

Integrated gene expression profiling reveals potential immune-related biomarkers for rheumatoid arthritis

Tong Ye (✉ 13965263207@163.com)

Suzhou Hospital of Anhui Medical University <https://orcid.org/0000-0003-1044-7984>

Zhou Haoyuan

Suzhou Hospital of Anhui Medical University

Zhao Peng

Suzhou Hospital of Anhui Medical University

Xu Haidong

General Hospital of Eastern Theatre Command

Research article

Keywords: Rheumatoid arthritis, differentially expressed genes, GEO, Immune infiltration

Posted Date: March 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-277931/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background As one of the most common causes of joint deformities, rheumatoid arthritis (RA) has seriously affected human health. In this study, RA biomarkers related to immunity were identified by bioinformatics methods.

Methods RA synovial microarray data were obtained from GEO database and preprocessed, and then differentially expressed genes (DEGs) were screened. Subsequently, the hub nodes were screened from the PPI network of DEGs by STRING database and cytoscape software. The hub nodes were verified in independent dataset. The immune infiltration in synovial tissue of RA was analyzed by CIBERSORT algorithm. In addition, the correlation between differentially infiltrated immune cells and key genes was analyzed.

Results Intersected DEGs were significantly enriched in biological processes, including immune response – activating signal transduction, B cell activation, T cell activation. Through PPI network and immune infiltration analysis, we finally identified 5 immune-related biomarkers including CCL5, CCR5, CD27, CXCL9 and TLR8.

Conclusion CCL5, CCR5, CD27, CXCL9 and TLR8 may be potential diagnostic biomarkers for RA treatment strategies.

Key Points

- Enrichment analysis of DEGs using ClusterProfiler package.
- Pearson correlation was used to analyze the relationship between key genes and immune cells.
- Immune infiltration analysis of synovium in patients with RA based on CIBERSORT algorithm.

Introduction

RA is a chronic systemic autoimmune disease with synovial inflammation and pannus formation as the main pathological changes[1,2]. With the gradually increase of inflammatory cell infiltration and bone tissue destruction, it eventually leads to joint deformity and dysfunction[3]. At present, limited by the understanding of the pathogenesis of RA, clinical treatment of RA is difficult to achieve complete or strict remission of symptoms. More often than not, remission does not last after the interruption of treatment[4]. In view of this, the research on the novel treatment scheme is the highlight of our efforts. Previous studies have shown that the cascade of innate immunity and adaptive immunity is a crucial immunopathogenic mechanism in the inflammatory process of RA[5]. Immune complex-mediated complement activation and cytokine disorder occupy a significant position in the progression of RA [6,7]. A variety of immune cells, represented by T cells, B cells, macrophages and neutrophils, participate in and modulate the autoimmune response of RA, speed up cartilage and bone erosion by disturbing the immune balance [8,9].

This study attempts to explore the pathogenesis of RA from the perspective of bioinformatics, analyze step by step and find out new immune-related biomarkers in RA.

Materials And Methods

Microarray data acquisition and preprocessing

Three RA datasets (GSE55235, GSE55457, GSE77298) were selected from the GEO database, in which the dataset GSE55235 and GSE55457 were based on the GPL570 platform, and the dataset GSE77298 was based on the GPL6947 platform. GSE55457 contains synovial tissue of rheumatoid arthritis (n = 13) and normal control tissue (n = 10), GSE55235 contains synovial tissue of rheumatoid arthritis (n = 10) and normal control tissue (n = 10), GSE77298 contains synovial tissue of rheumatoid arthritis (n = 16) and normal control tissue (n = 7). After background correction, normalization and summarization, the original expression data were annotated using their corresponding platform files.

DEGs screening

The limma package in R software was used to screen the differentially expressed genes (DEGs) from the data processed in GSE55235 and GSE55457. The Benjamini and Hochberg method were used to adjust the P value. Meanwhile, The filter criterias were set as follow: $|\log_2(\text{Fold Change})| > 1$ and the adjusted P value < 0.05 . In addition, we further screened the overlapping genes of the DEGs in GSE55235 and GSE55457, and the overlapping genes with consistent regulation (up regulated or down regulated) between the 2 datasets were employed for subsequent analysis.

Enrichment and PPI network analysis

ClusterProfiler[10] performed gene ontology (GO) and KEGG pathway analysis on the overlapping DEGs. With $P < 0.05$ as the threshold, GO items and signal pathways with significant enrichment were screened out. Besides, We used the online search for interacting genes/proteins (STRING) tool (<http://www.string-db.org>) to create an overlapping DEGs protein interaction network. Subsequently, in the Cytoscape software, the cytohubba plug-in was hired to analyze the PPI network. After the calculation, top 9 hub nodes ranked by MCC scores were identified.

