Identification of differentially expressed genes, signaling pathways and immune infiltration in rheumatoid arthritis by integrated bioinformatics analysis

Yanzhi Ge¹, Li Zhou¹, Zuxiang Chen¹, Yingying Mao², Ting Li³, Peijian Tong¹* and Letian Shan¹*

Abstract

Background: The disability rate associated with rheumatoid arthritis (RA) ranks high among inflammatory joint diseases. However, the cause and potential molecular events are as yet not clear. Here, we aimed to identify differentially expressed genes (DEGs), pathways and immune infiltration involved in RA utilizing integrated bioinformatics analysis and investigating potential molecular mechanisms.

Materials and methods: The expression profiles of GSE55235, GSE55457, GSE55584 and GSE77298 were downloaded from the Gene Expression Omnibus database, which contained 76 synovial membrane samples, including 49 RA samples and 27 normal controls. The microarray datasets were consolidated and DEGs were acquired and further analyzed by bioinformatics techniques. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of DEGs were performed using R (version 3.6.1) software, respectively. The protein-protein interaction (PPI) network of DEGs were developed utilizing the STRING database. Finally, the CIBERSORT was used to evaluate the infiltration of immune cells in RA.

Results: A total of 828 DEGs were recognized, with 758 up-regulated and 70 down-regulated. GO and KEGG pathway analyses demonstrated that these DEGs focused primarily on cytokine receptor activity and relevant signaling pathways. The 30 most firmly related genes among DEGs were identified from the PPI network. The principal component analysis showed that there was a significant difference between the two tissues in infiltration immune.

Conclusion: This study shows that screening for DEGs, pathways and immune infiltration utilizing integrated bioinformatics analyses could aid in the comprehension of the molecular mechanisms involved in RA development. Besides, our study provides valuable data related to DEGs, pathways and immune infiltration of RA and may provide new insight into the understanding of molecular mechanisms.

Keywords: Rheumatoid arthritis, Bioinformatics analysis, Differentially expressed genes, Immune infiltration

* Correspondence: tongpeijian@163.com; letian.shan@zcmu.edu.cn
1The First Affiliated Hospital, Zhejiang Chinese Medical University, Hangzhou 310053 Zhejiang, PR China
Full list of author information is available at the end of the article

© The Author(s). 2021 Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Introduction

Rheumatoid arthritis (RA) occurs in approximately 5 per 1000 people and can inevitably prompt severe joint damage and disability. Significant progress has been made over the past two decades with respect to the disease pathophysiology, optimal outcome measures, and effective treatment strategies, including the understanding of the comprehension in diagnosing and treating RA in the early stage [1]. The disability rate of RA ranks high among the arthritic which occurs in multiple-joint on the human body, and the incidence of this kind of arthritis is increasing year by year. The incidence of RA is occult, early diagnosis is difficult, and imaging manifestations occur comparatively late. At the point when RA is identified, the patients are usually at an advanced stage of this disease. RA would lead to multiple-joint dysfunction, disability, lower quality of life, respiratory illness, cardiovascular disease, and other comorbidities in patients not receiving intervention [2]. The etiology of RA is still ambiguous. All things considered, both genetic factors and environmental factors, contribute to the occurrence and development of RA [3].

At present, the frequently used methods for early detection and diagnosis of RA are magnetic resonance imaging, ultrasound and serological examination (including rheumatoid factor, anti-cyclic citrullinated peptide, etc). Be that as it may, there are confinements of these techniques, so the exactness and accuracy is not high [4, 5]. Computed tomography and X-rays can only detect lesions in its advanced stage, but do not detect early impairment. At the moment, the treatment of RA incorporates drug treatment, immunologic purging, functional training, surgical operation, and complementary and/or alternative medicine, and so forth [6, 7]. Therefore, it is critical to study the potential molecular mechanisms of RA synovial membrane and consequently identify more valid diagnostic techniques and more reliable molecular markers for detecting occurrence and evaluating prognosis, as well as to investigate more valid methods to control and prevent RA. Gene expression microarrays have been generally applied in studying gene expression profiles which provides a moderately new way for exploring genes and offers broad application prospects for drug-based molecular targeting. At present, large amounts of data have been published on Gene Expression Omnibus (GEO) [8]. Furthermore, integrating these databases can permit a more profound study of molecular mechanisms.

In this study, we downloaded four original microarray datasets (including GSE55235, GSE55457, GSE55584 and GSE77298) from the GEO database which incorporated a total of 76 samples, with 27 healthy controls and 49 RA samples. Differentially expressed genes (DEGs) in RA samples and control group (CG) were screened utilizing packages in R (version 3.6.1) software, and gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analysis of DEGs were additionally performed. After that, the protein-protein interaction (PPI) network was used to analyze the anticipated associations for a particular group of proteins through the STRING online database. Finally, the immune infiltration was analyzed by performing the CIBERSORT algorithm between RA and normal tissues, which was widely used to assess the relationship and the relative content of 22 types of immune cell subsets. Because of the limited samples and selection differences, not all the DEGs, relevant pathways and immune infiltration being screened out could be directly used as biomarkers.

