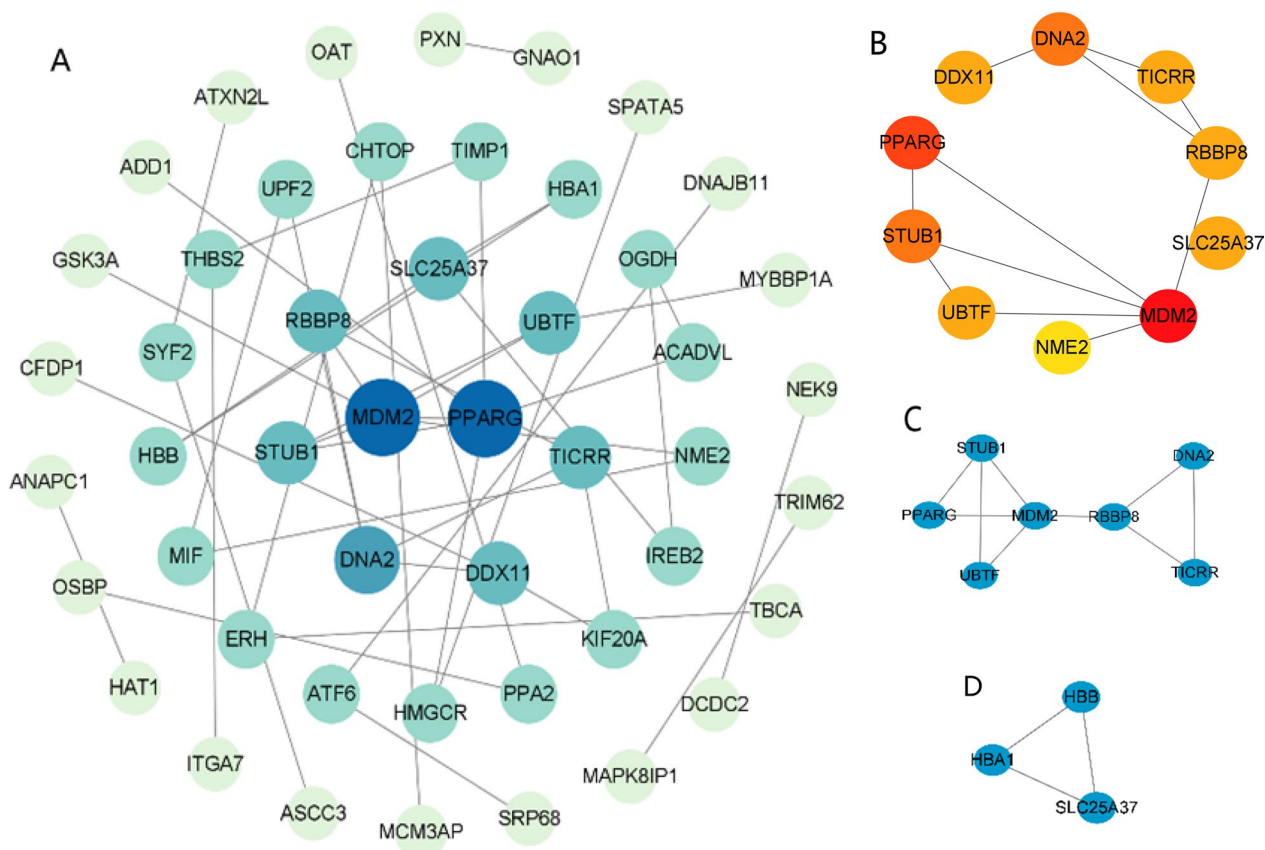
### Discussion

Recently, along with the continuous development of sequencing technologies and bioinformatics technology, studies on the molecular pathogenesis of diseases using bioinformatics tools have emerged in the field of biomedicine. Genome-wide genetic screens have identified a number of genes contributing to AS, but none has yet fully explained the disease process [31]. Previous studies have found that DEGs were abundantly expressed in synovial tissue, whole blood and peripheral blood mononuclear cells in AS [32–34]. To date, there has been no research highlighting the expression profiles of DEGs in serum. Taking this into account, we sought to reveal the role of serum DEGs in the pathogenesis of AS. In the present study, we adopted high-throughput sequencing to analyze DEG profiles in serum isolated from AS patients and controls. We successfully identified 100 DEGs and then performed a range of analyses (GO, KEGG pathway, GSEA, and PPI) in an attempt to uncover novel insights into DEG functions in AS.