Identification of hub genes

To evaluate the reliability of the top 9 hub nodes, we verified their expression levels in GSE77298, In which genes with significant differences and regulation direction consistent with overlapping DEGs were regarded as hub genes in RA.

Immune infiltration analysis

CIBERSORT algorithm is a reliable machine learning method based on linear support vector regression (SVR). It is widely used to evaluate the relative content and dynamic regulation process of 22 kinds of immune cells [23-24]. Based on GSE77298 dataset and linked CIBERSORT deconvolution method, the transcriptional characteristic matrix including 22 kinds of immune cells was simulated. The calculation times were set at 500 times. Kruskal-Wallis rank sum test was used to analyze the data with $P < 0.05$. Immune cells with significant difference in infiltration scores between RA group and control group were found by immune infiltration analysis, and then the correlation between immune cells and hub genes were analyzed by Pearson to reveal the immune-related biomarkers in RA.

Results

Identification of DEGs

First, we used Limma package to screen a total of 1185, 371 DEGs from GSE55235 and GSE55457, respectively. Subsequently, a total of 200 overlapping DEGs were selected from the DEGs of the two datasets (Fig. 1A, Fig. 1B). Finally, by comparing the logFC, of the overlapping DEGs in the two datasets, we found that the regulation direction of all the overlapping DEGs in the two datasets were surprisingly consistent.

Enrichment analyses of overlapping DEGs and PPI network construction

The results of GO enrichment analysis showed that the main biological processes (BP) of overlapping DEGs were immune response-activating signal transduction, B cell activation, T cell activation, lymphocyte mediated immunity, positive regulation of lymphocyte activation (Fig. 1C). The KEGG pathway analysis results demonstrated that overlapping DEGs were particularly enriched in cytokine-cytokine receptor interaction, chemokine signaling pathway, Th17 cell differentiation, B cell receptor signaling pathway (Fig. 1D). With the help of STRING database and cytoscape software, the PPI network containing 100 nodes and 200 edges of overlapping DEGs was constructed (Fig. 2A). Furthermore, we obtained 9 hub nodes from the PPI network through the cytohubba plug-in (Fig. 2B).

Validation of hub nodes

In order to further verify the expression level of 9 hub nodes in synovium of other RA patients, we selected GSE77298 as the test dataset. The results showed that among the 9 genes, the expression level of CCL5, CCR5, CD27, CXCL9 and TLR8 in RA group were significantly higher than that in control group, but there

was no significant difference in rest of the genes (Fig. 2C). We consider these 5 genes as the hub genes in RA.

Results of immune infiltration analysis

The CIBERSORT deconvolution method screened the synovial samples in GSE77298 with $P < 0.05$, and finally obtained 21 credible samples (Fig. 3A). The difference of immune infiltrating cells in synovium between RA and healthy control was analyzed by violin graph. The results showed that the plasma cells and M0 macrophages were significantly increased in RA group. In addition, NK cells activated, Monocytes, Dendritic cells resting and Mast cells resting were significantly decreased in RA synovium (Fig. 3B). By analyzing the correlation between the infiltration scores of the above 6 cells and the expression level of each key gene, we found that CCL5, CCR5, CD27 and CXCL9 were positively correlated with plasma cells; CCL5, CD27 and CXCL9 were negatively correlated with NK cells activated; CCR5 and TLR8 were negatively correlated with Dendritic cells resting; CCL5, CCR5 and CXCL9 were positively correlated with Mast cells resting.

Discussion

In this study, we first screened DEGs in synovium between RA and normal patients from 2 datasets (GSE55235, GSE55457). In order to optimize the reliability of the results, the DEGs, which exists in both datasets and regulates consistently were retained for follow-up analysis. Based on the previous step, we have got 30 intersecting DEGs. Interestingly, all of these genes are up-regulated. Then, with the help of GO enrichment analysis, we found that the BP significantly enriched by DEGs contained immune response-activating signal transduction, B cell activation, T cell activation, lymphocyte mediated immunity, etc. Meanwhile, KEGG enrichment analysis showed that DEGs were mainly associated with immune-related pathways containing Th17 cell differentiation and B cell receptor signaling pathway.