Materials and methods

Search strategy

We used the keyword “Rheumatoid arthritis” to search the GEO database (https://www.ncbi.nlm.nih.gov/geo/), and there was a sum of 5543 results for “Rheumatoid arthritis” in the GEO database from their inception up to November 7, 2019. By restricting the entry type (series), study type (expression profiling by array) and tissue sources (Homo sapiens), 5386 pieces of items that were not related to the purpose of this study were excluded. After further selection with title, summary and samples, we discovered the absence of required data in 153 items. Finally, 4 series from 2 platforms were included, and gene expression profiles of GSE55235, GSE55457, GSE55584 and GSE77298 were downloaded. Figure 1 depicted the details of the selection process.

Microarray data information

GSE55235, GSE55457 and GSE55584 were three multi-center genome-wide transcriptomic data sets (Affymetrix HG- U133 A) from a total of 79 individuals, including 20 healthy controls, as well as 26 osteoarthritis patients and 33 RA patients. The platform for GSE77298 was GPL570, [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array, which included 7 synovial tissue from healthy joints and 16 synovial tissue from rheumatoid arthritis joints. Platform and series matrix file(s) were downloaded from the GEO and saved as TXT files. R software (version 3.6.1) was used to process the downloaded files.

Integration of microarray data and DEGs

These four raw datasets were incorporated for the analysis. The coordinated microarray datasets were batch-normalized by R software (https://bioconductor.org/biocLite.R) utilizing limma packet (biocLite “limma”) analysis and then saved as a TXT file. The operating instruction codes, utilizing the R software, were processed.
automatically and the DEGs in CG and RA samples were analyzed by the limma package. Using R software to run the instruction code. Up or down-regulated genes were obtained independently and utilized for further analysis. The downloaded files (including platform and series of the matrix) were converted annotation package utilizing the R software.

**Data processing and identification of DEGs**

Volcano plot was used to display all of the up-regulated and down-regulated DEGs using the same limma package. Then, using pheatmap package and restricting the items to 100 DEGs, the hierarchical clustering was performed. The results were visualized using a heatmap. The ID associated with the probe name was converted into gene symbol using Perl programming (https://www.perl.org) language (version 5.30.0) as well as a genome-wide annotation library (http://www.bioconductor.org/packages/release/data/annotation/html/org.Hs.eg.db.html) and then saved in a TXT file. Adjusted \( P \)-value<0.05 and log fold change (logFC)>2 were considered as DEGs.

**GO and KEGG pathway enrichment analyses of DEGs**

The functional and pathway enrichment of the proteins encoded by candidate genes were analyzed, and these genes were annotated using the R software. GO and KEGG pathway analysis of DEGs were performed utilizing the appropriate packages (biocLite “DOSE”, “clusterProfiler” and “pathview”). In this study, we analyzed the DEGs that were significantly up and down-regulated as determined from integrated microarray RA data, and an adjusted \( P \)-value of<0.05 was considered statistically significant.

**PPI network integration**

The database STRING (https://string-db.org, version 11) was a precomputed worldwide resource for the exploration and analysis of interactions between known and predicted protein-protein interaction. Concerning a specific group of proteins, the network view analyzed the predicted associations. Each network node represented a different protein and the association between these nodes represented the interaction of biological molecules, which can be used for identifying interactions and associated pathways between these proteins encoded by DEGs in RA. The central nodes which were closely related to other corresponding proteins may be the core or key proteins and exert significant physiological functions.

**Immune infiltration by CIBERSORT analysis**

We used the CIBERSORT [9] to analyze the normalized data filtered by Perl programming language, and immune cells infiltration matrix were acquired. In this study, twenty-two immune cells included macrophages M2, plasma cells, neutrophils, mast cells activated, T cells CD8, macrophages M1, T cells gamma delta, B cells memory, monocytes, B cells naive, T cells follicular helper, NK cells activated, dendritic cells resting, T cells CD4 memory activated, T cells CD4 naive, NK cells...
resting, T cells regulatory (Tregs), dendritic cells activated, eosinophils, macrophages M0, T cells CD4 memory resting and mast cells resting. The percentage of immune cells in the gene expression matrix and the relationship between two immune cells by installing "corrplot" package was calculated [10]. After that, using "ggplot2" package, the principal component analysis (PCA) was performed to discover whether there was a difference between RA and normal synovial tissues.

**Results**

An aggregate of 828 DEGs was acquired, of which 758 were up-regulated and 70 down-regulated respectively (Fig. 2). Top of the 50 up and down-DEGs from the integrated data are shown in Table 1 separately. The smaller its adjusted P-value, the greater possibility of DEG and higher ranking in this experiment. R-heatmap software was utilized to draw a heatmap of the 50 up and 50 down-regulated DEGs, as shown in Fig. 3.

**GO term and KEGG pathway enrichment analysis of DEGs**

The GO term and KEGG pathway enrichment analyses of up and down-regulated genes with an adjusted P-value of<0.05 were obtained respectively. The results of the GO term in RA were shown in Table 2 and Fig. 4 (a and b). The visual analysis results of the KEGG enrichment of DEGs in RA were shown in Table 3 and Fig. 4 (c and d). The up-regulated genes were mainly enriched in cytokine receptor activity, G protein-coupled chemoattractant receptor activity, chemokine receptor activity, MHC protein complex binding, chemokine binding, chemokine receptor binding and cytokine activity. The down-regulated genes were mostly amassed in peroxidase activity and oxidoreductase activity, acting on peroxide as acceptor. In the KEGG analysis, the up-regulated genes were mainly enriched in the chemokine signaling pathway, hematopoietic cell lineage, cytokine-cytokine receptor interaction, viral protein interaction with cytokine and cytokine receptor, primary immunodeficiency, leishmaniasis, osteoclast differentiation, rheumatoid arthritis, cell adhesion molecules (CAMs). The down-regulated enriched KEGG pathways of DEGs included PPAR signaling pathway, regulation of lipolysis in adipocytes, adipocytokine signaling pathway, glucagon signaling pathway, AMPK signaling pathway, calcium signaling pathway, thyroid hormone synthesis, apelin signaling, cGMP-PKG signaling pathway. Besides, the pathway map for targeted RA (Fig. 5) was described using in KEGG pathway enrichment. The significantly enriched terms and pathways may enlighten our minds and assist us in further study of the role of DEGs in RA.