In this work, through annotation and functional enrichment analysis, we revealed that numerous genes associated with unfolded protein response and immunoreaction and apoptosis may play a key role in AS, GO and pathway analysis of DEGs demonstrated that the DEGs were mainly enriched in protein processing in ER,

response to endoplasmic reticulum stress (ERS), UPR, neutrophil activation involved in immune response and the intrinsic apoptotic signaling pathway. Studies have confirmed that misfolding of the human leucocyte antigen B27 allele (HLA-B27) forms non-native heavy chain dimeric structures [35]. Dimers may accumulate in the ER, resulting in increased ER stress and potentially leading to the onset of pro-inflammatory responses [36]. In addition, misfolding proteins within the ER can induce the UPR, which is a cellular stress response that initiates transcriptional changes whose function is to restore ER homeostasis [37]. Very recently it was shown that M1 macrophages produce a UPR and stimulated ER stress-related IL-23 in AS patients [38]. The UPR is also associated with aberrant distribution and function of plasmacytoid dendritic cells [39]. Taken together, the results of this study in association with other previous studies have indicated a basis for the vital role of ERS and UPR in the pathogenesis and progression of AS.

Thereafter, the STRING database was used to build a PPI network and Cytoscape was used to identify significant module and hub genes [40]. Analysis of the PPI network constructed based on AS-related DEGs identified direct or indirect crosstalk between these genes. According to the MCC algorithm from the CytoHubba plugin in Cytoscape, the top ten AS-related genes were



**Fig. 3** Protein–protein interaction (PPI) networks and modules. **A** PPI network of DEGs was analyzed using Cytoscape software. The size and color of the nodes corresponding to each gene were determined according to the degree of interaction. The size of the nodes reflects the degree value, where the larger the node, the greater the degree value. The closer to the blue node, the higher connectivity between two nodes. **B** PPI network for the top ten hub genes. **C** and **D** Graphic representation of top two significant modules of the PPI network. (**C** Module 1, **D** Module 2)

identified, including *PPARG*, *MDM2*, *DNA2*, *STUB1*, *UBTF*, *SLC25A37*, *TICRR*, *RBBP8*, *DDX11*, and *NME2*. A high degree value indicated that these genes play a pivotal role and modulate the functions of this network. Furthermore, a total of three cluster modules from this

