Mapping the Regulatory Programs of RNA Binding Protein Regulators in Rheumatoid Arthritis: Data from Single-Cell Transcriptome Analysis

Hongbin Luo
the First Affiliated Hospital of Fujian Medical University
Duoduo Lin
Fujian Medical University
Jie Wei
Fujian Medical University
Qunya Zheng
Fujian Medical University
Nanwen Zhang
Fujian Medical University
Peng Chen (13506989645@163.com)
the First Affiliated Hospital of Fujian Medical University

Research Article

Keywords: Single Cell Transcriptome Analysis, RNA Binding Proteins, rheumatoid arthritis, cell–cell communication, immunocytes

Posted Date: June 14th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3034461/v1

Abstract

Background

RNA binding proteins (RBPs), especially cell-specific RBPs are involved in critical processes such as alternative splicing of messenger RNAs and translational control, leading to the expression of cell-specific functional proteins. However, the expression pattern of RBPs in different cells of rheumatoid arthritis and their associated aberrant regulation remain largely unexplored.

Methods

We collected 2141 RNA binding protein genes (RBPs) from literature and identified cell populations present in rheumatoid arthritis and osteoarthritis control samples using single-cell data. We compared the changes in the relative proportions of cell classes between them and analyzed RBP expression patterns specific to different cell types. We investigated fibroblast cell populations and their cellular communication with different immune cells. Additionally, we used bulk RNA-seq data from rheumatoid arthritis and osteoarthritis samples to identify highly conserved variable splicing events and established a co-variation network of RBPs and these splicing events.

Results

We observed a greater number of down-regulated RBPs in each cell type, except for fibroblasts, endothelial cells, and macrophages, where the number of up-regulated genes was much higher. In fibroblasts from RA and OA patients, we identified 105 upregulated RBPs and 133 downregulated RBPs. These RBPs were co-expressed with genes enriched in various functional pathways, including extracellular matrix organization, cell adhesion, collagen fibril organization, and cytokine signaling. Cellular communication analysis demonstrated enhanced signaling pathways, like CXCL12-CXCR4, between fibroblasts and macrophages in RA. We identified a total of 715 differentially variable splicing events in our study, and alternative 5' and 3' splicing were the most prevalent. Some RBPs, such as MBNL2 in endothelial cells and U2AF1, SF3B6, and SF3B14 in fibroblast cells, may play a role in the pathogenesis of RA through splicing regulation.

Conclusion

In this study, we analyzed single-cell datasets to identify the inherent characteristics and abnormal expression patterns of RBPs in different cell types of patients with RA. Our findings revealed that certain cell-specific RBPs were associated with inflammatory signaling pathways and splicing regulation in RA. These findings suggest that the dysregulation of RBPs may contribute to the development of RA and highlight potential pathways for therapeutic interventions.

1 Introduction

Rheumatoid Arthritis (RA) is an autoimmune disease characterized by long-lasting joint inflammation, causing inflammation and eventual deformity, leading to a severe decline in daily activities, work performance, and overall quality of life(1). Synovial tissue in RA patients is a mixture of various cell types, including T cells, B cells, monocytes, and fibroblasts. Specifically, the RA fibroblast-like synovial cells (FLS) play a major part in starting and perpetuating the NF-κB pathway, which causes inflammation, excessive multiplying of cells, and the invasion of cartilage(2). A thorough understanding of the role these cells play and the underlying mechanisms is essential to develop effective treatments and identify new targets for RA therapy.

Over the past years, several studies have utilized single-cell transcriptome sequencing to analyze and interpret the interactions between different RA synovial tissue subpopulations of cells. This has enabled researchers to assess the impacts of treatment on RA and identify new potential targets for therapy (3, 4). Despite these advances, the triggers that cause inflammation in RA remain largely unknown. Therefore, a deeper understanding of the complex molecular mechanisms and cellular processes involved in RA is crucial to progress towards developing new and effective therapeutic options.

As essential regulators of gene expression, RNA binding proteins (RBPs) have a critical role in dictating the development and fate of specific RNA substrates (5). RBPs contribute to several aspects of RNA regulation, including but not limited to, mRNA splicing, RNA cleavage and polyadenylation, RNA localization, RNA stability and editing, and RNA translation (6). RBPs play a crucial role in maintaining cellular homeostasis by recognizing and regulating hundreds of transcripts. Dysfunctions in RBPs have been implicated in various diseases(7, 8), including immune-based disorders(9–12). Importantly, cell-specific RBPs are involved in critical processes such as alternative splicing of messenger RNAs and translational control, leading to the expression of cell-specific functional proteins. Aberrant regulation of RBPs can lead to dysfunctional cellular functions, which is associated with many clinical disorders, including rheumatoid arthritis(13). While recent studies have explored the function and mechanisms of RBPs in individual cells(14–16), the expression pattern of RBPs in different cells of rheumatoid arthritis and their associated aberrant regulation remain largely unexplored.

The introduction of single-cell transcriptome sequencing technology has enabled the identification of gene expression levels in individual cells within a population, thereby underscoring the presence of intercellular heterogeneity. Over the last few years, multiple research studies have implemented the single-cell RNA sequencing (scRNA-seq) approach to unveil the gene expression patterns characteristic of individual cells during the development and progression of RA. The transcriptional datasets collected from individual cells provide valuable understanding into the fundamental mechanisms of the disease, ultimately paving the way for the creation of novel, more precise therapeutic interventions. By examining these datasets, we aim to discover the inherent characteristics and atypical expression patterns of RBPs in different cell types affected by RA, thereby providing insight into the potential key involvement of RBPs in the disease's development.

2 Materials and methods

2.1 Retrieval and process of scRNA-seq data

For retrieving and processing the scRNA-seq data, we obtained the count matrix of unique molecular identifiers (UMI) for the scRNA-seq data from three OA and four RA samples, which were downloaded from GSE152805 and GSE200815 respectively. After conversion of this UMI count matrix into a Seurat object (17) using the R package Seurat (version 4.0.4), cells with UMI counts below 1000, genes detected in fewer than 500 cells, and UMI counts derived from mitochondrial genes exceeding 15% were treated as low-quality cells and removed from further analysis. Finally, genes detected in fewer than five cells were also removed from downstream analyses.

2.2 Retrieval and process of bulkRNA-seq data

We downloaded publicly available sequence data files for five rheumatoid arthritis tissue and five osteoarthritis tissue samples (www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89408) from the Sequence Read Archive (SRA). Using the NCBI SRA Tool, fastq-dump, we converted the SRA run files to fastq format. We trimmed out low-quality bases from the raw reads using the FASTX-Toolkit (v.0.0.13) (http://hannonlab.cshl.edu/fastx_toolkit/), and performed quality evaluation of the resulting clean reads with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc).

2.3 Reads alignment and differentially expressed gene (DEG) analysis

The RNA-seq data sequencing was analyzed as follows: First, the clean reads were aligned to the human GRCh38 genome using HISAT2 (v.2.2.1) (18). Uniquely mapped reads were then selected for further analysis, and the reads located on each gene were counted. The expression levels of the genes were estimated using FPKM (fragments per kilobase of exon per million fragments mapped). The DESeq2 software (version 1.30.1) was utilized to carry out differential gene expression analysis on the reads count file (19). For the differential expression analysis, DESeq2 was also used to determine the Fold Change (FC) and False Discovery Rate (FDR) between two or more samples to identify differentially expressed genes. The significant differential expression criteria were as follows: FC values more than or equal to 2 or lower than or equal to 0.5, with an FDR of lower than or equal to 0.05.

2.4 Alternative splicing analysis

The SUVA (v2.0) pipeline was utilized to quantify and define regulatory alternative splicing events (RAS) (20). The different splicing patterns for each group of alternative splicing events (ASEs) were analyzed, followed by the calculation of the pSAR (proportion of reads supporting each ASE event) for each SUV ASE event.

2.5 RBP-RAS Co-expression analysis

Initially, a comprehensive catalog of RBPs was compiled, consisting of 2,141 RBPs from four previous reports (6, 21–23). These RBPs were overlapped with the set of Regulatory Alternative Splicing (RAS) events (pSAR ≥ 50%), and co-expression analysis of RBP-RAS was conducted. To assess the relationship between the expression levels of each RBP and RAS, the Pearson correlation coefficient was used. RBP-RAS pairs with an absolute correlation coefficient greater than or equal to 0.6 and a p-value less than or equal to 0.01 were identified and selected for downstream analyses.