**Analyzing DEGs in RA using a PPI network**

The DEG expression products in RA were constructed by way of the STRING database to construct PPI networks (minimum required interaction score: 0.990). After deleting all isolated and partially disconnected nodes, an integrated network was built, as shown in Fig. 6a. The 30 most significant genes (Fig. 6b) which had been displaying statistical significant interaction were CDK1, KIF11, CDC20, CCNB1, CCNB2, MAD2L1, BUB1B, NDC80, AURKA, CCNA2, ISG15, NDC80, TTK, DLGAP5, LCP2, TPX2, CD247, CKS2, LCK, VAV1, CCL5, CD3E, FOXL1, KIF20A, MX1, NUSAP1, SYK, ZAP70, ZWINT and ASPM. Among these genes, CDK1, KIF11, and CDC20 possessed the highest node degree.

**Immune infiltration analyses**

PCA depicted no overlap of these two elliptical clusters and showed that there was a significant difference (P<0.05) in immune cell infiltration between the RA group and the healthy control group (Fig. 7a). The corheatmap (Fig. 7b) result showed that mast cells activated and eosinophils had a positive correlation (value = 0.68). T cells CD8 had a significant negative correlation with T cells CD4 memory resting (value = -0.64). Correlation heatmap (Fig. 7c) summarized the results obtained from 69 filtered gene expression matrix and the relative percent of the 22 immune cells was shown in Fig. 8a. Compared with normal tissue, the violin plot of the immune cell showed that, plasma cells, T cells CD8, T cells CD4 memory activated, T cells follicular helper, T cells
The characteristics of RA is synovitis, systemic inflammation, and the arrival of autoantibodies [2]. As a result, the synovial membrane breaks down the body’s immune system, causing chronic inflammation, destruction of cartilage and bone, and dysfunction to other essential organs [11–13]. It is reported that 50% of the risk for occurrence and development of RA is related to genetic factors. At the same time, smoking is an environmental risk factor for RA. The early onset of RA is not easy to identify, and in the meantime, cartilage and bone disintegration are frequently found in the end stages of this disease. RA occurrence and development can occur at any age, gender, and nationality for complex biological processes, and the positive rate of serum examination is low as well as non-specific. Consequently, it is miles critical to observe and study the mechanisms and development of RA at the molecular level. Based on this, differentially expressed genes (DEGs) have been efficaciously used to predict the response of therapeutic approaches for RA patients. As an example, the capability of certain genes (type I interferon-responsive) to predict the nonresponder of rituximab [14] and anti-tumor necrosis factor [15].

In this study, we integrated gene expression profile datasets from four specific groups (GSE55235, GSE55457, GSE55584 and GSE77298) and used R (version 3.6.1) to analyze these datasets. A total of 828 DEGs were identified using the limma package, consisting of 758 up-regulated genes and 70 down-regulated genes. The pinnacle 20 most significantly up-regulated genes were ADAMDEC1, IGHIM, IGJ, IGKC, IGLL3, IGLV1–44, IGLC1, TNFRSF17, IGLL5, CRTAM, CXCL9, IGLJ3, TRAT1, SDC1, TPD52, IGK, CD27, CXCL10, IL21R and IGHG1. The pinnacle 20 most significantly down-regulated genes were SLCL19A2, PLIN1, KL9F, ADCY2, PPAP2B, EB2F, ADH1B, KLF4, PPARGC1A, GABARAP1, TRHDE, PHKA1, FBXW12, TCEAL2, PCK1, PCDH9, MAFF, LEP, RERGL and SGCA. Constructing a PPI network (minimum required interaction score = 0.990) of DEG-encoding proteins from the STRING database and screening the 30 most significant related genes. The enriched GO of DEGs were analyzed by R software, and correlation analysis confirmed that the up-regulated genes have been specifically involved in cytokine receptor activity, G protein-coupled chemoattractant receptor activity, chemokine receptor activity, MHC protein complex binding, chemokine binding, chemokine receptor binding and cytokine activity, and that the down-regulated DEGs were mainly involved in peroxidase activity and oxidoreductase activity, acting on peroxide as acceptor. This finding is consistent with the knowledge that cytokine, chemokine and peroxidase activity play crucial roles in the RA occurrence and progression.