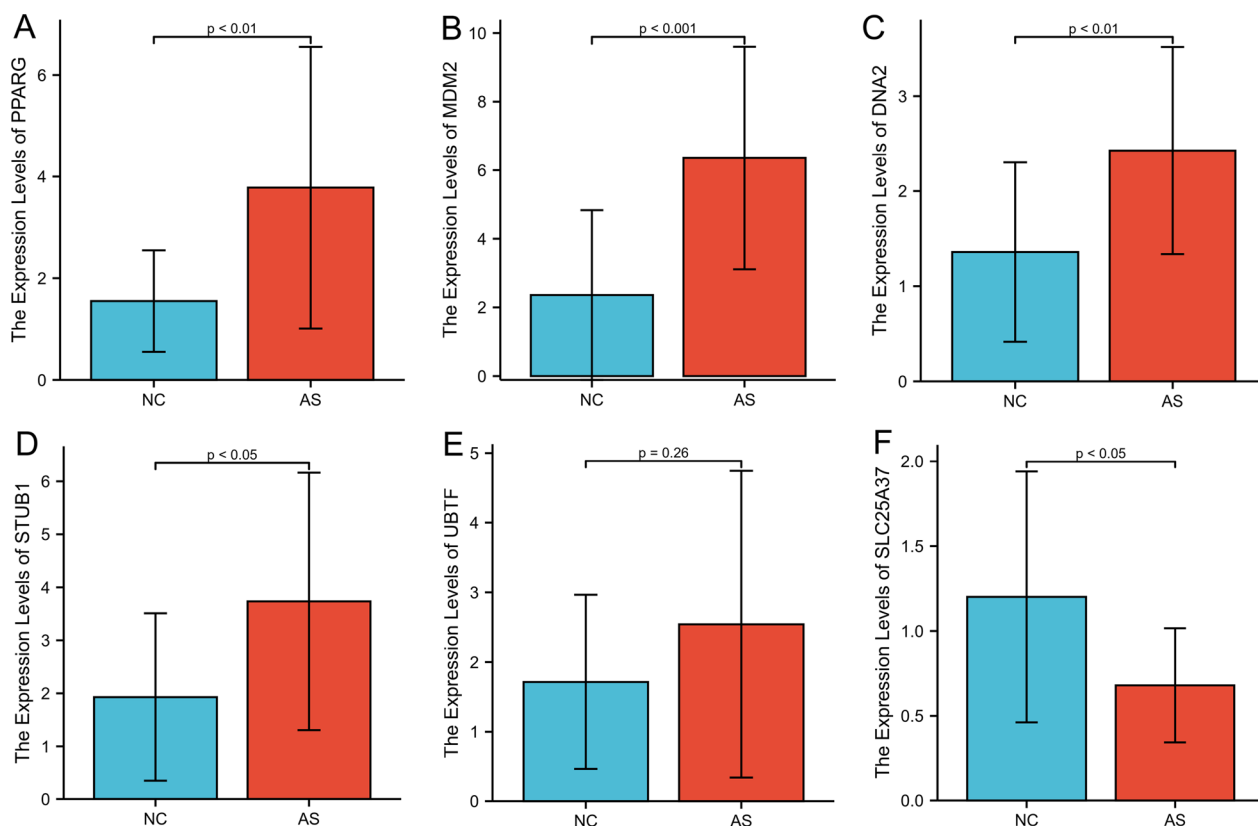
PPI network were extracted by MCODE analysis. Modularization contributed to analyzing the biological functions of the intricate networks of our RNA-seq data in AS. Finally, to validate the results of our bioinformatics analyses, we used qRT-PCR to determine the expression of hub genes that were related to AS, including *PPARG*, *MDM2*, *DNA2*, *STUB1*, *UBTF*, and *SLC25A37*. The results of our qRT-PCR assays provided further verification that our high-throughput sequencing results were reliable. Next, to assess the diagnostic utility of these hub genes, a ROC curve analysis was performed. The AUCs for five hub genes were more than 0.7, suggesting that these genes could effectively distinguish between samples from AS patients and normal controls. Thus, they should be novel and efficient serum indicators of AS in patients.

Several studies on these hub genes are related to the occurrence and development of autoimmune diseases. *PPARG* encodes a member of the peroxisome proliferator-activated receptor (PPAR) subfamily of nuclear receptors. PPARs, especially *PPARG*, contribute to the inhibition of key pro-inflammatory genes such as NF-κB,

**Table 4** The degree values of top ten hub genes

| Gene            | Degree |
|-----------------|--------|
| <i>PPARG</i>    | 6      |
| <i>MDM2</i>     | 6      |
| <i>DNA2</i>     | 4      |
| <i>STUB1</i>    | 3      |
| <i>UBTF</i>     | 3      |
| <i>SLC25A37</i> | 3      |
| <i>TICRR</i>    | 3      |
| <i>RBBP8</i>    | 3      |
| <i>DDX11</i>    | 3      |
| <i>NME2</i>     | 2      |



