Original Article

Drug discovery in rheumatoid arthritis with joint effusion identified by text mining and biomedical databases

Zhuo Li[1,2], Jun Fu[2], Yalei Cao[3,4], Chi Xu[2], Xinli Han[1,2], Wupeng Zhang[1,2], Zelong Song[1,2], Jiying Chen[2]

[1]School of Medicine, Nankai University, Tianjin, China; [2]Department of Orthopedic Surgery, Chinese PLA General Hospital, Beijing, China; [3]Department of Urology, Peking University Third Hospital, Beijing, China; [4]Department of Andrology, Peking University Third Hospital, Beijing, China

Contributions: (I) Conception and design: J Chen, Z Li; (II) Administrative support: J Chen; (III) Provision of study materials or patients: Z Li, X Han, W Zhang, Z Song; (IV) Collection and assembly of data: Z Li, J Fu, Y Cao; (V) Data analysis and interpretation: Z Li, C Xu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jiying Chen, MD. Department of Orthopedic Surgery, Chinese PLA General Hospital, Beijing, China.
Email: jiyingchen_301@163.com.

Background: Rheumatoid arthritis is a long-term systemic disease that primarily affects multiple synovial joints throughout the body. Some patients with severe joint effusion even require repeated arthrocentesis or arthroscopic debridement to relieve symptoms, which causes them much suffering mentally and physically. This text-mining study was designed to find potential drugs that target key genes in this disease.

Methods: Firstly, we performed text mining by two keywords ("rheumatoid synovitis" and "joint effusion") to get a common set of genes. Secondly, Gene Ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analysis performed on these genes, and protein-protein interaction (PPI) network was constructed. Subsequently, the significant genes clustered in the PPI network were chose to execute gene-drug interaction analysis for potential drug discovery.

Results: Through text mining, 68 overlapping genes were identified as an initial set of key genes. Construction of the initial gene set's PPI network showed that 25 genes clustered in a significant gene module. Ultimately, 8 out of 25 genes could be targetable by a total of 19 drugs.

Conclusions: The final 8 genes (*PTGS2*, *TNF*, *VEGFA*, *IL1B*, *CCL2*, *VWF*, *IL6*, and *ESR1*) and 19 drugs may provide significant therapeutic value for rheumatoid arthritis patients with joint effusion.

Keywords: Rheumatoid arthritis; joint effusion; drug discovery; text mining; bioinformatics analysis

Submitted Jan 01, 2021. Accepted for publication Mar 23, 2021.
doi: 10.21037/apm-20-2631b
View this article at: http://dx.doi.org/10.21037/apm-20-2631b

Introduction

Rheumatoid arthritis (RA) is a long-term systemic disease that primarily affects multiple synovial joints throughout the body. The prevalence of RA is approximately similar between different regions globally, ranging from 0.5% to 1% in adults (1-3). RA most often strikes in middle age, and is 2.5 times more common in women than in men (4,5). It usually causes joint pain, stiffness, swelling, or joint effusion, accompanied by extra-articular manifestations such as anemia, respiratory, or nervous system lesions. RA was responsible for approximately 38,000 deaths in 2013, placing a massive health management burden on various countries (6-8). The exact causes of RA are unclear today. Still, research has shown that both genetic [such as shared epitope (SE) of HLA-DR] and environmental (such as infection, tobacco) factors play a significant role (9,10). The primary pathogenic mechanism is the human immune system attacking the joints, resulting in inflammation and hyperplasia of the joint capsule, which usually affects bone and cartilage. A variety of cells, including T lymphocytes, B lymphocytes, macrophage-like synoviocytes (MLS),

and fibroblast-like synoviocytes (FLS) exhibit abnormal behavior that induces the production of multiple pro-inflammatory factors. And multiple inflammatory cascade responses, including overexpression of tumor necrosis factor, are involved in the development of rheumatoid arthritis. Typically, the inflammatory response mediated by TNF and IL6 can lead to severe joint destruction, making these two genes two key therapeutic targets (11,12).