The detection of autoantibodies (including RA, A-CCP, CRP) in RA patients is an identification that distinguishes the disease from other inflammatory arthritis, such as psoriatic arthritis, reactive arthritis and osteoarthritis. In addition to the clinical symptoms and signs arising from arthritis processes in the joints, muscle weakness around joints is also commonly reported by RA patients [16–18]. Takashi Yamada et al. [18] discovered that altered Ca2+ and free radical signaling (such as reactive oxygen and reactive nitrogen species) can result in RA-based muscle weakness. In a general way, RA with CCP + RF+ subjects had excessively high citrulline-specific IgG binding, and CCP + RF- and CCP-RF+ subjects had modest binding to array peptides [19]. As a systemic autoimmune disease, RA is characterized by inflammation and angiogenesis in the synovium. Many cytokines and inflammatory medium are observed in synovial tissues and synovial fluids, whose function is to display angiogenic properties. Inhibitor of DNA binding 1, one of the transcription factors, is a marker of cellular self-renewal. This factor within the bone marrow causes a significant reduction of endothelial progenitor cell association with tumor-related vasculogenesis [20, 21]. Amélie Simon et al. [22] observed that microscopic polyangiitis is vasculitides typical of necrotizing inflammation for small-sized vessels and is usually connected with serum positivity for those anti-neutrophil cytoplasmic

**Table 1**

| DEGs                                      | Gene symbol                                      |
|------------------------------------------|--------------------------------------------------|
| Up-regulated                             | ADAMDEC1, IGHIM, IGJ, IGKC, IGLL3, IGLV1–44, IGLC1, TNFRSF17, IGLL5, CRTAM, CXCL9, IGLJ3, TRAT1, SDC1, TPD52, IGK, CD27, CXCL10, IL21R, and IGHG1. |
| Down-regulated                           | SLC19A2, PLIN1, KL9F, ADCY2, PPAP2B, EB2F, ADH1B, KLF4, PPARGC1A, GABARAP1, TRHDE, PHKA1, FBXW12, TCEAL2, PCK1, PCDH9, MAFF, LEP, RERGL, and SGCA. |

**DEGs Differentially expressed genes, RA Rheumatoid arthritis**
Fig. 3 Heatmap of the top 100 DEGs according to the adjusted \( P \)-value and logFC. Red indicates higher gene expression and green indicates lower gene expression.
antibodies. In most conditions, anti-neutrophil cytoplasmic antibodies are directed against two constituents of neutrophil primary granules as well as monocyte lysosomes: myeloperoxidase or proteinase 3.

Furthermore, the up-regulated enriched Kyoto Encyclopedia of Genes and (KEGG) pathways of DEGs included the chemokine signaling pathway, hematopoietic cell lineage, cytokine-cytokine receptor interaction, viral protein interaction with cytokine and cytokine receptor, primary immunodeficiency, leishmaniasis, osteoclast differentiation, rheumatoid arthritis, cell adhesion molecules (CAMs). The down-regulated enriched KEGG pathways of DEGs included PPAR signaling pathway, regulation of lipolysis in adipocytes, adipocytokine signaling pathway, glucagon signaling pathway, AMPK signaling pathway, calcium signaling pathway, thyroid hormone synthesis, apelin signaling, cGMP-PKG signaling pathway. Relative studies have demonstrated that fibroblast-like synoviocytes play a crucial role by producing cytokines in all stages of RA. Once fibroblast-like synoviocytes are activated during RA, a series of inflammatory factors and proteases will be produced involved in the inflammatory response, causing progressive destruction of bone and cartilage [23].

RA is associated with an increase in mortality. Previous research displayed that the occurrence rate of the

Table 2 GO analysis of up-regulated and down-regulated DEGs

| ID        | Description                              | Adjusted P-values | Gene symbol | Count |
|-----------|------------------------------------------|-------------------|-------------|-------|
| Up-regulated |                                          |                   |             |       |
| GO0004696 | cytokine receptor activity               | 8.17E-09          | IL21R/IL7R/IL2RG/CCR5/CCR2/CSF2RB/CCR7/CXCR4/ CXCR6/CXCR3/CD4/CCR1/IL2RA/IL1RA/CX3CR1/ CD74/CCR2/IL1RL1/CCR5/CCR6 | 20    |
| GO0001637 | G protein-coupled chemoattractant receptor activity | 1.30E-07          | CCR5/CCR2/CCR7/CXCR4/CXCR6/CXCR3/CCR7/ CX3CR1/CCR1/CCR5/CCR6 | 11    |
| GO0004950 | chemokine receptor activity              | 1.30E-07          | CCR5/CCR2/CCR7/CXCR4/CXCR6/CXCR3/CCR7/ CX3CR1/CCR1/CCR5/CCR6 | 11    |
| GO0023023 | MHC protein complex binding             | 1.30E-07          | HLA-DOB/MS4A1/CD8A/HLA-DMB/HLA-DMA/HLR1/ LILRB2/LILRB1/HLA-DRA/TAIBP1/CD74 | 11    |
| GO0019956 | chemokine binding                        | 7.20E-07          | ITGA4/CXCR5/CCR2/CCR7/CXCR4/CXCR6/CXCR3/CCR1/ CX3CR1/CCR6 | 10    |
| Down-regulated |                                          |                   |             |       |
| GO0004650 | peroxidase activity                     | 0.044092206       | DUOX2/GPX3/PTGS2 | 3     |
| GO0016684 | oxidoreductase activity, acting on peroxide as acceptor | 0.044092206       | DUOX2/GPX3/PTGS2 | 3     |