2.6 scRNA-seq data preprocessing and quality control

After performing quality control of the UMI count matrix, the data underwent log normalization. Subsequently, the top 2000 variable genes were used to create potential anchors with the FindIntegrationAnchors function of Seurat. To further reduce the dimensionality of the scRNA-Seq dataset, principal component analysis (PCA) was applied to the integrated data matrix. The top 50 PCs were then selected for downstream analyses using the Elbowplot function of Seurat. Main cell clusters were identified using the FindClusters function of Seurat with the default resolution setting (res = 0.6). A total of 18 major cell types were then identified and visualized using tSNE or UMAP plots. Cell types were annotated using ScTypetools (24), which identified gene markers for each cell cluster using the "FindMarkers" function in the Seurat package (v4.0.4). Additionally, pre-existing marker genes of osteoarthritis (25) were also used for this cell type annotation.

2.7 scRNA-seq Differential gene expression analysis

To identify the DEGs, the Seurat package FindMarkers/FindAllMarkers function was utilized with a one-tailed Wilcoxon rank-sum test, and the p-values were adjusted for multiple testing using the Bonferroni correction. In computing DEGs, we ensured that the expression difference of all genes on a natural log scale was at least 0.5, with a difference in the percentage of detected cells of at least 0.15. Furthermore, the adjusted p-value was set to less than 0.05, ensuring that the results obtained were statistically significant.

2.8 RBP genes analysis

We first compiled a comprehensive catalog of 2,141 RBPs from four prior reports [5–8]. Then, we utilized the UMI count matrix of RBPs as input for Seurat to perform cell clustering. Using the "FindAllMarkers" function of the Seurat package, we selected differentially activated RBPs. To identify co-expression associations between differential RBP and targeted genes, we employed the "grn" algorithm from the SCENIC (6) python workflow (version 0.11.2) with the default parameters (http://scenic.aertslab.org). The networks of the modules consisting of differential RBP and their target genes were visualized using

Cytoscape (v3.9.1) (https://cytoscape.org/). Additionally, through co-expression analysis by extracting the differential RBP from fibroblasts, we could identify co-expression associations between differential RBP and their targeted genes.

2.9 Functional enrichment analysis

To determine the functional roles of the studied genes, KOBAS2.0 (26) was used to identify Gene Ontology (GO) terms and KEGG pathways. Enriched categories were detected through the hypergeometric test and significance of each term was assessed using the Benjamini-Hochberg FDR controlling procedure. This ensured the statistical significance of the enriched GO terms and KEGG pathways, providing meaningful insights into the function of the target genes.

2.10 Cell–cell communication

The determination of cell-to-cell interactions in different cell types, based on the expression of known ligand-receptor pairs, was carried out using CellChat (27) (v1.0.0). To identify potential communication networks that could be disrupted or induced in intervertebral disc degeneration, we followed the official workflow and inputted the normalized counts into CellChat. Preprocessing functions, such as identifyOverExpressedGenes, identifyOverExpressedInteractions, and projectData, were applied with standard parameters. We used the human protein-protein interactions database as a priori network information. Core functions such as computeCommunProb, computeCommunProbPathway, and aggregateNet were applied using standard parameters and fixed randomization seeds for the main analysis. The function netAnalysis_signallingRole was applied on the netP dataslot to identify the senders and receivers in the network. The results obtained provide significant insights into the potential cell-to-cell communication networks involved in intervertebral disc degeneration.

2.11 Other statistical analysis

To showcase the sample clustering using the first two components, we conducted PCA using the R package factoextra (https://cloud.r-project.org/package=factoextra). The clustering was based on the PCA results using Euclidean distance, and performed using the R package pheatmap (https://cran.r-project.org/web/packages/pheatmap/index.html). Additionally, to compare two groups of replicates, we utilized the R package speckle (version: 0.0.3) (28). The analysis results provide valuable insights into the sample clustering and comparisons between replicate groups, contributing to a better understanding of the underlying factors in the experiment.

3 Results

3.1 ScRNA-seq analysis of synoviocytes samples from patients with osteoarthritis and rheumatoid arthritis identified different cell types.

In this study, we obtained synovial tissue single-cell transcriptome data from the published dataset GSE200815 for RA patients (4 cases) and selected synovial tissue single-cell transcriptome data from GSE152805 for OA patients (3 cases) as a control for integration analysis. We then utilized RBP for single-cell subclustering to analyze the mechanisms of RBP regulation at the single-cell level that differed between rheumatoid arthritis and OA. Moreover, we collected bulk RNA-seq data from the GSE89408 dataset, and analyzed variable splicing events that were significantly regulated in RA compared to OA controls using the newly published SUVA software. We constructed a co-variant regulatory network of RBP and variable splicing events, as illustrated in Fig. 1A. The analyses offered significant insights into the unique mechanisms of RBP regulation and variable splicing that differentiate RA and OA. These insights contribute to a deeper understanding of the underlying factors of RA.

Following rigorous data quality control, a total of 33,226 single cells were included in the transcriptome profile analysis. Following normalization of the transcriptomic expression of these cells, principal component downsizing analysis was applied, and the top 50 principal components were selected for UMAP (Uniform Manifold Approximation and Projection) downsizing and visualization. Through unbiased cluster analysis, we identified 17 cell subgroups (Fig. 1B, Figure S1A-1B). Based on the characteristic expressed genes of the clusters, combined with previously reported synovial cell marker genes, we identified 9 different cell types (Fig. 1C-D).

We proceeded to compare the relative abundance of cell subtypes across all sample groups, and found that the RA and OA groups exhibited notably more pronounced alterations. In the RA group, we observed marked increases in proportions of cells including C2, C12, C7: Endothelial, C6: T, and C14: B, comparing to the OA group. Conversely, proportions of cells such as C15: Mastcells, C8, C4, and C0: Fibroblasts were found to decrease (Fig. 1E-F). Gene ontology enrichment analysis showed that up-regulated and down-regulated genes of each cell type in two comparison groups were enrich in several biological processes (Figure S1C-1D). We conducted a differential expression analysis between RA and OA groups for each cell type separately, and Fig. 1G shows the number of genes and differentially expressed RBPs in each cell type. Our analysis revealed that in each cell type, the count of down-regulated RBPs exceeded that of up-regulated ones. Nonetheless, when it came to fibroblasts, endothelial cells, and macrophages, the overall number of up-regulated genes vastly outnumbered that of down-regulated genes. Furthermore, In Fig. 1H, it can be observed that the overall AUC levels of RBPs were comparatively lower in the RA group as opposed to the OA group, suggesting reduced RBP activity in the RA group, which may affect transcriptional and post-transcriptional gene regulation. Collectively, these results provide a comprehensive understanding of the changes in cell composition and changes in rheumatoid arthritis and OA control synovial cells.

3.2 ScRNA-seq analysis identified heterogeneity and regulatory module of cellular specific RBP expression module.

Our aim was to investigate changes in RBP expression patterns in both the OA control group and the RA group. We conducted unsupervised clustering using Seurat based on the expression of 2,141 reported RBP genes. We observed that the resulting clusters of RBP expression (referred to as RBP-expressing cell clusters) were highly cell-type specific and correlated with disease status (Fig. 2A, Figure S2A-B). By analyzing the composition of RBP-expressing cellular taxa in different cell types, we found that for most cell types, the pattern of RBP expression was highly specific to that particular cell type, with an overwhelmingly dominant RBP-expressing cellular taxon (Fig. 2B). These findings provide evidence for cell type-specific RBP expression patterns in both disease states and control groups. Gene ontology enrichment analysis showed that these RBPs were enrich in several biological processes (Figure S2C). By shedding light on the function of RBPs, our findings help advance the comprehension of their role in synovial tissue during rheumatoid arthritis and osteoarthritis.

In addition, we found that the composition of RBP expression taxa changed significantly between the RA and OA groups for almost all cell types. For example, the main RBP expression taxon in the RA group was R1 for endothelial cells, whereas it was R13 for the OA group (Fig. 2B). This result suggests that different cell types have specific RBP expression patterns that are highly correlated with disease status. Specifically, different RBP expression taxa have specific marker RBP genes, and we demonstrated the expression of these RBPs in different sample subgroups and cell types (Fig. 2C). For instance, osteoglycine (OGN) and Filamin B (FLNB) is one of the markers of RBPs of R2, and we observed that these genes are mainly highly expressed in fibroblasts (Fig. 2D, Figure S2D). On the other hand, Muscleblind-like splicing regulator 2 (MBNL2) is one of the markers of RBPs of R1 and is predominantly expressed in endothelial cells of the RA group (Fig. 2E).