In the management of RA, a crucial principle is that disease-modifying treatment should be given early after diagnosis to delay the progression of this disease or even achieve complete remission (13). NSAIDs are no longer first-line drugs because of their adverse effect profile and their inability to improve long-term prognosis of the disease. And early administration of no more than 10 mg of prednisone equivalent glucocorticosteroids has been shown to slow radiographic progression within six months (14). Overall, DMARDs (such as methotrexate) and the burgeoning biologic agents (such as TNF inhibitors) have their own pharmacological characteristics, mechanisms of action, and clinical indications (15,16). However, there are few studies on which drugs should be chosen in RA patients with severe joint effusion. Some patients even require repeated arthrocentesis or arthroscopic debridement to relieve symptoms, which causes physical and mental suffering.

Although drug development has a standard and rigorous process, it is sometimes more efficient to explore new indications for existing drugs, which can significantly reduce costs (17). One example is the successful treatment of erectile dysfunction with sildenafil. The rapid development of bioinformatics has brought new progress in the field of biomedicine. Based on the combination of text mining and data analysis, it will be more efficient to identify gene targets or signal pathways and to explore further the treatment model of the disease. Particularly in drug discovery, text mining can provide compelling evidence from a new perspective (18-20).

In our study, we carried out literature mining through keywords retrieval so as to preliminarily screen for the target genes. Next, we performed enrichment analyses of the target genes and further generated a set of relevant genes with a higher interconnection level by constructing protein-protein interaction (PPI) networks. With in-depth analysis of drug-gene interaction, a list of potential candidate drugs that specifically act on one of the focused genes was identified. Compared with bioinformatics in the cancer field, there are fewer researchers working on rheumatoid arthritis via text mining and data analysis. In summary, we explored potential drugs for the RA patients associated with severe joint effusion, which provided a brand new perspective and theoretical basis for the prevention and treatment of rheumatoid arthritis. We present the following article in accordance with the MDAR checklist (available at http://dx.doi.org/10.21037/apm-20-2631b).

Methods

Data acquisition

The open-access database pubmed2ensembl (http://pubmed2ensembl.ls.manchester.ac.uk) was used as a tool to mine associated genes. Based on existing biological literature, pubmed2ensembl can filter and extract the genes associated with the keywords entered (21). We selected "rheumatoid synovitis" (RS) and "joint effusion" (JE) as keywords respectively for text mining. Then unique gene symbols were extracted from each result. The overlapping genes were the intersection of two gene sets, and then for the next further analysis.

Functional and signal pathway enrichment analysis

The Gene Ontology (GO) analysis can analyze and annotate genes in terms of cellular components, biological processes, and molecular function. The Kyoto Encyclopedia of Genes and Genomes (KEGG) (22) was created in 1995 by the Kanehisa laboratory to provide insight into advanced functional and biological systems using genome sequencing and other high-throughput experimental techniques. The web-based enrichment analysis tool (https://david.ncifcrf.gov/) was selected for GO and KEGG enrichment analysis of the gene set corresponding to the RS and JE intersection.

PPI analysis and gene module analysis

PPI networks consist of individual proteins that interact with each other to participate in various aspects of life processes such as biological signaling, regulation of gene expression, energy and material metabolism, and cell cycle. The well-known online analysis tool STRING (23) (http://string-db.org/) was used to analyze the PPI network for the selected genes. The STRING (Version 11) covers more than five thousand species and contains 24.6 million proteins and about three billion PPIs. Interactions in STRING derived from five primary sources: genomic

context predictions, high-throughput lab experiments, co-expression, automated text mining and previous knowledge in databases (24). As the input set, all overlapping gene symbols were uploaded into STRING and analyzed through "Multiple Protein" mode. In order to obtain more detailed data for drug mining, we set the interaction score threshold at 0.400. Besides, we downloaded the result file in TSV format and rebuilt the PPI network of target genes via Cytoscape (Version 3.7.1) (25). The STRING app in Cytoscape was used to visualize and embellish the PPI network. MCODE, an app that can calculate topological parameters of every node in the PPI network, was used to classify the significant nodes and to select highly relevant gene modules. The Degree Cutoff and the Node Score Cutoff were set at 2 and 0.2, respectively. Subsequently, the genes in highly interconnected gene modules were chosen to performed final drug discovery.

Drug-gene interaction and functional analysis of potential genes

DGIdb (http://www.dgidb.org) (26) contains over 40,000 genes and 10,000 drugs involved in over 100,000 drug-gene interactions or belonging to one of 42 potentially druggable gene categories. We uploaded module genes into the DGIdb database (Version 3.0) to search for drugs or compounds that may have therapeutic value. Besides, all drugs were approved by the FDA. Subsequently, the genes which have precise interactions with specific drugs were selected and analyzed for functional enrichment as target genes.