GO gene ontology, DEGs differentially expressed genes

Fig. 4 GO and KEGG pathway enrichment analysis of DEGs in GSE55235, GSE55457, GSE55584 and GSE77298. a GO terms in the enrichment analysis of the up-regulated genes. b GO terms in the enrichment analysis of the down-regulated genes. c KEGG terms in the enrichment analysis of the up-regulated genes. d KEGG terms in the enrichment analysis of the down-regulated genes.
malignancies in RA patients has been reported to be high [6]. A review of scientific studies compiled in Romania demonstrated findings that anemia and other chronic disease manifestations are relatively common in approximately 6–10% of RA patients, and are all related to worse outcomes in particular functional impairment and mortality [24, 25]. The adaptive immune system is closely connected with the generation of the anti-tumor immune response. For that reason, RA patients with gastrointestinal cancer history must be carefully monitored while receiving the treatment of disease-modifying antirheumatic drugs [26]. However, in many factors, tumor necrosis factor-α is recognized as performing biological functions associated with the pathogenesis of RA [27]. Its capabilities include chemokine amplification, endothelial cell activation, leukocyte accumulation [28],

| Table 3 KEGG pathway of up-regulated and down-regulated DEGs |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|
| ID                             | Description                                             | Adjusted P-values | Gene symbol                                               | Count |
|--------------------------------|----------------------------------------------------------|-------------------|-----------------------------------------------------------|-------|
| Up-regulated                    |                                                          |                   |                                                           |       |
| hsa04062 Chemokine signaling pathway | 1.90E-14                                                | CXCL9/CXCL10/CCL18/CXCL6/CCL5/CXCR5/STAT1/STAT2/ CCR2/ITGAM/PRGK2/NF1/CCR1/RAF2/CCR1/CXCL5/ CXCL11/CXCR4/CCL13/ADcy7/TNFSF6/CXCR3/DDOQ2/PRKcz/XCL1/JAK2/PRKACB/PIK3CG/CCR1/WAS/AVI1/CXCL1/PIK3CD/HCK/CX3CR1/CCL19/CCL7/CXCR5/ARBB2/GNAI2/STAT2/CXCR6/FRG/PRKCD | 43    |
| hsa04640 Hematopoietic cell lineage | 8.51E-13                                                | HLA-DOB/M541/IILR/ITG4A/CD38/CD2/CD2A/CD3D/ HLA-DMB/CSF1R/CD19/3ADG7/HILA-MDA/CD3E/ HLA-DPB1/CD37/HILA-DQB1/CR1/HILA-DR/CD14/CD4/ IL2RA/CSF3R/FLT3LG/CD18/CD24/CD5G/CD7/CFS2RA | 29    |
| hsa04060 Cytokine-cytokine receptor interaction                  | 4.76E-12                                                | TNFSF17/CXCL9/CD27/CXCL20/IILR/ITG18/TNFSF11/CXCL6/CCL5/L7R/L2RG/CXCR5/IIL52/CX2R2B/CXCL13/ CSF1R/CCR7/IIL7/CXCL5/IIL23A/CXCL11/CXCR4/CCL13/CXCR6/CXCR3/TNFSF10/LTB/XCL1/IIL2RA/TNFSF4/CD4/ CCR1/IIL10RA/CXCL1/IIL2RA/CSF3R/IIL15/CR3XCR1/CFAS/ IIL18/CCL19/IFNA21/CX4GL/GCL7/TNFSF13/IL1R1/CRXRS/ CSF2RA/CXCR6 | 50    |
| hsa04061 Viral protein interaction with cytokine and cytokine receptor | 6.00E-12                                                | CXCL9/CXCL10/CCL18/CXCL6/CCL5/L2RG/CXCR5/CXCR2/ CXCL13/CSF1R/CCR7/CCL5/CXCL11/CXCR4/CCL13/CXCR3/TNFSF10/XCL1/IIL10RA/CXCL1/IIL2RA/CXCR3/XCR1/IIL8/ CCL19/CXCR5/CXCR6 | 28    |
| hsa05340 Primary immunodeficiency                                  | 7.15E-11                                                | CD79A/LCK/BLNK/IILR/L2RG/PTPRC/CD8A/CD3D/CD19/ CD3E/TAP1/PTK5/ZA50/CD4/ICOS/IIL7/CD40LG | 17    |
| hsa05140 Leishmaniasis                                              | 1.16E-10                                                | HLA-DOB/ITG4A/STAT1/HILA-MDB/PRGK2/NF1/ITG23/ADAC7/ HILA-MDA/HLA-DPB1/CR1/IIL15/HLA-MDA/CR2/CD32/ HILA-DMA/CD3E/CD19/CD8A/CD3D/ACD-2/CD2A/CD3D/ CCR7/ITGAM/PRGK2/NF1/ITG23/ADAC7/ | 23    |
| hsa04380 Osteoclast differentiation                                | 4.24E-10                                                | TNFSF11/AK/LCK/BLNK/STAT1/NF1/CSF1R/CRG2/CTG2/3ADG7/ CYBA/ULRB6/ULRB1/CD1B3/ULRB3/FCGR3B/IRP5/ULRB3/CD1C/ LTB/XCL1/IFNA21/CD4GL/GCL7/TNFSF13/IL1R1/CRXRS/CSF2RA/CCR6 | 29    |
| hsa05323 Rheumatoid arthritis                                      | 1.27E-09                                                | MMP1/HILA-DDB/TNFSF11/CXCL6/CCL5/HILA-MDB/MP3/TG2B/ ITGAL/3ADG7/IIL5/HILA-MDA/CXCL5/IIL23A/HILA-DPB1/CD86/ HILA-DQB1/LTB/HILA-DR/CXCL11/CTL4A/IIL18/TNFSF13/ACPS | 24    |
| Down-regulated                                                           |                                                          |                   |                                                           |       |
| hsa03320 PPAR signaling pathway                                     | 0.000189284                                             | PLN1/PCK1/ADIPQ/ACADL/FABP4/LPL | 6    |
| hsa04923 Regulation of lipolysis in adipocytes                      | 0.000398572                                             | PLN1/ADCY2/NPY1R/FABP4/PTG5S | 5    |
| hsa04920 Adipocytokine signaling pathway                            | 0.000818815                                             | ADCY2/ADCY21/PEK1/LEP/ADIPQ/ACABC | 5    |
| hsa04922 Glucagon signaling pathway                                 | 0.004887657                                            | ADCY2/ADCY21/PHK1/ADCY21/PEK1/ACABC | 5    |
| hsa04152 AMPK signaling pathway                                      | 0.006451484                                            | ADCY2/ADCY21/PEK1/ADIPQ/ACABC | 5    |
| hsa04020 Calcium signaling pathway                                  | 0.006451484                                            | ADCY2/PEK1/EDNRB/PLN/NNC2/AGTR1 | 6    |
| hsa04918 Thyroid hormone synthesis                                 | 0.008113772                                            | ADCY2/ATP1A2/DUOX/GXG3 | 4    |
| hsa04371 Apelin signaling pathway                                   | 0.008113772                                            | PLN1/ADCY2/ADCY21/23710AGTR1 | 5    |
| hsa04022 cGMP-PKG signaling pathway                                 | 0.016012172                                            | ADCY2/EDNRB/ATP1A2/PLN/AGTR1 | 5    |
| hsa04925 Aldosterone synthesis and secretion                        | 0.016012172                                            | ADCY2/ADIPX2/KNK3/AGTR1 | 4    |