3.3 Functional RBPs are largely regulated in Fibroblasts cells between RA and OA samples

In particular, fibroblast-like synoviocyte (FLS) in RA have been identified as key actors in both the activation and maintenance of the RA-induced NF-κB pathway, which in turn promotes local proliferation, production of pro-inflammatory cytokines, and cartilage invasion. To better understand the potential functions of RBPs, which are key regulators of transcriptional and post-transcriptional processes, we isolated fibroblast cell classes and conducted secondary clustering and analysis using RBP-Genes. The resulting RBP-expressing cell classes differed significantly between the OA and RA groups. Specifically, the RA group primarily expressed FR1, FR3, FR4, FR5, FR5, FR6, and FR7, while the OA group mainly expressed FR0, FR2, FR8, and FR9. These observations indicate that the expression pattern of RBPs

differs considerably between RA and OA and may have significant implications for the transcriptional and post-transcriptional regulation of fibroblasts (Fig. 3A-B, Figure S3A-B).

In our analysis comparing fibroblasts from RA and OA patients, we identified 105 upregulated RBPs and 133 downregulated RBPs. Figure 3C displays the top 20 upregulated and downregulated RBPs. These differentially expressed RBPs were co-expressed with genes enriched in various functional pathways, such as extracellular matrix organization, cell adhesion, collagen fibril organization, and cytokine signaling. These pathways suggest potential roles for RBPs in regulating fibroblast proliferation, antigen presentation, and pro-inflammatory responses (Fig. 3D-E, Figure S3C). Our study also revealed specific RBPs associated with important cellular pathways in RA, such as YBX3 and EIF4A1 (Fig. 3F-G). Additionally, we identified significant differences in the expression of important splicing factors, including U2AF1, SF3B6, and SF3B14, between RA and OA groups (Figure S3D, Figure S2E). These observations suggest that variations in splicing regulation may contribute to the disease pathology of RA.

3.4 RBP-mediated fibroblasts subpopulations contributed to the aberrant activation of signaling pathways with immune cells during rheumatoid arthritis.

scRNA-seq has been applied successfully to predict potential LR interactions, revealing the importance of crosstalk between different cell types in various disease mechanisms. To investigate further the potential role of RBPs in the interactions between fibroblasts and immune cells, we performed cellular communication analysis of different RBP-expressing cell populations of fibroblasts and immune cells. Based on our research, we discovered that the level of interconnection and strength between fibroblasts and immune cells was noticeably higher in the RA group when compared to the OA group (Fig. 4A). In addition, fibroblasts and macrophages displayed the highest communication strength among the different immune cells (Fig. 4B). We then identified several signaling pathways, such as CXCL12-CXCR4, that were enhanced in RA and known to be associated with the disease (Fig. 4C-D). Lastly, changes in ligand expression levels of these key signaling pathways could be regulated by RBPs (Fig. 4E).

3.5 Identification of highly conserved RA-associated AS events co-disturbed with differentially expressed RBPs in RA and OA patients.

Disruption of normal RBP function can lead to cellular dysfunction by affecting post-transcriptional processes, such as variable splicing of RNA. To explore further the role of RBPs associated with RA in variable splicing regulation, we downloaded bulk RNA-seq data from the GSE89408 dataset, which included five RA and five OA samples as controls. Given the complex nature of human variable splicing, we utilized the SUVA software tool, which is a recently developed tool, to identify significant differences in

variable splicing events between OA controls and RA. Our analysis using SUVA identified 715 differential variable splicing events, mainly alternative 5' and 3' splicing (Fig. 5A). Among the SUVA-identified splice events, alternative 5' splice site, cassette exon, exon skipping, and alternative 3' splice site were the most frequently detected differentially variable splicing events (Fig. 5B).

A single splicing event usually encompasses two transcripts, which make up only a small portion of the overall gene expression. To identify the leading transcripts where splicing occurs, we screened out 556 splicing events (pSAR > = 50%) that accounted for the dominant transcripts for further analysis (Fig. 5D). PCA analysis based on the splicing ratio of these dominant transcripts clearly separated the two sample sets, suggesting that the RNA splicing landscape is closely associated with the development of rheumatoid arthritis (Fig. 5D). These 556 transcripts hosting the identified splicing events are enriched in multiple functional pathways, including cytoskeleton organization, GTPase activation (which is associated with cellular processes such as programmed cell death regulation), and damage repair (Fig. 5E).

RBPs have been identified as crucial regulators of variable splicing. To predict the potential regulatory relationship between RBPs and variable splicing during rheumatoid arthritis development, Initially, we identified genes that exhibited a significant difference in expression levels between RA and OA controls. Our findings indicated that a large number of genes were activated during disease progression, with 2183 genes showing up-regulation and 823 genes showing down-regulation, as illustrated in Fig. 5F. Out of these genes, 209 RBPs showed up-regulation, while 232 were down-regulated, as shown in Fig. 5G-H. We then overlapped the differential RBPs obtained from scRNA-seq and bulk RNA-seq to identify conserved regulated RBP genes, finding three co-up-regulated (SYNE2, S100A9, IFIT3) and four co-down-regulated genes (RNASE1, GRN, FN1, SORBS2) (Fig. 5G, H, I). To predict the potential regulatory role of RBPs on variable splicing, we performed a co-variation analysis using these co-regulated RBPs and RAS. After screening (|correlation| >= 0.6, p-value < = 0.01), we identified RASs that were significantly associated with these RBPs (Fig. 5J).

4 Discussion

In this study, we utilized single-cell data to identify and annotate cell populations from samples of RA and OA groups. According to our results, the number of down-regulated RBPs was higher than that of up-regulated RBPs in each cell type, except for fibroblasts, endothelial cells, and macrophages. We observed distinct expression patterns of RBP genes specific to different cell types along with changes between the OA and RA groups. Specifically, we focused on the fibroblast cell population and analyzed the RBP genes specifically expressed in the RA group. Our results revealed that RBP-mediated fibroblast subpopulations contributed to the aberrant activation of signaling pathways with immune cells during rheumatoid arthritis. Furthermore, we identified 556 transcripts hosting the identified splicing events that were enriched in multiple functional pathways, including cytoskeleton organization, GTPase activation, and damage repair in RA patients. The findings obtained in our study offer novel perspectives on the intricate regulatory mechanisms of RBPs in the pathogenesis of RA.

In recent times, a variety of evidence has emerged that highlights the crucial role played by RBPs in controlling posttranscriptional regulatory mechanisms that influence the immune system in both healthy and pathological conditions. These regulatory mechanisms impact protein expression patterns by directly controlling mRNA through processes like pre-mRNA splicing and maturation, mRNA transport to the cytoplasm, and management of mRNA stability, storage, and translation. According to our results, the number of down-regulated RBPs was higher than that of up-regulated RBPs in each cell type. These findings are in agreement with previous reports. By employing silencing protocols, studies conducted both in vitro and in vivo showed that the downregulation of RBPs can significantly impact key regulatory mechanisms contributing to the pathogenesis of arthritis (29, 30).

Autoimmune diseases arise from an imbalance between immune response activation and suppression. The dysregulation is the result of overproduction of pro-inflammatory cytokines, particularly IL-6 and TNF. The expression of pro-inflammatory mediator genes is tightly regulated at the post-transcriptional level, which is mediated by immune-associated RBPs. In the context of RA, decreased levels of RBP activity suggest that the post-transcriptional regulation of pro-inflammatory mediator genes may be compromised. TTP is a well-researched RNA-binding protein that plays a significant role in destabilizing the mRNA of pro-inflammatory cytokines. Given its ability to target and reduce the mRNA levels of several inflammatory cytokines, including TNF-a, it is generally considered to play a crucial function in the development of RA pathology (31). This finding is supported by the research of Yang et al. (32), whose study reported decreased levels of TTP mRNA in PBMCs collected from patients with RA. Additionally, the authors identified an association between a SNP (rs3746083) located in the gene encoding for TTP and RA susceptibility in Chinese RA patients.

We observed a noteworthy finding that specific RBP expression profiles are highly associated with the disease state. For instance, Osteoglycin (OGN) is actively involved in regulating various cellular processes, for instance, cell growth, differentiation, and adhesion. It belongs to the SLRP (small leucine-rich proteoglycan) family of proteins (33–35). OGN has been implicated in a range of physiological and pathological processes, including skeletal development (36), tissue inflammation, cardiovascular disease (37), and cancer (38, 39). Furthermore, OGN has been shown to modulate collagen fibril organization and maturation, thereby influencing connective tissue integrity. Our analysis revealed that OGN is predominantly expressed in fibroblasts, indicating its potential role in regulating fibroblast-specific functions. In addition, MBNL2 is an essential regulator of alternative splicing, and its expression has been linked to the regulation of endothelial cell differentiation (40–42). Our analysis revealed that MBNL2 is mainly expressed in endothelial cells of the RA group (Fig. 2E), indicating its possible involvement in the pathogenesis of RA via its role in splicing regulation in endothelial cells.