To uncover the vital members of these DEGs, we built the PPI network. 9 hub nodes in the PPI network have been found using cytohubba plug-in. In the GSE77298 dataset, we verified the expression of hub nodes in the synovium of RA patients, and found that the expression of CCL5, CCR5, CD27, CXCL9 and TLR8 was consistent with the previous. We regarded these 5 genes as key genes in RA. Previous studies have shown that CCL5 and CXCL9 are up-regulated in the synovium of RA patients, which is of great significance for synovial growth and inflammation[11,12]. CCR5, a G protein-coupled receptor, Silencing CCR5 can inhibit inflammation, synovial activity and promote apoptosis in RA rats by inhibiting MAPK pathway[13]. CD27 is a member of the tumor necrosis factor receptor superfamily. Recent studies have shown that CD19⁺CD24^{hi}CD27⁺B cells accumulate excessively in the synovium and might be involved in joint destruction in RA[14]. As a member of the Toll-like receptor family, TLR8 inhibitors can reduce the production of TNF in human rheumatic joint synovial culture[15].

Next, we used CIBERSORT deconvolution method to analyze the immune infiltration of the samples in GSE77298. The results showed that there were significant differences in the proportion of plasma cells, M0 macrophages, NK cells activated, Monocytes, Dendritic cells resting and Mast cells resting in RA group compared with the control group. Finally, we studied the correlation between each key gene and all differentiated infiltrating cells. The results of correlation analysis provide us with a number of information, and here we suggest interpreting it from a global perspective. As shown in figure 4, Plasma cells are the solely immune cells that are positively related to nearly all key genes, but there is no significant correlation between key genes and M0 macrophages. Meanwhile, the rest of the immune cells were negatively correlated with key genes or had no significant negative correlation trend. There are a large number of lymphocytes, plasma cells and monocytes in the synovium of chronic RA with diffuse and localized infiltration. If the inhibitor is developed with the key genes as the target, we speculate that its greatest effect is to limit or reduce the infiltration of plasma cells in the synovium, so as to improve the progression of the disease.

Abbreviations

RA: Rheumatoid arthritis; GEO: Gene Expression Omnibus; DEGs: Differentially expressed genes; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: Protein–protein interaction; BP: Biological process; STRING: Search Tool for the Retrieval of Interacting Genes; MCC: Matthews correlation coefficient

Declarations

Authors' contributions

TY and XHD designed the study and developed the retrieve strategy. TY, ZHY and ZP searched and screened the summaries and titles. TY and ZHY drafted the article. All authors read and approved the final draft.

Funding

This research was funded by the National Natural Science Foundation of China(Grant. 81501925).

Availability of data and materials

The data was downloaded from the public GEO database.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

References

1. McInnes IB, Schett G (2017) Pathogenetic insights from the treatment of rheumatoid arthritis. *Lancet* 389(10086):2328-2337
2. Falconer J, Murphy AN, Young SP, Clark AR, Tiziani S, Guma M, Buckley CD (2018) Review: synovial cell metabolism and chronic inflammation in rheumatoid arthritis. *Arthritis Rheumatol* 70(7):984-999
3. Holers VM, Demoruelle MK, Kuhn KA, Buckner JH, Robinson WH, Okamoto Y, Norris JM, Deane KD (2018) Rheumatoid arthritis and the mucosal origins hypothesis: protection turns to destruction. *Nat Rev Rheumatol* 14(9):542-557
4. Abbasi M, Mousavi MJ, Jamalzahi S, Alimohammadi R, Bezvan MH, Mohammadi H, Aslani S (2019) Strategies toward rheumatoid arthritis therapy; the old and the new. *J Cell Physiol*, 234(7):10018-10031
5. Chen JQ, Papp G, Szodoray P, Zeher M (2016) The role of microRNAs in the pathogenesis of autoimmune diseases. *Autoimmun Rev* 15(12):1171-1180
6. Coutant F, Miossec P (2020) Evolving concepts of the pathogenesis of rheumatoid arthritis with focus on the early and late stages. *Curr Opin Rheumatol* 32(1):57-63
7. Arend WP, Firestein GS (2012) Pre-rheumatoid arthritis: predisposition and transition to clinical synovitis. *Nat Rev Rheumatol* 8(10):573-586
8. Fang Q, Zhou C, Nandakumar KS (2020) Molecular and cellular pathways contributing to joint damage in rheumatoid arthritis. *Mediators Inflamm* 2020:3830212
9. Orr C, Vieira-Sousa E, Boyle DL, Buch MH, Buckley CD, Cañete JD, Catrina AI, Choy EHS, Emery P, Fearon U, et al (2017) Corrigendum: synovial tissue research: a state-of-the-art review. *Nat Rev Rheumatol* 14(1):60
10. Yu G, Wang L, Han Y, He Q (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16(5):284-28
11. Huang Y, Zheng SY, Wang R, Tang CP, Zhu JQ, Li J (2019) CCL5 and related genes might be the potential diagnostic biomarkers for the therapeutic strategies of rheumatoid arthritis. *Clin Rheumatol* 38(9):2629-2635
12. Yoshida S, Arakawa F, Higuchi F, Ishibashi Y, Goto M, Sugita Y, Nomura Y, Niino D, Shimizu K, Aoki R, Hashikawa K, Kimura Y, Yasuda K, Tashiro K, Kuhara S, Nagata K, Ohshima K (2012) Gene expression analysis of rheumatoid arthritis synovial lining regions by cDNA microarray combined with laser microdissection: up-regulation of inflammation-associated STAT1, IRF1, CXCL9, CXCL10, and CCL5. *Scand J Rheumatol* 41(3):170-9
13. Lan YY, Wang YQ, Liu Y (2019) CCR5 silencing reduces inflammatory response, inhibits viability, and promotes apoptosis of synovial cells in rat models of rheumatoid arthritis through the MAPK signaling pathway. *J Cell Physiol* 234(10):18748-18762