We just re-analyzed the open accessed datasets, and no ethical approval and informed consent were required by the local ethics committees. The study conformed to the provisions of the Declaration of Helsinki (as revised in 2013).

Statistical analysis

Fisher's Exact test was used to analyzed GO and KEGG enrichments via online tool DAVID or R software (version 3.6.3). $P<0.05$ was considered statistically significant.

Results

Data acquisition

The flow chart of our study was shown in *Figure 1*. Based on the text mining strategy described method section, 107 genes were related to "rheumatoid synovitis", and 410 genes were related to "joint effusion" . Besides, there were 68 overlapping genes of two gene sets.

Functional and signal pathway enrichment analysis

The online tool DAVID was used to perform enrichment analysis of the target genes and to visualize the data. *Figure 2* and *Table 1* showed the top six items most significantly enriched in each category (biological process, BP; cellular component, CC; or molecular function, MF) during the functional enrichment analysis. In the BP category, inflammation response, regulation of inflammation response, acute inflammation response, regulation of response to stimulus or external stimulus, regulation of defense response was most dramatically enriched. In the enrichment of CC category, extracellular space, extracellular region or region part, secretory granule lumen, secretory granule was significantly enhanced. As for the MF section, cytokine receptor binding, cytokine activity, receptor binding, growth factor activity, protease binding showed a higher degree of enrichment. *Figure 3* showed the 20 most prominent pathways in the KEGG pathway analysis. The five pathways in which the 68 genes were most significantly involved were cytokine-cytokine receptor interaction, rheumatoid arthritis, IL-17 signaling pathway, TNF signaling pathway, and malaria.

PPI analysis and gene module analysis

The original data was processed by STRING database, Cytoscape software, STRING app, and MCODE app sequentially. Then we obtained a PPI network with 60 nodes and 542 edges (*Figure 4A*). There were eight genes that were not involved in constructing the PPI network, so they were excluded from the gene set. According to the aforementioned screening method, we got a gene module with 25 nodes, 253 edges (*Figure 4B*). Twenty-five genes with high levels of interaction were as follows: *ELANE*, *CCL2*, *CCL4*, *CCL5*, *PTGS2*, *TNFRSF11B*, *CRP*, *IL1B*, *MMP3*, *MMP1*, *VWF*, *IL10*, *B2M*, *TIMP1*, *TIMP2*, *TNF*, *CSF2*, *CXCL8*, *ALB*, *CD40LG*, *IL6*, *IL4*, *VEGFA*, *ESR1*, *BDNF*.