Fibroblasts, which are a crucial cell type in the synovium of RA patients, are well-documented to play a pivotal role in driving the disease process. Through our analysis of fibroblasts derived from RA and OA patients, we identified a total of 105 upregulated RBPs and 133 downregulated RBPs. These differentially expressed RBPs were found to be co-expressed with genes involved in various functional pathways, including those related to extracellular matrix organization, cell adhesion, collagen fibril organization, and

cytokine signaling. Moreover, our study identified specific RBPs that were associated with critical cellular pathways. Our study revealed that certain RBPs were significantly enriched in RA patients, including YBX3 and eIF4A1. YBX3 is classified under RBP, which plays a significant part in the post-transcriptional regulation of mRNA stability and translation (43). It has been instrumental in a wide variety of biological processes, including spermatogenesis, development, as well as cellular differentiation and proliferation (44). On the other hand, eIF4A1, which is a DEAD-box RNA helicase, functions in two major reactions during translation initiation. Firstly, the loading of mRNA onto the 43S pre-initiation complex is dependent on the activity of eIF4A1. Secondly, by unwinding RNA secondary structures, eIF4A1 allows the translocation of the PIC along the 5'UTR of the mRNA that has high structural content (45). These results suggest that dysregulated expression of RBPs such as YBX3 and eIF4A1 may contribute to the pathogenesis of RA and highlight their potential as therapeutic targets for this debilitating disease.

The process of alternative splicing plays a crucial role in regulating gene expression across eukaryotic organisms. This particular process enables the creation of several unique mRNA species from a single gene. Alternative splicing exhibits a varied range of mechanisms, including exons skipping (removal of specific cassette exons), mutual exclusive exon choice, alternative splicing events (affecting the boundaries between introns and exons, leading to variations in transcript diversity), and intron retention. The alternates splicing phenomena can result in mRNAs with varying untranslated regions (UTRs) or coding sequences, which further impact mRNA stability, translation, and localization. Moreover, variations in the reading frame due to alternative splicing can produce diverse protein isoforms serving distinct functions and localizations (46). Dysregulation of this process can lead to changes in gene expression and alterations in protein function, contributing to various disease states. In our study, we identified a total of 715 splicing events that were differentially variable, with alternative 5' and 3' splicing being the most prevalent. Among these splicing events, alternative 5' splice site, cassette exon, exon skipping, and alternative 3' splice site were the most commonly detected. Interestingly, our analysis revealed that 556 transcripts that host the identified splicing events were enriched in various functional pathways, including cytoskeleton organization, GTPase activation (associated with programmed cell death regulation), and damage repair. These findings suggest that dysregulated alternative splicing may contribute to the pathogenesis of RA and highlight potential pathways for therapeutic intervention.

RNA-binding proteins have a vital role to play in the process of alternative splicing, where they take up binding in pre-mRNA transcripts and facilitate splicing. Acting either as activators or repressors of splicing events, RBPs can influence the inclusion of particular exons in the final mRNA transcript either by promoting or inhibiting the same. Some RBPs can also interact with splice sites or splicing factors to modulate the splicing process. Dysregulated expression or function of RBPs can lead to aberrant alternative splicing patterns, which can be attributed to the development and progression of several disorders, including some cancer (47) and neurodegenerative disorders (48). In our study, we found that MBNL2 was predominantly expressed in endothelial cells of the RA group, indicating its potential involvement in the pathogenesis of RA through splicing regulation in endothelial cells. Additionally, we observed significant differences in the expression of key splicing factors, including U2AF1, SF3B6, and SF3B14, in fibroblast cells between the RA and osteoarthritis (OA) groups. U2AF1 is an RNA-binding

protein that plays a crucial role in the recognition of the 3' splice site during pre-mRNA splicing. U2AF1 is essential for initiating U2 small nuclear ribonucleoprotein particle (snRNP) recruitment to the splice site (49). Several studies have suggested that cells expressing mutant U2AF1 exhibit changes in alternative splicing patterns. Mutations in U2AF1 may affect the binding affinity of the protein to RNA, leading to aberrant splicing patterns and contributing to the development of various diseases, including myelodysplastic syndromes (50, 51) and lung cancer (52). The SF3b complex is an essential part of the spliceosome and is a predominant component of U2 and U11/U12 snRNPs (53). It plays a crucial role in recognizing the branch-point sequence (BPS), stabilizing the U2snRNA/BPS duplex, and preventing any untimely transesterification (54). SF3b6, a distinctive feature of SF3b in humans, contains an RNA recognition motif (RRM) that directly interacts with the branch-point adenine. However, studies demonstrate the lack of specificity of SF3b6 for the BPS, as it does not differentiate between adenosine monophosphate, adenine, or single-stranded RNA over double-stranded RNA/DNA (55). Structural analysis of SF3b6 reveals that its RNA binding region in RRM is obstructed by its own helix α3, the C-terminal tail, and an interacting region from SF3b1. SF3b14, also known as SAP130, is a subunit of the SF3b complex. SF3b14 is a critical player in the identification of the BPS during the processing of pre-mRNAs. It works in conjunction with the other subunits of the SF3b complex and aids in stabilizing the U2snRNA/BPS duplex, which is crucial for the formation of the spliceosome. This stabilization mechanism is essential for the proper splicing of pre-mRNAs and ensures a high-quality end product (57). These findings underscore the critical role of RBPs in regulating alternative splicing and highlight their potential as therapeutic targets for diseases related to aberrant splicing.

In conclusion, we analyzed single-cell RNA sequencing datasets to identify the inherent characteristics and abnormal expression patterns of RBPs in different cell types of patients with RA. Our findings revealed that certain cell-specific RBPs were associated with inflammatory signaling pathways and splicing regulation, highlighting their potential role in the pathogenesis of RA. These results suggest the dysregulation of RBPs may contribute to RA and provide a basis for identifying potential pathways for therapeutic intervention in the treatment of RA. Additionally, there were certain limitations to our study that should be considered. For instance, our results were primarily based on data from public databases, and we did not have cellular or animal experimental results to validate our findings. Therefore, further research is required to fully comprehend the specific role and mechanism of certain RPB in the development and progression of RA.

Abbreviations

ASEs alternative splicing events

BPS branch-point sequence

DEG Differentially expressed gene

FC fold change

FDR false discovery rate

FLS fibroblast-like synovial cells

GO Gene Ontology

KEGG Kyoto Encyclopedia of Genes and Genomes

OA Osteoarthritis

OGN Osteoglycin

PCA Principal component analysis

pSAR proportion of each SUVA AS event

RA Rheumatoid arthritis

RAS regulatory alternative splicing events

RBPs RNA binding proteins

RRM RNA recognition motif

scRNA-seq single-cell RNA sequencing

snRNP small nuclear ribonucleoprotein particle

SRA Sequence Read Archive

UMI unique molecular identifier

UTRs untranslated regions

Declarations

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

The article was written by Hongbin Luo, while experiments were conducted by Duoduo Lin, Jie Wei, and Qunya Zheng. Peng Chen and Nanwen Zhang reviewed the article. All authors have reviewed and approved the submitted version of the article.

Funding

This work was supported by Fujian Provincial Health Technology Project (No. 2020GGA043), Joint Funds for the Innovation of Science and Technology, Fujian Province (No. 2019Y9111), Natural Science Foundation of Fujian Province of China (No. 2023J01610 and 2023J01307), and Fujian Provincial Finance Project (No. BPB-LHB2021).

List of Abbreviations

Ethics approval for consent to participate

Not applicable.

Consent for publication

Not applicable.

Acknowledgements

Not applicable.

Availability of data and material

The scRNA-seq data from three OA and four RA samples were downloaded from GSE152805(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE152805) and

GSE200815(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE200815). The bulkRNA-seq data was downloaded from GSE89408(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE89408).

References

1. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. Lancet (London, England). 2016;388(10055):2023-38.
2. Nygaard G, Firestein GS. Restoring synovial homeostasis in rheumatoid arthritis by targeting fibroblast-like synoviocytes. Nature reviews Rheumatology. 2020;16(6):316-33.
3. Stephenson W, Donlin LT, Butler A, Rozo C, Bracken B, Rashidfarrokhi A, et al. Single-cell RNA-seq of rheumatoid arthritis synovial tissue using low-cost microfluidic instrumentation. Nature communications. 2018;9(1):791.
4. Zhang F, Wei K, Slowikowski K, Fonseka CY, Rao DA, Kelly S, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nature immunology. 2019;20(7):928-42.
5. Glisovic T, Bachorik JL, Yong J, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. FEBS letters. 2008;582(14):1977-86.