KEGG Kyoto Encyclopedia of Genes and Genomes, DEGs differentially expressed genes
Fig. 5 KEGG pathway enrichment analysis and pathway map for RA

Fig. 6 a PPI network (828 DEGs filtered into the PPI network that contained 103 nodes and 168 edges). b The predicted association rank (from low to high) of the top 30 genes in the PPI network.
experiencing cardiovascular comorbidity [29], acceleration destruction of osteoclast and chondrocyte, and demonstrating metabolic syndrome [30]. Related studies have reported that peroxisome proliferator-activated receptor (PPAR)-γ may additionally induce activation Wnt/β-catenin signaling [31]. Numerous studies have indicated that decreased expression of adipocyte genes such as nuclear receptors PPARg in the RA synovial tissue [32, 33], and PPARg mediates mesenchymal stem cells as well as fibroblast-like synovial cells differentiation into adipocytes [34]. PPAR-γ activators performed significantly anti-inflammatory and anti-degeneration roles in rheumatoid arthritis [35]. According to a study of adjuvant-induced arthritis in rats’ synovium, inhibition PPAR-γ expression by T0070907 or PPAR-γ siRNA could significantly promote the proliferation of fibroblast-like synoviocytes and expressions of c-Myc, Cyclin D1, MMP-1, and MMP-9, except for TIPM-1 [31, 36]. Meanwhile, compared with normal tissues, PPAR-γ was obviously reduced whether immunohistochemical technique or protein detection by western blot [36]. A lot of experimental results were commendably consistent with the result in this integrated study and suggested that PPAR-γ might play a pivotal role during RA synovial tissue activation. As for the gene expression of AMPK in those newly diagnosed RA patients, a master regulator of the metabolic process was decreased in the peripheral blood leukocytes and elevated levels of TGF-β1 in plasma accounts for the occurrence of RA pathogenesis [37]. Recent data evidence suggested that S100A8/A9 was a member of the Ca2+ binding S100 protein family and had become a hot topic as a critical alarmin modulating the inflammatory response. Using small-molecule inhibitors that block off S100A8/A9 activity can exhibit beneficial functions on disease relative activities in animal models of autoimmune diseases such as RA [38, 39].