Drug-gene interaction and functional analysis of potential genes

Through querying the DGIdb database, we found that

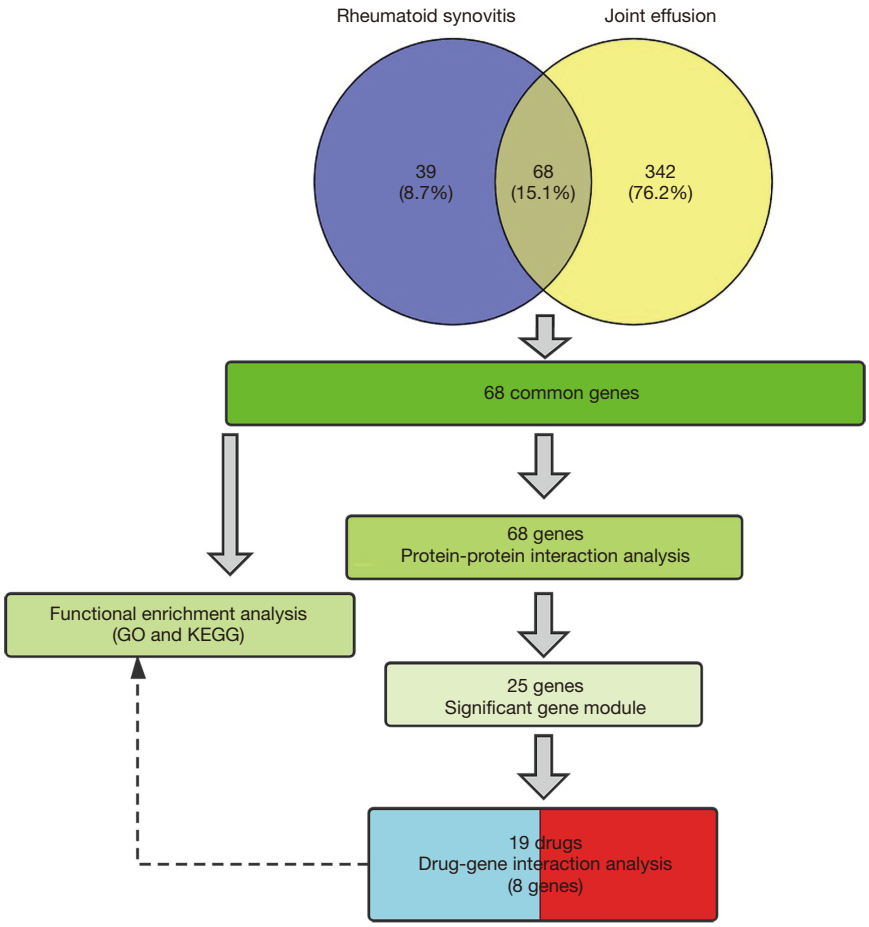

Figure 1 Overall data mining procedure. Text mining was used to identify genes associated with rheumatoid synovitis and joint effusion by using pubmed2ensemble. Extracted 68 common genes were then analyzed for GO/KEGG enrichment by using DAVID. Further, the data were sequentially processed by the STRING database, Cytoscape software, STRING app, and MCODE app to obtain a gene module containing 25 genes. The final drug list (19 drugs) was obtained by gene-drug interaction analysis using the DGIdb database. GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

8 out of 25 genes could correspond to 19 specific drugs (*Figure 5A*, *Table 2*). In our study, the potential gene targets of drugs were *PTGS2* (6 drugs), *TNF* (4 drugs), *VEGFA* (3 drugs), *IL1B* (2 drugs), *CCL2*, *VWF*, *IL6*, and *ESR1* (1 drug each). With four kinds of drug-gene interaction types, there were many suitable types of these drugs, such as antineoplastic agent, anti-inflammatory agent, immunosuppressive agents, agents for age-related macular degeneration (AMD), and fertility agents. Subsequently, we performed functional enrichment analysis of the final eight genes. The top 10 items in each category that meet the statistical criteria and have a high degree of enrichment were presented (*Figure 5B*, *Table 3*). The most significant terms were positive regulation of nitric oxide biosynthetic process (BP, P=1.34E-09), extracellular space (CC, P=0.00265906), and cytokine activity (MF, P=1.39E-06).

Discussion

Rheumatoid arthritis is one of the most common chronic autoimmune diseases worldwide today, and it causes immeasurable physical and psychological damage to the patients. Through a genome-wide association study meta-analysis, Okada *et al.* identified 42 novel risk loci in RA patients, which may reveal the underlying genes and pathways that contribute to the pathogenesis of RA (27).