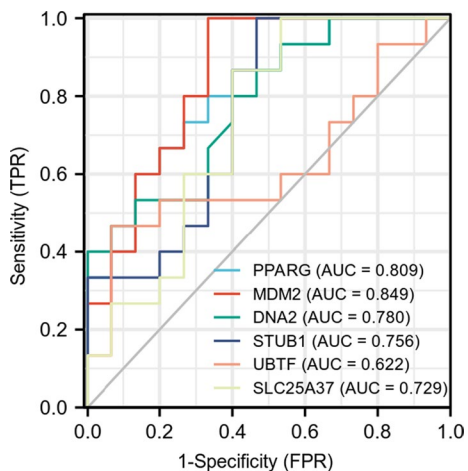


**Fig. 4** qRT-PCR-based validation of the expression of six differentially expressed genes in control and AS patient serum

TNF $\alpha$ , TGF $\beta$ , and the interleukins IL-1 $\alpha$  and IL-6 [41, 42]. In a very recent study of CD14+ monocytes from systemic lupus erythematosus patients, it was reported that there was an emergence of an immunosuppressive M2-phenotype upon TLR-induced epigenetic activation of PPARG expression [43]. Moreover, lipocalin

2 modulated by PPARG could be a potential pathway involved in concurrent inflammation and ankylosis in inflammatory bowel diseases and ankylosing spondylitis [44]. MDM2 is a multi-functional protein that is involved in both the p53 and NF- $\kappa$ B signaling pathways [45]. DNA induction of MDM2 promotes proliferation of human renal mesangial cells and alters peripheral B cell subsets in pediatric systemic lupus erythematosus [46]. Another study found a potential association between the del1518 variants in MDM2 and rheumatoid arthritis and indicates that combinatorial genotypes and haplotypes in the *MDM2* locus may be linked to rheumatoid arthritis [47]. These hub genes thus regulate immune-related diseases and have the potential to serve as diagnostic and therapeutic targets in these diseases.

Although potential DEGs in AS were identified based on bioinformatics, there were still some limitations to this study. First, the number of samples in this study was limited. Second, further in vitro and in vivo experiments to validate the DEGs and their potential mechanisms are lacking. Therefore, further research is warranted to address the possible limitations of this study in terms of biased results and conclusions.



**Fig. 5** Visualization and details of the ROC curve

## Conclusion

Using high-throughput RNA sequencing, we analyzed the DEGs profile in NC and AS groups in serum tissue. Many of these DEGs were enriched in several signaling pathways associated with ERS, which can provide novel clues for understanding the mechanism driving AS. In addition, we found that PPARG and MDM2 can be used as novel potential molecular targets for the diagnosis and treatment of AS. In conclusion, our results may provide a theoretical basis for further studies to elucidate the molecular mechanism of AS and provide more therapeutic targets for future clinical interventions.

## Acknowledgements

The authors would like to thank Guangzhou Genesee Biotech Co., Ltd. (Guangzhou, China) for their assistance during high-throughput sequencing and data analysis, and also thank LetPub ([www.letpub.com](http://www.letpub.com)) for its linguistic assistance during the preparation of this manuscript. All authors read and approved the final manuscript.

## Author contributions

YB and JK conceived and designed the experiments. All authors performed the experiments. JK YB and XZ analyzed the data. YB and JK were significant contributors to the manuscript. All authors revised the manuscript and have agreed to the publication of this manuscript. All authors read and approved the final manuscript.

## Funding

This research was supported by the Applied Research Project of Qingdao Postdoctoral Researchers (No. RZ2100001531) and the Clinical Medicine +X Project of the Affiliated Hospital of Qingdao University (No. QDFY+X2021037).

## Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository, under accession number GSE205812 are publicly accessible at <http://www.ncbi.nlm.nih.gov/geo>.