We constructed a PPI network of the protein encoded by DEGs and identified the subsequent top 10 closely related genes: CDK1, KIF11, CDC20, CCNB1, CCNB2, MAD2L1, BUB1B, NDC80, AURKA and CCNA2. These genes are key nodes for the construction a PPI network and play a distinct role in the pathogenesis of RA. In accordance with the proinflammatory CDK signaling, p16INK4A protein as a Cyclin-Dependent Kinases inhibitor in synovial fibroblasts also demonstrated an
inhibitory action in the development of RA [40]. Ectopic expression of $\text{p16}^{\text{INK4A}}$ protein can also suppress LPS-induced IL-6 expression in macrophages [41], and simultaneously enhance the observations that CDK inhibitory proteins relative features to counteract inflammation [42]. Interleukin-6 (IL-6) signaling was a critical target in inflammatory pathways [43]. In patients with RA, the high level of IL-6 and IL-6R are found in both serum and synovial fluid of related joints affected by the disease. IL-6 is a cytokine serving several biological and biochemical functions that affect the immune and vasculature system. Generally speaking, conventional IL-6 signaling is in charge of the anti-inflammatory capabilities of IL-6, conversely, trans-signaling is in charge of the pro-inflammatory properties of IL-6. Consequently, disorders of the IL-6 axis can result in the onset or progression of disease states, especially in autoimmune and inflammatory dysregulation [44]. Activation of epidermal growth factor receptor (EGFR) signaling leads to the propagation and metabolism of synovial fibroblasts in RA. Beyond that, in addition to its function in propagation and metabolism, EGFR can generate cytokine in synovial tissues during the pathogenesis of RA. Some animal experiments have yielded potentially prospective results aiming at target EGFR involving RA. As a result, pharmacologic modulations or its ligands targeting EGFR may reveal undiscovered methods for the treatment of RA [45]. EGFR receptor is a tyrosine kinase. At present, only NEK6 and CDK1 kinases can phosphorylate KIF11 at Ser1033 and Thr926 respectively, causing the combination of microtubules and KIF11 in the process of mitotic spindle assembly [46–48]. Some inflammatory cytokines are controlled by the expression of the c-Fos. Both IL-1β and c-Fos are interacted with each other, including its gene expression and activities, and causing a cross-linking effect that is a vital mechanism to arthritic joint destruction. As a result, the blockade of IL-1β, c-Fos, or link between both can be effective
therapeutically as a treatment method for RA patients' joint destruction [49]. Researchers at the University of Chicago found through mice experiments that the inhibition of c-Myc or c-Raf-1 can significantly decrease the invasiveness of RA synovial fibroblasts. Besides, dominant-negative mutants c-Raf-1 reduced the expression of phosphorylated c-Jun in vivo as well as the expression of disease-relevant MMPs [50].

In the current study, using CIBERSORT analytical tool, the relative percent and content of 22 immune cell subsets, what we have performed, were the most comprehensive analysis in RA synovial tissues to date. We found that the proportion of mast cells activated, eosinophils and T cells CD4 memory resting was high in the normal synovial tissue, while T cells CD8 accounted for RA synovial high expression. A study has found that T cells, particularly CD8 infiltration in the synovial tissues, were considered as predictors of RA development and the presence of antibodies against citrullinated peptides [51]. In the correlation analysis of the infiltration degree of immune cells, CD8+ T cells were negatively correlated with the infiltration degree of resting CD4+ memory T cells, which meant the binding of CD4 + T cells and RA synovial antigen in the microenvironment was decreased [52].

In the last few years, DEGs identification and further bioinformatics analysis was performed and these results may provide new perspective for the study of RA. However, a limitation of this study was not experimentally validated, which may need to perform in the future studies.

Conclusions
It is beneficial for the research community to study this network to further examine and understand the interaction between RA relevant DEGs. These findings may help us to enhance our general understanding of the molecular mechanism of RA. However, further relevant molecular biological experiments are required to affirm the function of the identified genes associated with RA.

Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: National Natural Science Foundation of China (Grant No. 81774331 and 82074464). The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Availability of data and materials
The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests.

Author details
1The First Affiliated Hospital, Zhejiang Chinese Medical University, Hangzhou 310053 Zhejiang, PR China. 2Department of Epidemiology and Biostatistics, Zhejiang Chinese Medical University, Hangzhou 310053 Zhejiang, PR China. 3The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003 Zhejiang, PR China.

Received: 6 July 2020 Accepted: 8 December 2020

Published online: 04 January 2021

References
1. Myasoedova E, Crowson CS, Kremer HM, Themeau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising?: results from olmsted county, Minnesota. Arthritis Rheum. 2010;62(6):1576–82.
2. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet. 2010;376(9746):1094–108.
3. Viatte S, Barton A. Genetics of rheumatoid arthritis susceptibility, severity, and treatment response. Semin Immunopathol. 2017;39(4):395–408.
4. Groves C, Chandramohan M, Chew NS, Aslam T, Hellwell PS. Clinical examination, ultrasound and MRI imaging of the painful elbow in psoriatic arthritis and rheumatoid arthritis: which is better, ultrasound or MRI, for imaging enthesis? Rheumatol Ther. 2017;4(1):71–84.
5. Kourilovitch M, Galarza-Maldonado C, Ortiz-Prado E. Diagnosis and classification of rheumatoid arthritis. J Autoimmun. 2014;48-49:26–30.
6. De Cock D, Hynich K. Malignancy and rheumatoid arthritis: epidemiology, risk factors and management. Best Pract Res Clin Rheumatol. 2018;32(6):869–86.
7. Misuiniene N, Baranauksaitė A. diagnosis and treatment of rheumatoid arthritis. Medicina (Kaunas). 2002;38(1):11–6.
8. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. Ncbi ge: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41(Database issue):D991–5.
9. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773–82.
10. Harrison WT, Yeates RM, Phillips ML, Nenoff TM. New framework connectivity patterns in templated networks: the creatinine zinc phosphites c4n3oh7.Zn(h2o)hpo3, and (c4n3oh7)2.Znhpo3.H2o. Inorg Chem. 2003;42(5):1493–8.
11. Cho Y. Understanding the dynamics: pathways involved in the pathogenesis of rheumatoid arthritis. Rheumatol (Oxford). 2012;51(Suppl 5):v3–11.
12. Liu L, Wong CW, Han M, Farhoudi HP, Liu G, Liu Y, et al. Meta-analysis of preclinical studies of mesenchymal stromal cells to treat rheumatoid arthritis. EBioMedicine. 2019;47:563–77.
13. Smolen JS, Aletaha D, Rudolph K. The pathogenesis of rheumatoid arthritis: new insights from old clinical data? Nat Rev Rheumatol. 2012;8(4):235–43.
14. Raterman HG, Vossberm S, de Ridder S, Nurmohamed MT, Lems WF, Boers M, et al. The interferon type i signature towards prediction of non-response to rituximab in rheumatoid arthritis patients. Arthritis Res Ther. 2012;14(2):R95.
15. Stuhlmueller B, Haupl T, Hernandez WM, Grutzke A, Kuban R, Tandon N, et al. Cd11c as a transcriptional biomarker to predict response to anti-tnf monotherapy with adalimumab in patients with rheumatoid arthritis. Clin Pharmacol Ther. 2010;87(3):311–21.
16. Fraser A, Vallow J, Preston A, Cooper RG. Predicting ‘normal’ grip strength for rheumatoid arthritis patients. Rheumatology (Oxford). 1999;38(6):521–8.
17. Stenstrom CH, Minor MA. Evidence for the benefit of aerobic and strengthening exercise in rheumatoid arthritis. Arthritis Rheum. 2003;49(3):428–34.