14. Guo XF, Xu TT, Zheng JT, Cui XJ, Li M, Wang K, Su M, Zhang HF, Zheng K, Sun CL, Song SL, Liu HJ (2020) Accumulation of synovial fluid CD19CD24CD27 B cells was associated with bone destruction in rheumatoid arthritis. *Sci Rep* 10(1):14386
15. Sacre SM, Lo A, Gregory B, Simmonds RE, Williams L, Feldmann M, Brennan FM, Foxwell BM (2008) Inhibitors of TLR8 reduce TNF production from human rheumatoid synovial membrane cultures. *J Immunol* 181(11):8002-9

Figures

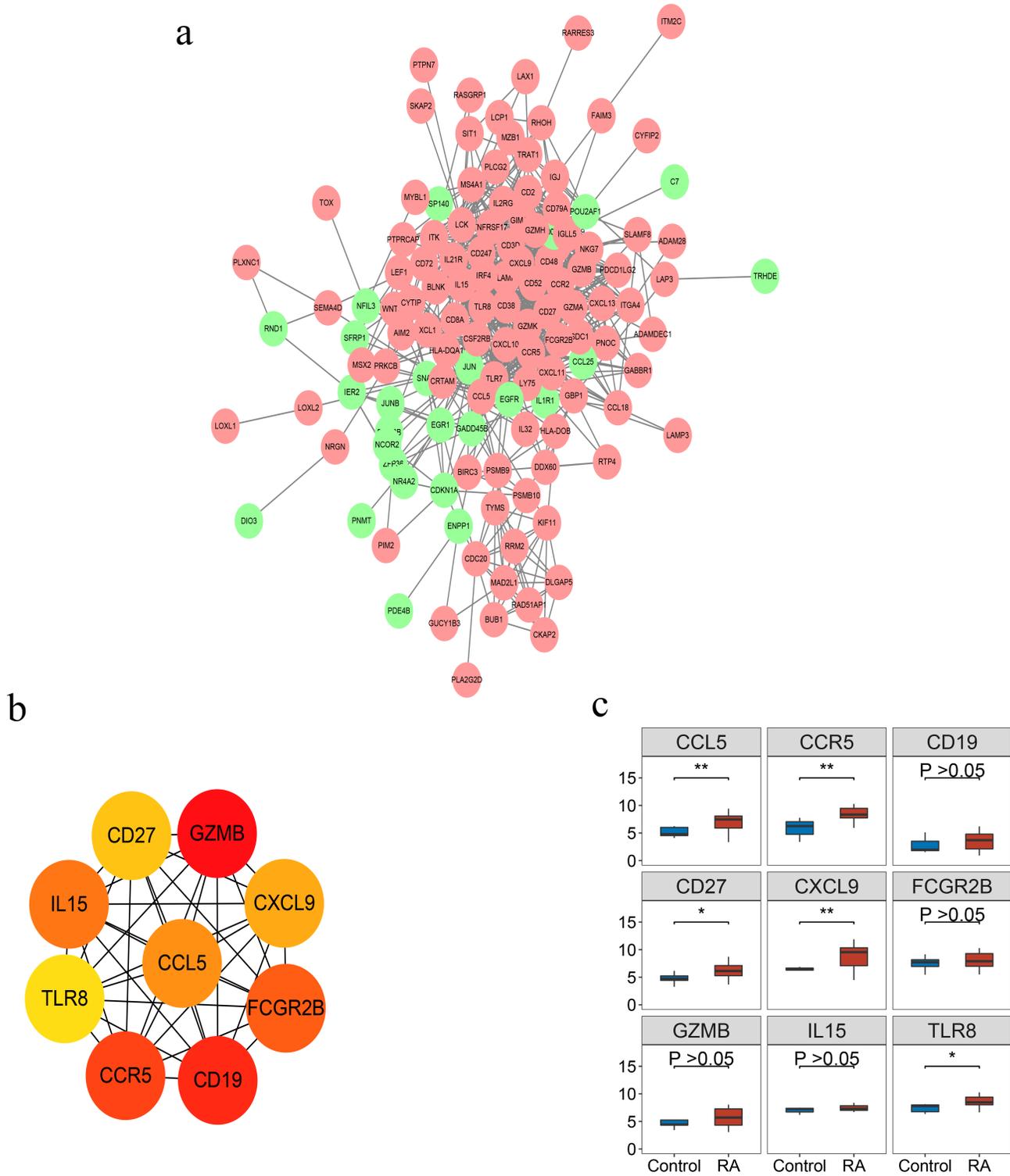


Figure 2

(a) PPI network of intersection DEGs. (b) 9 hub nodes in the PPI network. (c) Verification of hub node expression in GSE77298

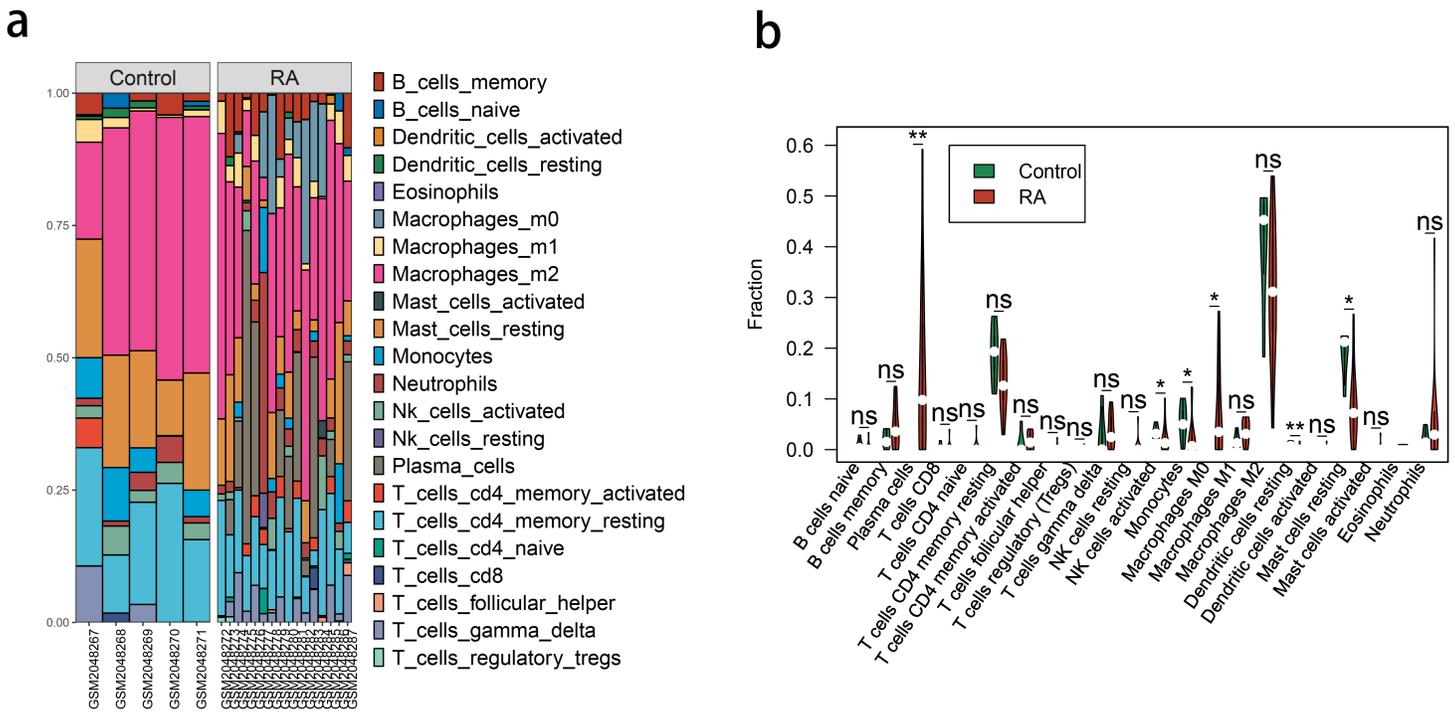


Figure 3

(a-b) Analysis of immune infiltration in synovial tissues of RA patients.