## Declarations

### Ethics approval and consent to participate

This research was approved and reviewed by the Medical Ethics Review Committee of the Affiliated Hospital of Qingdao University (approval number: QYFY WZLL 27251). Prior written informed consent was obtained from all the patients to participate in the study. All methods were carried out in accordance with relevant guidelines and regulations.

### Consent for publication

Not applicable.

### Competing interests

The authors have no conflicts of interest to declare.

### Author details

<sup>1</sup>Department of Spinal Surgery, The Affiliated Hospital of Qingdao University, Qingdao 266000, Shandong, China. <sup>2</sup>Department of Rheumatology, The Affiliated Hospital of Qingdao University, Qingdao 266000, Shandong, China.

Received: 12 February 2023 Accepted: 21 May 2023

Published online: 30 May 2023

## References

- Dean LE, Jones GT, Macdonald AG, et al. Global prevalence of ankylosing spondylitis. *Rheumatology*. 2013;53(4):650–7. <https://doi.org/10.1093/rheumatology/ket387>.

- Rudwaleit M, van der Heijde D, Landewe R, et al. The development of assessment of spondyloarthritis international society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis*. 2009;68(6):777–83. <https://doi.org/10.1136/ard.2009.108233>.
- Zhu W, He X, Cheng K, et al. Ankylosing spondylitis: etiology, pathogenesis, and treatments. *Bone Res*. 2019;7(1):22. <https://doi.org/10.1038/s41413-019-0057-8>.
- Stolwijk C, van Tubergen A, Castillo-Ortiz JD, et al. Prevalence of extra-articular manifestations in patients with ankylosing spondylitis: a systematic review and meta-analysis. *Ann Rheum Dis*. 2013;74(1):65–73. <https://doi.org/10.1136/annrheumdis-2013-203582>.
- Peluso R, Minno MNDD, Iervolino S, et al. Enteropathic spondyloarthritis: from diagnosis to treatment. *Clin Dev Immunol*. 2013;2013:1–12. <https://doi.org/10.1155/2013/631408>.
- Dean LE, Macfarlane GJ, Jones GT. Five potentially modifiable factors predict poor quality of life in ankylosing spondylitis: results from the scotland registry for ankylosing spondylitis. *J Rheumatol*. 2017;45(1):62–9. <https://doi.org/10.3899/jrheum.160411>.
- Ghasemi-Rad M, Attaya H, Llesha E, et al. Ankylosing spondylitis: a state of the art factual backbone. *World J Radiol*. 2015;7(9):236. <https://doi.org/10.4329/wjrv.7.9.236>.
- Ermann J. Pathogenesis of axial spondyloarthritis—sources and current state of knowledge. *Rheum Dis Clin N Am*. 2020;46(2):193–206. <https://doi.org/10.1016/j.rdc.2020.01.016>.