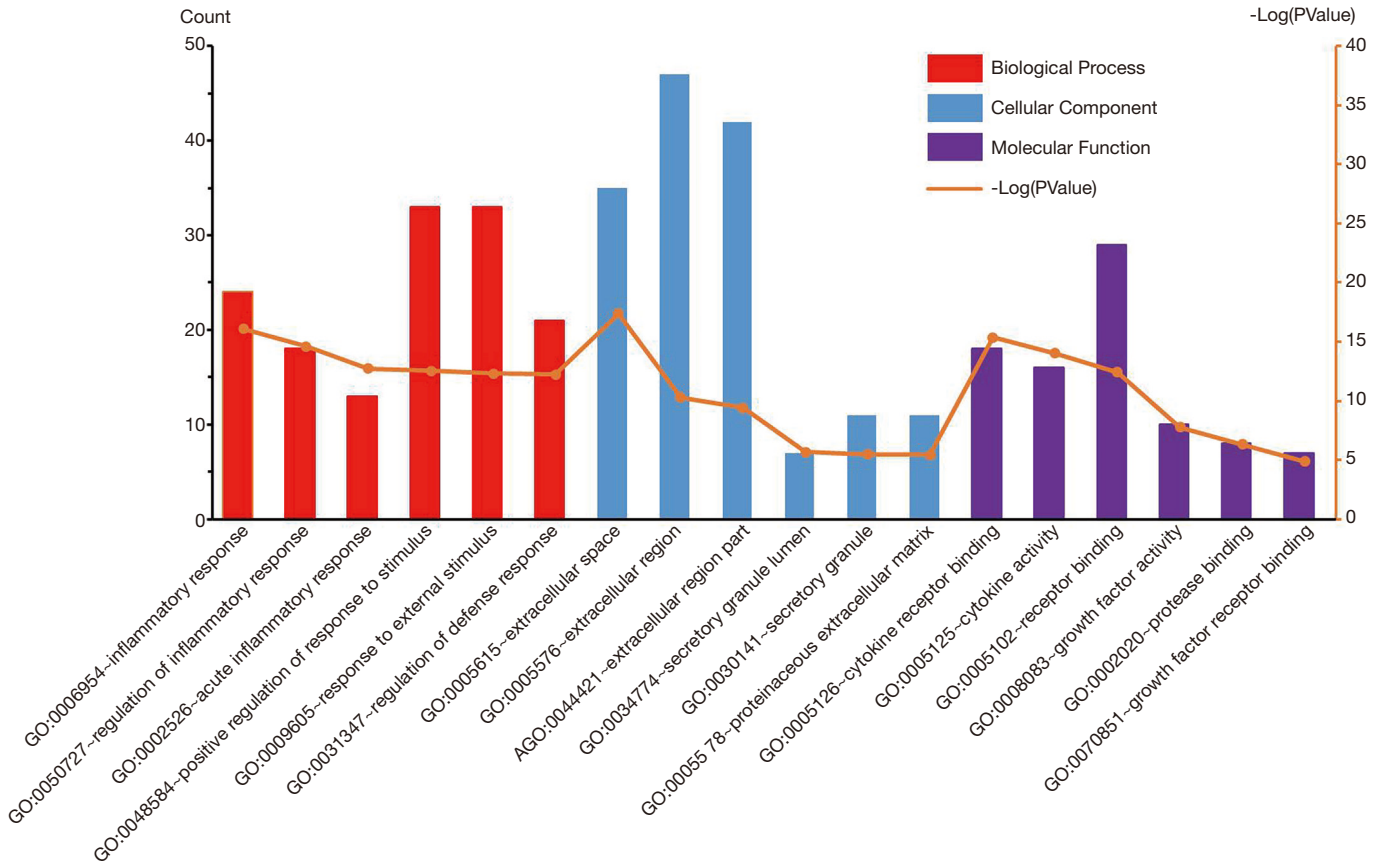

Figure 2 The top six significant GO terms of common genes. The bar charts represent the counts of genes classified in the BP, CC and MF respectively; the orange line chart represents the significance of enrichment terms. The horizontal coordinate represents each GO term, the left vertical coordinate represents the number of enriched genes, and the right vertical coordinate represents –log(P value). GO, gene ontology; BP, biological process; CC, cellular component; MF, molecular function.

Table 1 The gene ontology enrichment of the 68 common genes

Category	ID	Term	Count	P value
BP_FAT	GO:0006954	Inflammatory response	24	8.01E-17
BP_FAT	GO:0050727	Regulation of inflammatory response	18	2.57E-15
BP_FAT	GO:0002526	Acute inflammatory response	13	1.79E-13
BP_FAT	GO:0048584	Positive regulation of response to stimulus	33	2.78E-13
BP_FAT	GO:0009605	Response to external stimulus	33	4.98E-13
BP_FAT	GO:0031347	Regulation of defense response	21	5.71E-13
CC_FAT	GO:0005615	Extracellular space	35	4.29E-18
CC_FAT	GO:0005576	Extracellular region	47	5.34E-11
CC_FAT	GO:0044421	Extracellular region part	42	3.98E-10
CC_FAT	GO:0034774	Secretory granule lumen	7	2.05E-06

Table 1 (*continued*)

Table 1 (*continued*)

Category	ID	Term	Count	P value
CC_FAT	GO:0030141	Secretory granule	11	3.38E-06
CC_FAT	GO:0005578	Proteinaceous extracellular matrix	11	3.64E-06
MF_FAT	GO:0005126	Cytokine receptor binding	18	4.41E-16
MF_FAT	GO:0005125	Cytokine activity	16	9.20E-15
MF_FAT	GO:0005102	Receptor binding	29	3.78E-13
MF_FAT	GO:0008083	Growth factor activity	10	1.76E-08
MF_FAT	GO:0002020	Protease binding	8	4.59E-07
MF_FAT	GO:0070851	Growth factor receptor binding	7	1.31E-05