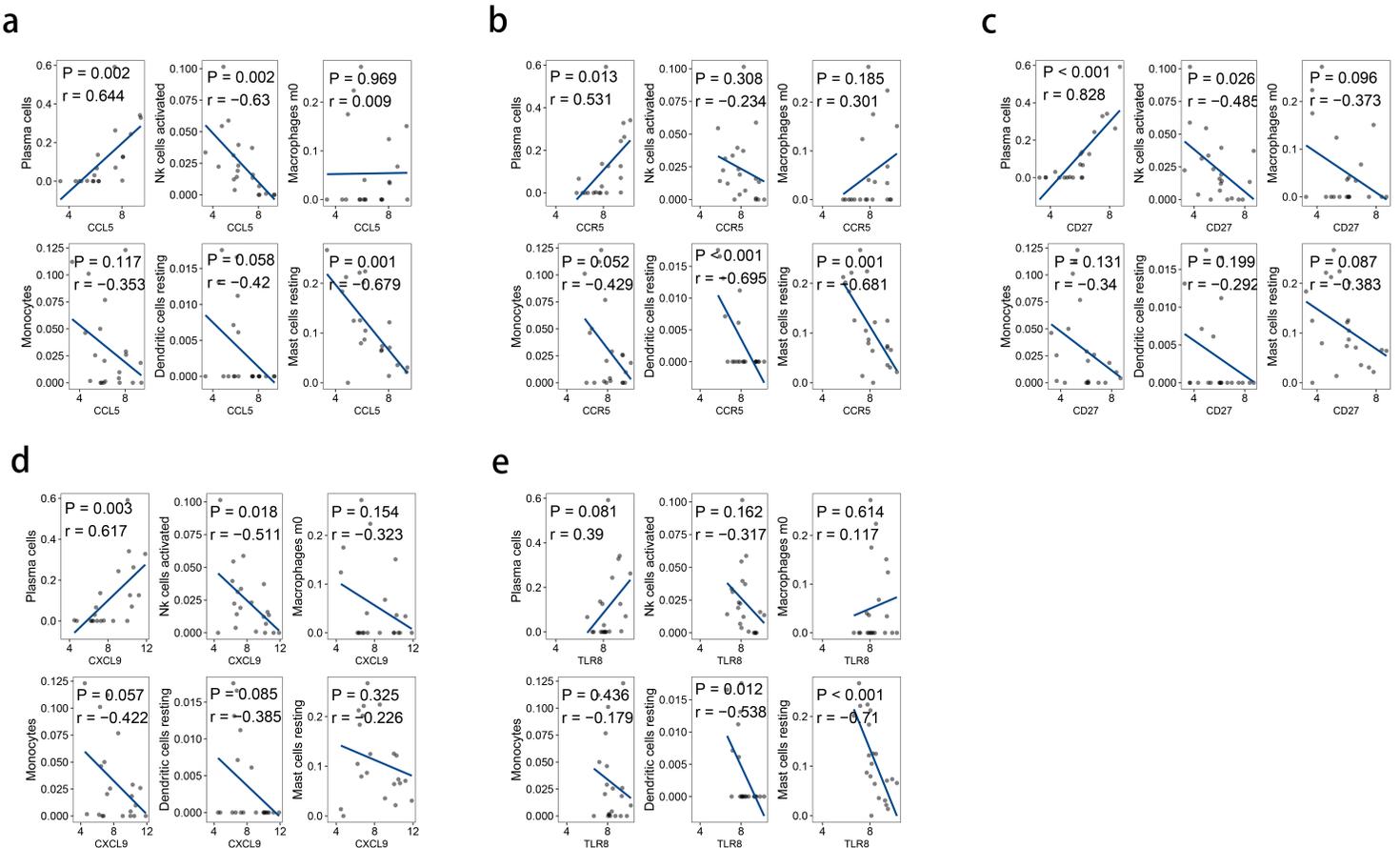


Figure 4

(a-e) Pearson correlation analysis of key genes and immune cells.