18. Yamada T, Steine MM, Kunne E, Lanner JT. Muscle weakness in rheumatoid arthritis: the role of ca (2+) and free radical signaling. ElBioMedicine. 2017;23:12–5.

19. Zheng Z, Mengaet AM, Fahy NM, Bawadekar M, Holmes CL, Ong IV, et al. Disordered antigens and epitope overlap between anti-citullinated protein antibodies and rheumatoid factor in rheumatoid arthritis. Arthritis Rheumatol. 2020;72(2):622–62.

20. Edhayan G, Ohara RA, Stinson WA, Amin MA, Isozaki T, Ha CM, et al. Inflammatory properties of inhibitor of DNA binding 1 secreted by synovial fibroblasts in rheumatoid arthritis. Arthritis Res Ther. 2016;18:37.

21. Gao D, Nolan DJ, Mellick AS, Bambino K, McDonnell K, Mittal V, Endothelial progenitor cells control the angiogenic switch in mouse lung metastasis. Science. 2008;319(5860):195–8.

22. Kleinert S, Waldner M, Wendler J, Kunzmann V. Disease-modifying antirheumatic drugs in rheumatoid arthritis patients with a history of colorectal cancer. Z Rheumatol. 2016;75(1):41–6.

23. Simon A, Subra JF, Guillaumin P, Jeaenin P, Pignon P, Blanchard S, et al. Detection of anti-pentrin-3 autoantibodies in anca-associated vasculitis. PLoS One. 2016;11(1):e0147091–e0147091.

24. Firestein GS, McInnes IB. Evidence that cytokines play a role in rheumatoid arthritis. Immunol Rev. 2010;233(1):233–55.

25. Firestein GS, McInnes IB. Immunopathogenesis of rheumatoid arthritis. Immunol. 2017;46:183–96.

26. Kleinert S, Waldner M, Wendler J, Kunzmann V, Munch-Baase N. Disease-modifying antirheumatic drugs in rheumatoid arthritis patients with a history of colorectal cancer. Z Rheumatol. 2016;75(1):41–6.

27. Bartok B, Firestein GS. Fibroblast-like synoviocytes: key effector cells in rheumatoid arthritis. Immunol Rev. 2010;233(1):233–55.

28. Brennan FM, McInnes IB. Evidence that cytokines play a role in rheumatoid arthritis. J Clin Invest. 2008;118(1):537–45.

29. Feldmann M, Maini RN. Anti-tnf alpha therapy of rheumatoid arthritis: what have we learned? Annu Rev Immunol. 2001;19:163–96.

30. Dixon WG, Symmons DP. What effects might anti-tnfalpha treatment be having on the growth rate of the disease? Ann Rheum Dis. 2007;66(9):1132–3.

31. Li XF, Sun YY, Bao J, Chen X, Li YH, Yang Y, et al. Functional role of ppar-g in the regulation of adipogenesis. J Am Heart Assoc. 2013;2:6(4):e000441.

32. Brenner M, Laragione T, Gulko PS. Increased synovial expression of Tnf-alpha in rheumatoid arthritis. Arthritis Rheum. 2004;50(9):2794–8.

33. Avci AB, Feist E, Burmester GR. Targeting il-6 or il-6 receptor in rheumatoid arthritis: What's the difference? BioDrugs. 2013;26(6):531–46.

34. Yamasaki S, Nakashima T, Kawakami A, Miyashita T, Tanaka F, Ida H, et al. The impaired gene expression of adenosine monophosphate-activated AMPK-a, a key metabolic enzyme in leukocytes of newly diagnosed rheumatoid arthritis patients. Mol Biol Rep. 2019;46(6):6353–60.

35. Balliet A, Trocccm C, Berthier S, Arlotto M, Grange L, Chenau J, et al. Synovial fluid proteomic fingerprint: s100a8, s100a9 and s100a12 proteins discriminate rheumatoid arthritis from other inflammatory joint diseases. Rheumatol. 2010;49(4):671–82.

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.