- Zeng L, Lindstrom MJ, Smith JA. Ankylosing spondylitis macrophage production of higher levels of interleukin-23 in response to lipopolysaccharide without induction of a significant unfolded protein response. *Arthritis Rheum*. 2011;63(12):3807–17. <https://doi.org/10.1002/art.30593>.
- Rashid T, Ebringer A. Ankylosing spondylitis is linked to klebsiella—the evidence. *Clin Rheumatol*. 2006;26(6):858–64. <https://doi.org/10.1007/s10067-006-0488-7>.
- Sveaas S, Berg I, Provan S, et al. Circulating levels of inflammatory cytokines and cytokine receptors in patients with ankylosing spondylitis: a cross-sectional comparative study. *Scand J Rheumatol*. 2015;44(2):118–24. <https://doi.org/10.3109/03009742.2014.956142>.
- Tan M, Zhang Q, Liu T, et al. Autophagy dysfunction may be involved in the pathogenesis of ankylosing spondylitis. *Exp Ther Med*. 2020;20:3578–86. <https://doi.org/10.3892/etm.2020.9116>.
- Li Z, Sun Y, He M, et al. Differentially-expressed MRNAs, microRNAs and long noncoding RNAs in intervertebral disc degeneration identified by RNA-sequencing. *Bioengineered*. 2021;12(1):1026–39. <https://doi.org/10.1080/21655979.2021.1899533>.
- Wang X, Liotta L. Clinical bioinformatics: a new emerging science. *J Clin Bioinform*. 2011;1(1):1. <https://doi.org/10.1186/2043-9113-1-1>.
- Fattahi F, Kiani J, Khosravi M, et al. Enrichment of up-regulated and down-regulated gene clusters using gene ontology, MiRNAs and LncRNAs in colorectal cancer. *Comb Chem High Throughput Screen*. 2019;22(8):534–45. <https://doi.org/10.2174/1386207321666191010114149>.
- Xu Z, Wang X, Zheng Y. Screening for key genes and transcription factors in ankylosing spondylitis by RNA-Seq. *Exp Ther Med*. 2017. <https://doi.org/10.3892/etm.2017.5556>.
- Li C, Qu W, Yang X. Comprehensive LncRNA and mRNA profiles in peripheral blood mononuclear cells derived from ankylosing spondylitis patients by rna-sequencing analysis. *Medicine*. 2022;101(4):e27477. <https://doi.org/10.1097/md.00000000000027477>.
- Zak DE, Penn-Nicholson A, Scriba TJ, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet*. 2016;387(10035):2312–22. [https://doi.org/10.1016/s0140-6736\(15\)01316-1](https://doi.org/10.1016/s0140-6736(15)01316-1).
- Linden SVD, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. *Arthritis Rheum*. 1984;27(4):361–8. <https://doi.org/10.1002/art.1780270401>.
- Lu W, Lu Y, Wang C, et al. Expression profiling and bioinformatics analysis of exosomal long noncoding RNAs in patients with myasthenia gravis by RNA sequencing. *J Clin Lab Anal*. 2021;35(4):e23722. <https://doi.org/10.1002/jcla.23722>.
- Everaert C, Helmsmoortel H, Decock A, et al. Performance assessment of total RNA sequencing of human biofluids and extracellular vesicles. *Sci Rep*. 2019;9(1):1–16. <https://doi.org/10.1038/s41598-019-53892-x>.

22. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods*. 2012;9(4):357–9. <https://doi.org/10.1038/nmeth.1923>.
23. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinform*. 2011;12(1):1–16. <https://doi.org/10.1186/1471-2105-12-323>.
24. Robinson MD, McCarthy DJ, Smyth GK. EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2009;26(1):139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
25. Kanehisa M. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30. <https://doi.org/10.1093/nar/28.1.27>.
26. Yu G, Wang L-G, Han Y, et al. ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS J Integr Biol*. 2012;16(5):284–7. <https://doi.org/10.1089/omi.2011.0118>.
27. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102(43):15545–50. <https://doi.org/10.1073/pnas.0506580102>.
28. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2020;49(D1):D605–12. <https://doi.org/10.1093/nar/gkaa1074>.
29. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504. <https://doi.org/10.1101/gr.1239303>.
30. Chin C-H, Chen S-H, Wu H-H, et al. CytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8(S4):1–7. <https://doi.org/10.1186/1752-0509-8-s4-s11>.
31. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007; 39(11): 1329–1337. <https://doi.org/10.1038/ng.2007.17>
32. Thomas GP, Duan R, Pettit AR, et al. Expression profiling in spondyloarthritis synovial biopsies highlights changes in expression of inflammatory genes in conjunction with tissue remodelling genes. *BMC Musculoskelet Disord*. 2013;14(1):1–9. <https://doi.org/10.1186/1471-2474-14-354>.
33. Pimentel-Santos FM, Ligeiro D, Matos M, et al. Whole blood transcriptional profiling in ankylosing spondylitis identifies novel candidate genes that might contribute to the inflammatory and tissue-destructive disease aspects. *Arthritis Res Ther*. 2011;13(2):1–8. <https://doi.org/10.1186/ar3309>.
34. Huang D, Liu J, Cao Y, et al. RNA sequencing for gene expression profiles in peripheral blood mononuclear cells with ankylosing spondylitis RNA. *Biomed Res Int*. 2020;2020:1–13. <https://doi.org/10.1155/2020/5304578>.
35. Antoniou AN, Lenart I, Guiliano DB. Pathogenicity of misfolded and dimeric HLA-B27 molecules. *Int J Rheumatol*. 2011;2011:1–9. <https://doi.org/10.1155/2011/486856>.
36. Layh-Schmitt G, Colbert RA. The interleukin-23/interleukin-17 axis in spondyloarthritis. *Curr Opin Rheumatol*. 2008;20(4):392–7. <https://doi.org/10.1097/bor.0b013e328303204b>.
37. Ryno LM, Wiseman RL, Kelly JW. Targeting unfolded protein response signaling pathways to ameliorate protein misfolding diseases. *Curr Opin Chem Biol*. 2013;17(3):346–52. <https://doi.org/10.1016/j.cbpa.2013.04.009>.
38. Rezaeiemanesh A, Mahmoudi M, Amirzargar AA, et al. Upregulation of unfolded protein response and ER stress-related IL-23 production in M1 macrophages from ankylosing spondylitis patients. *Inflammation*. 2022;45:665–76. <https://doi.org/10.1007/s10753-021-01575-z>.
39. Liu C, Chou C, Chen C, et al. Aberrant distribution and function of plasmacytoid dendritic cells in patients with ankylosing spondylitis are associated with unfolded protein response. *Kaohsiung J Med Sci*. 2020;36(6):441–9. <https://doi.org/10.1002/kjm2.12184>.
40. Liu B-H. Differential coexpression network analysis for gene expression data. New York: Springer; 2018.
41. Liu Y, Wang J, Luo S, et al. The roles of PPAR $\gamma$  and its agonists in autoimmune diseases: a comprehensive review. *J Autoimmun*. 2020;113:102510. <https://doi.org/10.1016/j.jaut.2020.102510>.
42. Fougerat A, Montagner A, Loiseau N, et al. Peroxisome proliferator-activated receptors and their novel ligands as candidates for the treatment of non-alcoholic fatty liver disease. *Cells*. 2020;9(7):1638. <https://doi.org/10.3390/cells9071638>.
43. Liu Y, Luo S, Zhan Y, et al. Increased expression of PPAR- $\gamma$  modulates monocytes into a M2-like phenotype in SLE patients: an implicative protective mechanism and potential therapeutic strategy of systemic lupus erythematosus. *Front Immunol*. 2021;11:579372. <https://doi.org/10.3389/fimmu.2020.579372>.
44. Lin A, Inman RD, Streutker CJ, et al. Lipocalin 2 links inflammation and ankylosis in the clinical overlap of inflammatory bowel disease (IBD) and ankylosing spondylitis (AS). *Arthr Res Ther*. 2020;22(1):1–11. <https://doi.org/10.1186/s13075-020-02149-4>.
45. Thomasova D, Mulay SR, Bruns H, et al. P53-independent roles of MDM2 in NF- $\kappa$ B signaling: implications for cancer therapy, wound healing, and autoimmune diseases. *Neoplasia*. 2012;14(12):1097–101. <https://doi.org/10.1593/neo.121534>.
46. Zhang C, Chen J, Cai L, et al. DNA induction of MDM2 promotes proliferation of human renal mesangial cells and alters peripheral B cells subsets in pediatric systemic lupus erythematosus. *Mol Immunol*. 2018;94:166–75.
47. Gansmo LB, Lie BA, Mæhlen MT, et al. Polymorphisms in the TP53-MDM2-MDM4-axis in patients with rheumatoid arthritis. *Gene*. 2021;793:145747.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