6. Hentze MW, Castello A, Schwarzl T, Preiss T. A brave new world of RNA-binding proteins. Nature reviews Molecular cell biology. 2018;19(5):327-41.
7. Lukong KE, Chang KW, Khandjian EW, Richard S. RNA-binding proteins in human genetic disease. Trends in genetics : TIG. 2008;24(8):416-25.
8. Gebauer F, Schwarzl T, Valcárcel J, Hentze MW. RNA-binding proteins in human genetic disease. Nature reviews Genetics. 2021;22(3):185-98.
9. Hashimoto S, Kishimoto T. Roles of RNA-binding proteins in immune diseases and cancer. Seminars in cancer biology. 2022;86(Pt 3):310-24.
10. Jeltsch KM, Heissmeyer V. Regulation of T cell signaling and autoimmunity by RNA-binding proteins. Current opinion in immunology. 2016;39:127-35.
11. Mino T, Takeuchi O. Post-transcriptional regulation of immune responses by RNA binding proteins. Proceedings of the Japan Academy Series B, Physical and biological sciences. 2018;94(6):248-58.
12. Yoshinaga M, Takeuchi O. RNA binding proteins in the control of autoimmune diseases. Immunological medicine. 2019;42(2):53-64.
13. Lodde V, Floris M, Zoroddu E, Zarbo IR, Idda ML. RNA-binding proteins in autoimmunity: From genetics to molecular biology. Wiley interdisciplinary reviews RNA. 2023:e1772.
14. Agarwal V, Lopez-Darwin S, Kelley DR, Shendure J. The landscape of alternative polyadenylation in single cells of the developing mouse embryo. Nature communications. 2021;12(1):5101.
15. Brannan KW, Chaim IA, Marina RJ, Yee BA, Kofman ER, Lorenz DA, et al. Robust single-cell discovery of RNA targets of RNA-binding proteins and ribosomes. Nature methods. 2021;18(5):507-19.
16. Feng H, Moakley DF, Chen S, McKenzie MG, Menon V, Zhang C. Complexity and graded regulation of neuronal cell-type-specific alternative splicing revealed by single-cell RNA sequencing. Proceedings of the National Academy of Sciences of the United States of America. 2021;118(10).
17. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature biotechnology. 2018;36(5):411-20.
18. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nature methods. 2015;12(4):357-60.
19. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology. 2014;15(12):550.
20. Cheng C, Liu L, Bao Y, Yi J, Quan W, Xue Y, et al. SUVA: splicing site usage variation analysis from RNA-seq data reveals highly conserved complex splicing biomarkers in liver cancer. RNA biology. 2021;18(sup1):157-71.
21. Ali A, Shafarin J, Abu Jabal R, Aljabi N, Hamad M, Sualeh Muhammad J, et al. Ferritin heavy chain (FTH1) exerts significant antigrowth effects in breast cancer cells by inhibiting the expression of c-MYC. FEBS open bio. 2021;11(11):3101-14.
22. Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, et al. Comprehensive Identification of RNA-Binding Domains in Human Cells. Molecular cell. 2016;63(4):696-710.

23. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nature reviews Genetics. 2014;15(12):829-45.
24. Ianevski A, Giri AK, Aittokallio T. Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. Nature communications. 2022;13(1):1246.
25. Chou CH, Jain V, Gibson J, Attarian DE, Haraden CA, Yohn CB, et al. Synovial cell cross-talk with cartilage plays a major role in the pathogenesis of osteoarthritis. Scientific reports. 2020;10(1):10868.
26. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic acids research. 2011;39(Web Server issue):W316-22.
27. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. Nature communications. 2021;12(1):1088.
28. Phipson B, Sim CB, Porrello ER, Hewitt AW, Powell J, Oshlack A. propeller: testing for differences in cell type proportions in single cell data. Bioinformatics (Oxford, England). 2022;38(20):4720-6.
29. Chen J, Cascio J, Magee JD, Techasintana P, Gubin MM, Dahm GM, et al. Posttranscriptional gene regulation of IL-17 by the RNA-binding protein HuR is required for initiation of experimental autoimmune encephalomyelitis. Journal of immunology (Baltimore, Md : 1950). 2013;191(11):5441-50.
30. Nieminen R, Vuolteenaho K, Riutta A, Kankaanranta H, van der Kraan PM, Moilanen T, et al. Aurothiomalate inhibits COX-2 expression in chondrocytes and in human cartilage possibly through its effects on COX-2 mRNA stability. European journal of pharmacology. 2008;587(1-3):309-16.
31. Yamasaki S. Recent advances in the role of RNA-binding protein, tristetraprolin, in arthritis. Immunological medicine. 2018;41(3):98-102.
32. Yang X, Chen B, Zhang M, Xu S, Shuai Z. Tristetraprolin Gene Single-Nucleotide Polymorphisms and mRNA Level in Patients With Rheumatoid Arthritis. Frontiers in pharmacology. 2021;12:728015.
33. Van Aelst LN, Voss S, Carai P, Van Leeuwen R, Vanhoutte D, Sanders-van Wijk S, et al. Osteoglycin prevents cardiac dilatation and dysfunction after myocardial infarction through infarct collagen strengthening. Circulation research. 2015;116(3):425-36.
34. Funderburgh JL, Corpuz LM, Roth MR, Funderburgh ML, Tasheva ES, Conrad GW. Mimecan, the 25-kDa corneal keratan sulfate proteoglycan, is a product of the gene producing osteoglycin. The Journal of biological chemistry. 1997;272(44):28089-95.
35. Dunlevy JR, Beales MP, Berryhill BL, Cornuet PK, Hassell JR. Expression of the keratan sulfate proteoglycans lumican, keratocan and osteoglycin/mimecan during chick corneal development. Experimental eye research. 2000;70(3):349-62.
36. Starup-Linde J, Viggers R, Handberg A. Osteoglycin and Bone-a Systematic Review. Current osteoporosis reports. 2019;17(5):250-5.

37. Rienks M, Papageorgiou A, Wouters K, Verhesen W, Leeuwen RV, Carai P, et al. A novel 72-kDa leukocyte-derived osteoglycin enhances the activation of toll-like receptor 4 and exacerbates cardiac inflammation during viral myocarditis. Cellular and molecular life sciences : CMLS. 2017;74(8):1511-25.
38. Lomnytska MI, Becker S, Hellman K, Hellström AC, Souchelnytskyi S, Mints M, et al. Diagnostic protein marker patterns in squamous cervical cancer. Proteomics Clinical applications. 2010;4(1):17-31.
39. Xu T, Zhang R, Dong M, Zhang Z, Li H, Zhan C, et al. Osteoglycin (OGN) Inhibits Cell Proliferation and Invasiveness in Breast Cancer via PI3K/Akt/mTOR Signaling Pathway. OncoTargets and therapy. 2019;12:10639-50.
40. Rau F, Freyermuth F, Fugier C, Villemin JP, Fischer MC, Jost B, et al. Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. Nature structural & molecular biology. 2011;18(7):840-5.
41. Charizanis K, Lee KY, Batra R, Goodwin M, Zhang C, Yuan Y, et al. Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron. 2012;75(3):437-50.
42. Wang ET, Cody NA, Jog S, Biancolella M, Wang TT, Treacy DJ, et al. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. Cell. 2012;150(4):710-24.
43. Nie M, Balda MS, Matter K. Stress- and Rho-activated ZO-1-associated nucleic acid binding protein binding to p21 mRNA mediates stabilization, translation, and cell survival. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(27):10897-902.
44. Cooke A, Schwarzl T, Huppertz I, Kramer G, Mantas P, Alleaume AM, et al. The RNA-Binding Protein YBX3 Controls Amino Acid Levels by Regulating SLC mRNA Abundance. Cell reports. 2019;27(11):3097-106.e5.
45. Schmidt T, Dabrowska A, Waldron JA, Hodge K, Koulouras G, Gabrielsen M, et al. eIF4A1-dependent mRNAs employ purine-rich 5'UTR sequences to activate localised eIF4A1-unwinding through eIF4A1-multimerisation to facilitate translation. Nucleic acids research. 2023;51(4):1859-79.
46. Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. Nature reviews Molecular cell biology. 2017;18(7):437-51.
47. Qin H, Ni H, Liu Y, Yuan Y, Xi T, Li X, et al. RNA-binding proteins in tumor progression. Journal of hematology & oncology. 2020;13(1):90.
48. Rybak-Wolf A, Plass M. RNA Dynamics in Alzheimer's Disease. Molecules (Basel, Switzerland). 2021;26(17).
49. Wu S, Romfo CM, Nilsen TW, Green MR. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature. 1999;402(6763):832-5.
50. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nature genetics. 2011;44(1):53-7.

51. Yip BH, Steeples V, Repapi E, Armstrong RN, Llorian M, Roy S, et al. The U2AF1S34F mutation induces lineage-specific splicing alterations in myelodysplastic syndromes. The Journal of clinical investigation. 2017;127(6):2206-21.
52. Chen C, Zhou P, Zhang Z, Liu Y. U2AF1 mutation connects DNA damage to the alternative splicing of RAD51 in lung adenocarcinomas. Clinical and experimental pharmacology & physiology. 2022;49(7):740-7.
53. Zhang Z, Will CL, Bertram K, Dybkov O, Hartmuth K, Agafonov DE, et al. Molecular architecture of the human 17S U2 snRNP. Nature. 2020;583(7815):310-3.
54. Will CL, Schneider C, MacMillan AM, Katopodis NF, Neubauer G, Wilm M, et al. A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. The EMBO journal. 2001;20(16):4536-46.
55. Perea W, Schroeder KT, Bryant AN, Greenbaum NL. Interaction between the Spliceosomal Pre-mRNA Branch Site and U2 snRNP Protein p14. Biochemistry. 2016;55(4):629-32.
56. Schellenberg MJ, Dul EL, MacMillan AM. Structural model of the p14/SF3b155 · branch duplex complex. RNA (New York, NY). 2011;17(1):155-65.
57. Yazhini A, Sandhya S, Srinivasan N. Rewards of divergence in sequences, 3-D structures and dynamics of yeast and human spliceosome SF3b complexes. Current research in structural biology. 2021;3:133-45.

Figures

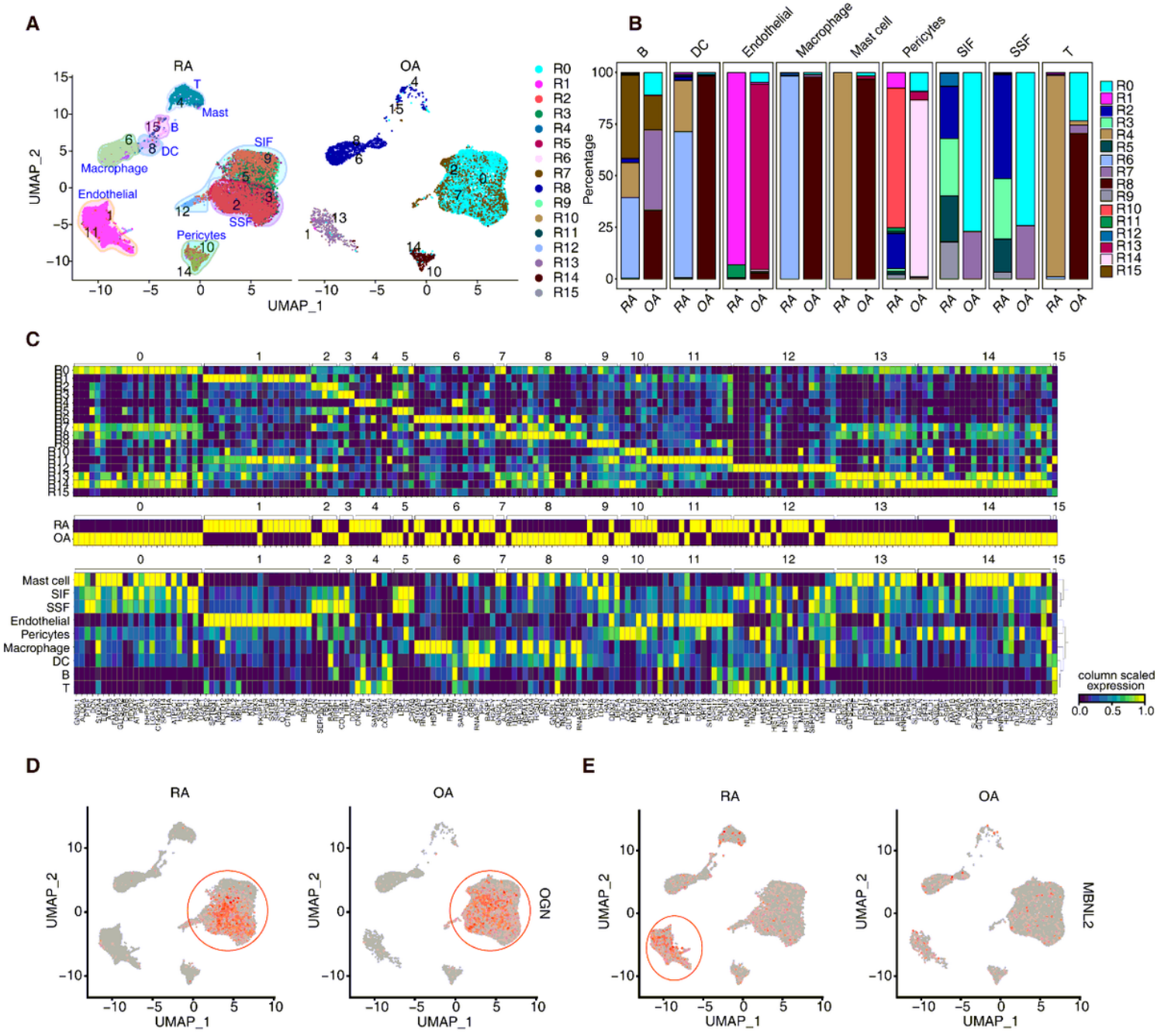

Figure 1

ScRNA-seq analysis of synoviocytes samples from patients with osteoarthritis and rheumatoid arthritis identified different cell types.

A. Schematic illustration of scRNA-seq and bulk RNA-seq data processing.

B-C. UMAP plot of composite single-cell transcriptomic profiles from all 7 samples from RA and OA. Colors indicate cell clusters along with annotations.

D. Dot plots showing the expression of representative genes annotated by cell types in each cell type.

E. Bar plot comparing the proportions of cell populations of each cell type within each sample group.The P-value were calculated in the speckle R package (version: 0.0.3). *P ≤ 0.05, **≤ 0.01, ***≤0.001.

F. Rank order based on decreasing values of the relative frequency ratio between RA and OA groups.

G. The barplot showing the number of up-regulated and down-regulated DEG and RBP in each cell type comparing RA and OA samples.

H. The boxplot depicts activation of RBPs in each cell types inferred by SCENIC-calculated AUCell scores.

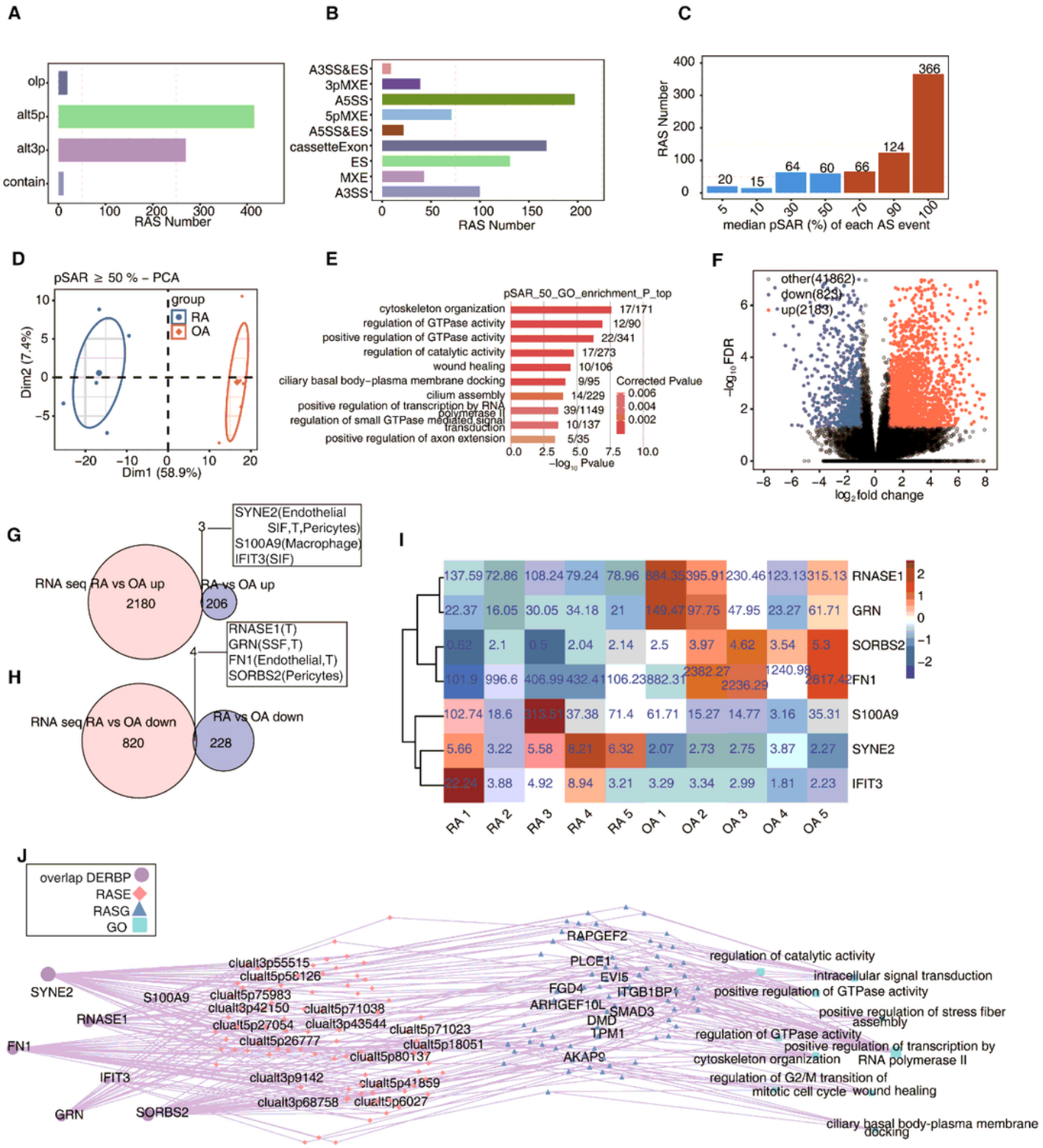

Figure 2

ScRNA-seq analysis identified heterogeneity and regulatory module of cellular specific RBP expression module.

A. UMAP plot of scRNA-seq profile. Cells are colored according to cell clusters based on RBP-Genes expression module.

B. Stacked bar plot comparing the proportions of cell populations of each RBP expression module within each sample group for different cell types.

C. Unsupervised clustering heatmap showing relative expression (z score, column scaled) levels of RBP markers of each RBP expression module in single-cell dataset according to different clinical variables containing cell types, and sample groups.

D. Gene expression level of OGN were represented in the UMAP plot split by different sample groups.

E. Gene expression level of MBNL2 were represented in the UMAP plot split by different sample groups.

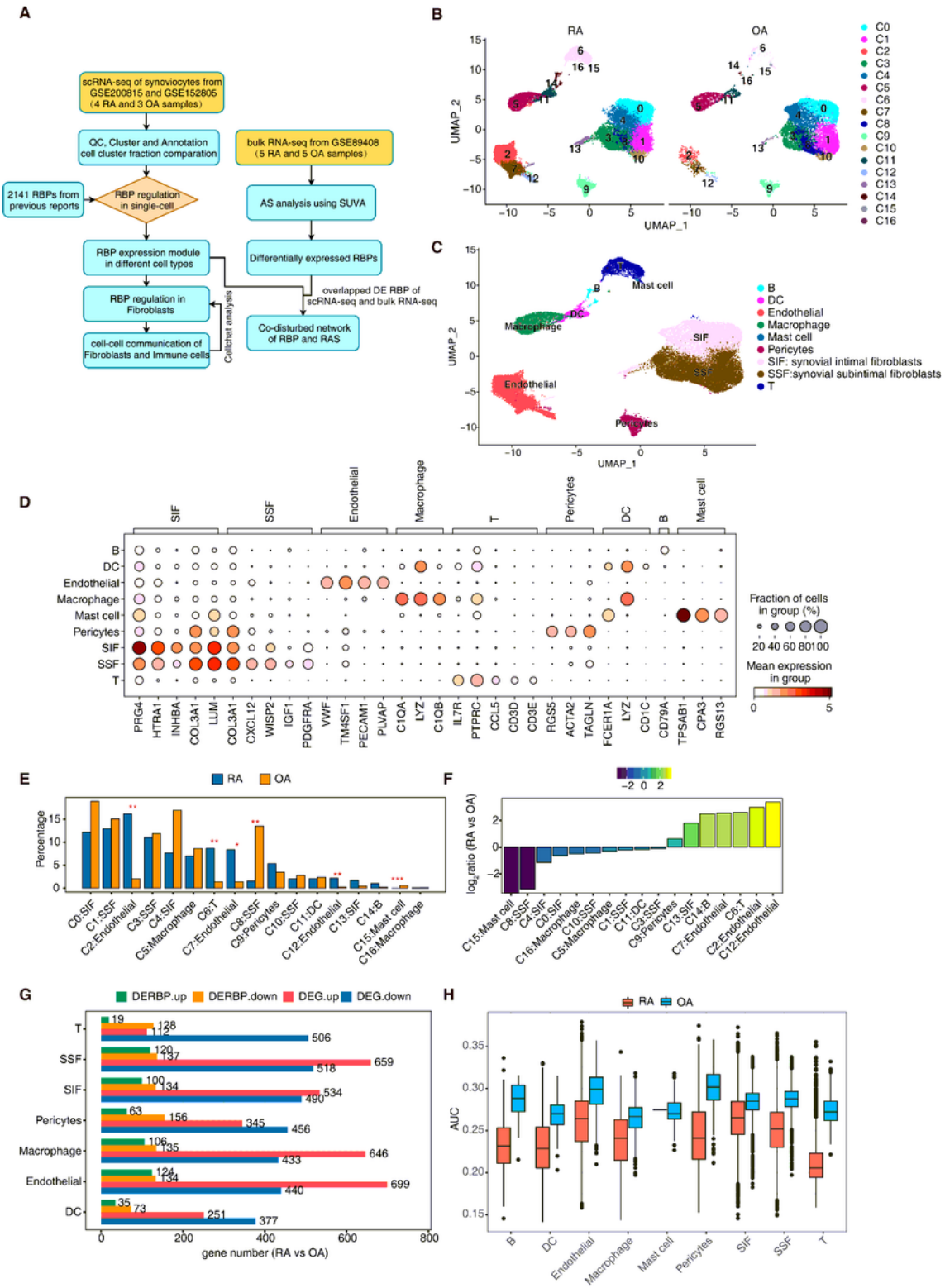

Figure 3

Functional RBPs are largely regulated in Fibroblasts cells between RA and OA samples.

A. UMAP plot of scRNA-seq profile. Fibroblasts cells are colored according to cell clusters based on RBP-Genes expression module.

B. UMAP displays the distribution of fibroblasts cells. The color indicates the RA and OA group.

C. Unsupervised clustering heatmap showing relative expression (z score, row scaled) levels of top20 up-regulated and down-regulated RBP between RA and OA.

D. Cytoscape shows the co-expression networks comprising up and down regulated DERBP in C. Edges connect DERBP-target gene pairs while nodes represent genes. DERBP are displayed in larger font size and red color. Co-expression associations of DERBP and target genes in fibroblasts were built by the “grn” algorithm from SCENIC.

E. Gene ontology enrichment analysis of biological processes of differential RBP co-expressed fibroblasts genes. Top 3 terms were selected for each cluster and heatmap shows the enrichment q-value of these terms (scaled by column).

F. Gene expression level of YBX3 were represented in the UMAP plot split by different sample groups.

G. Gene expression level of EIF4A1 were represented in the UMAP plot split by different sample groups.

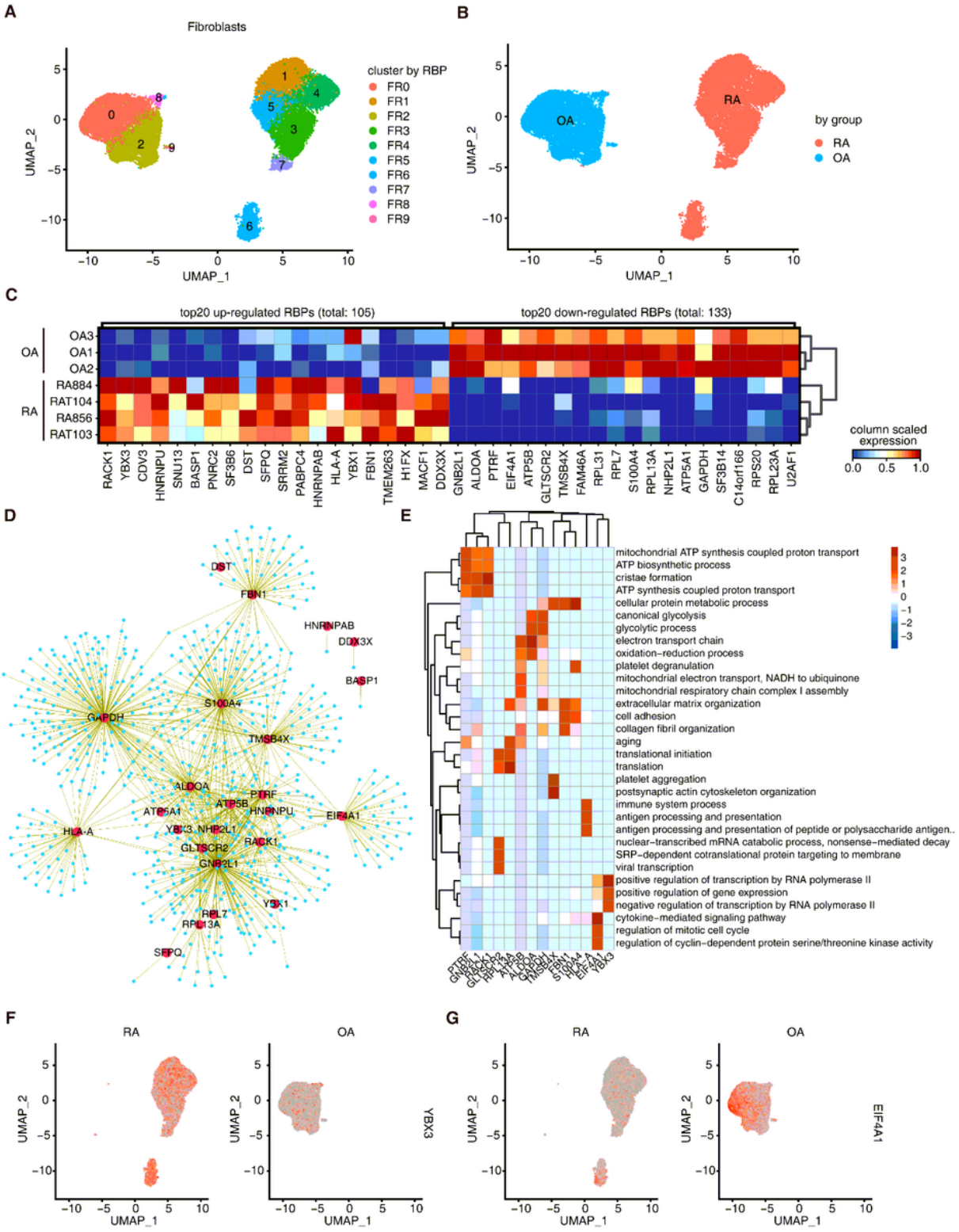

Figure 4

RBP-mediated fibroblasts subpopulations contributed to the aberrant activation of signaling pathways with immune cells during rheumatoid arthritis.

A. The number and strength of interactions in RA and OA.

B. Comparing the interaction number among RBP-Genes expression module of fibroblasts and Immune cells in RA and OA samples.

C. Scatter plot displaying up-regulated ligand-receptor pairs in RA sample group from fibroblasts cells to immune cells comparing with OA samples.

D. Cytoscape shows the co-expression networks comprising RBP. Edges connect RBP- ligand gene pairs while nodes represent genes. RBP are displayed in larger font size and red color. Co-expression associations of RBP and ligand genes were built by the “grn” algorithm from SCENIC.

E. Violin plot of CXCL12 and CXCR4 in each cell type split by different sample groups.

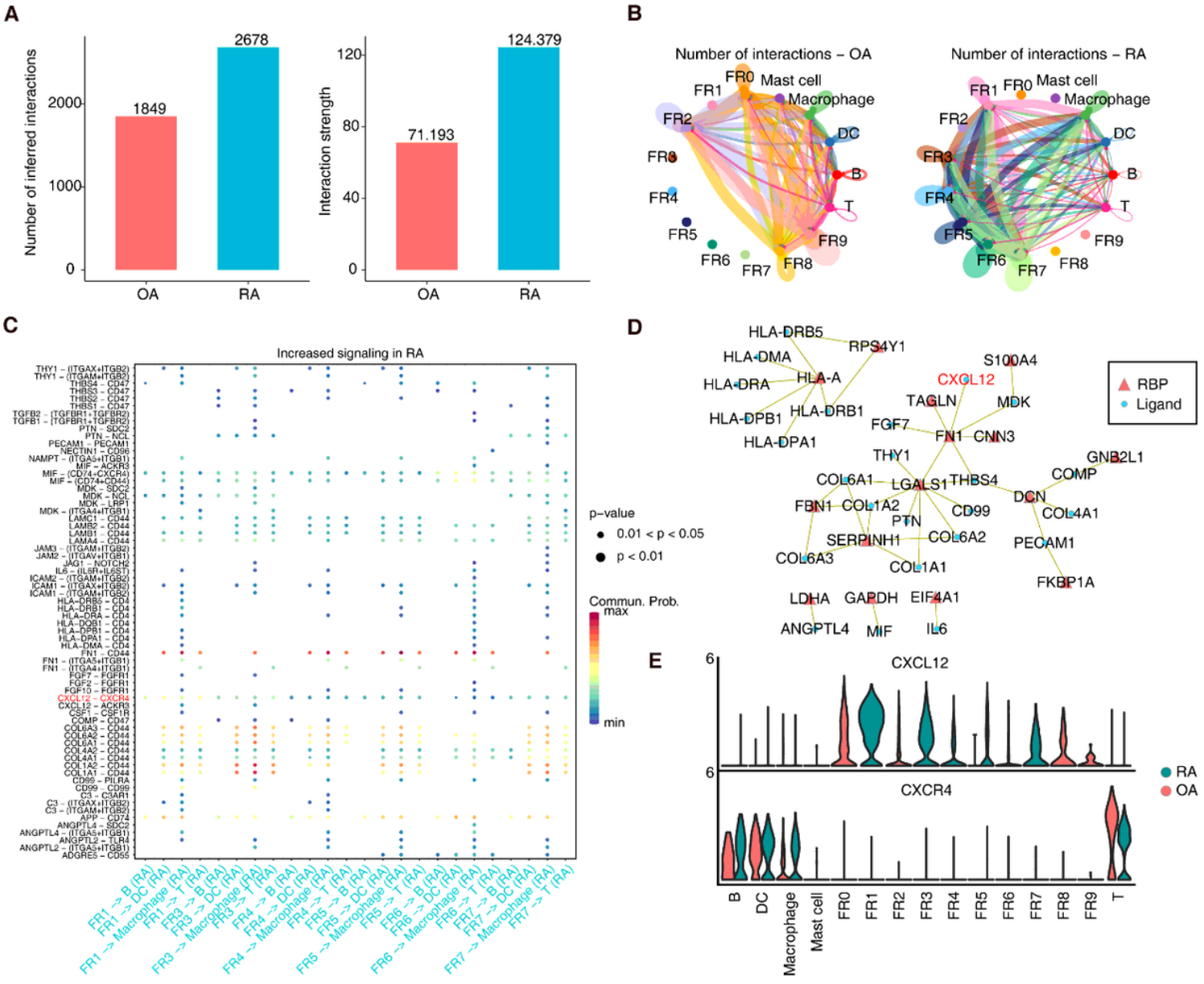

Figure 5

Identification of highly conserved RA-associated AS events co-disturbed with differentially expressed RBPs in RA and OA patients.

A. Bar plot showing number of regulated AS (RAS) detected by SUVA comparing RA with OA samples.

B. Splice junction constituting RAS events detected by SUVA was annotated to classical AS event types. And the number of each classical AS event types were showed with bar plot.

C. Bar plot showing RAS with different pSAR. RAS which pSAR (Reads proportion of SUVA AS event) ≥ 50% were labeled.

D. Principal component analysis (PCA) based on splicing ratio of RAS with pSAR ≥ 50%. The ellipse for each group is the confidence ellipse.

E. Bar plot showing the most enriched GO biological process results of RAS (PSAR ≥ 50%).

F. Volcano plots presenting all DEGs between RA and OA samples with DESeq2. FDR ≤ 0.05 and FC (fold change) ≥ 2 or ≤ 1/2.

G-H. Venn diagram showing the overlap of DE RBPs from bulk RNA-seq and ScRNA-seq dataset.

I. The heatmap diagram showing the FPKM of 7 overlapped DERBP.

J. The co-disturbed network among expression of co-regulated RBPs showed in C, and splicing ratio of RAS events (pSAR>=50%) was constructed. |Pearson's correlation| >=0.6 and pvalue <=0.01 were retained for RBP and RAS correlation. Ellipses represent RBP. Squares in around indicate RAS.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- S1.docx
- S2.docx
- S3.docx
- S4.docx
- S5.docx