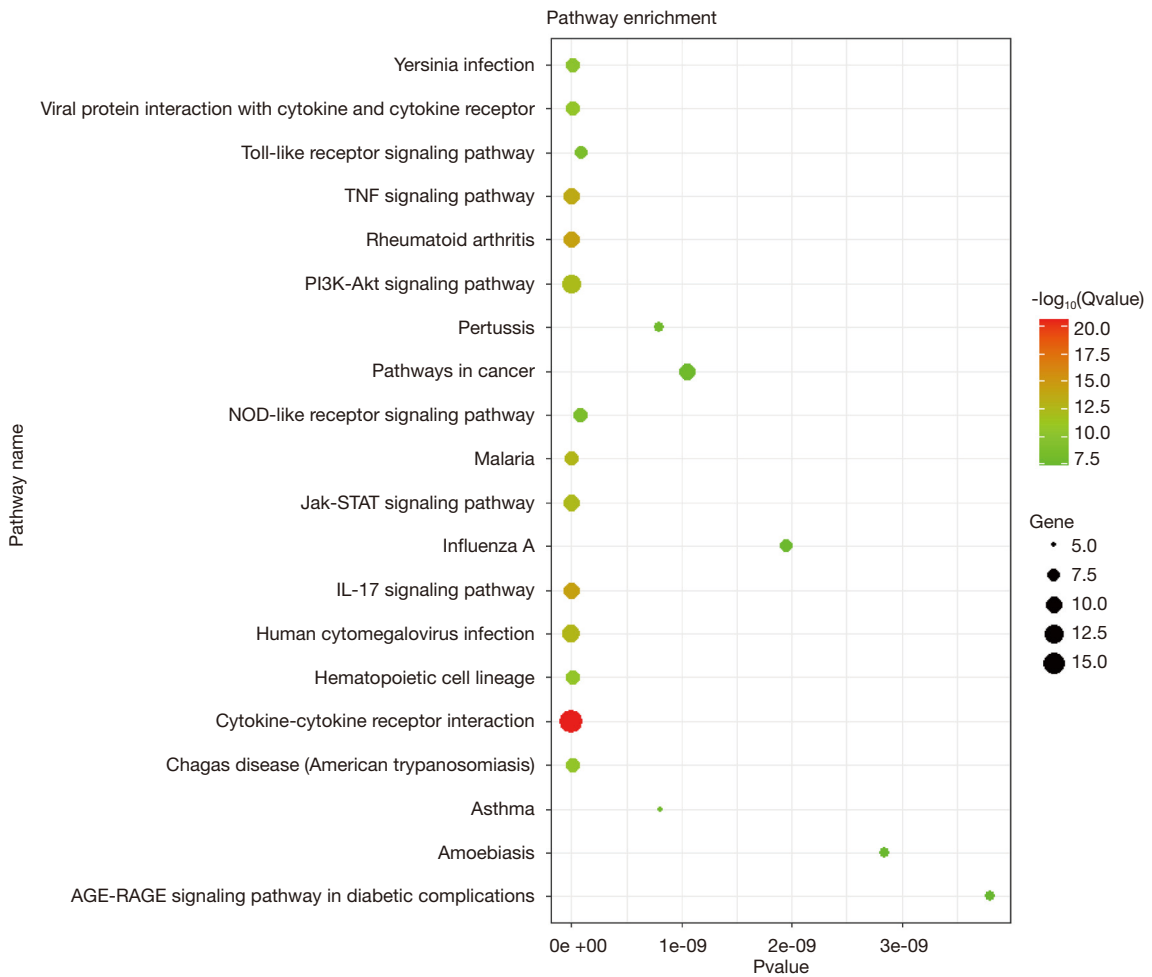

Figure 3 Scatter plot of enriched KEGG pathways statistics. Q-value is corrected P-value ranging from 2.18E-21 to 4.00E-08. The color and size of the dots represent the range of the Q-value and the common genes mapped to the indicated pathways, respectively. Top 20 enriched pathways are shown in the figure. KEGG, Kyoto Encyclopedia of Genes and Genomes.

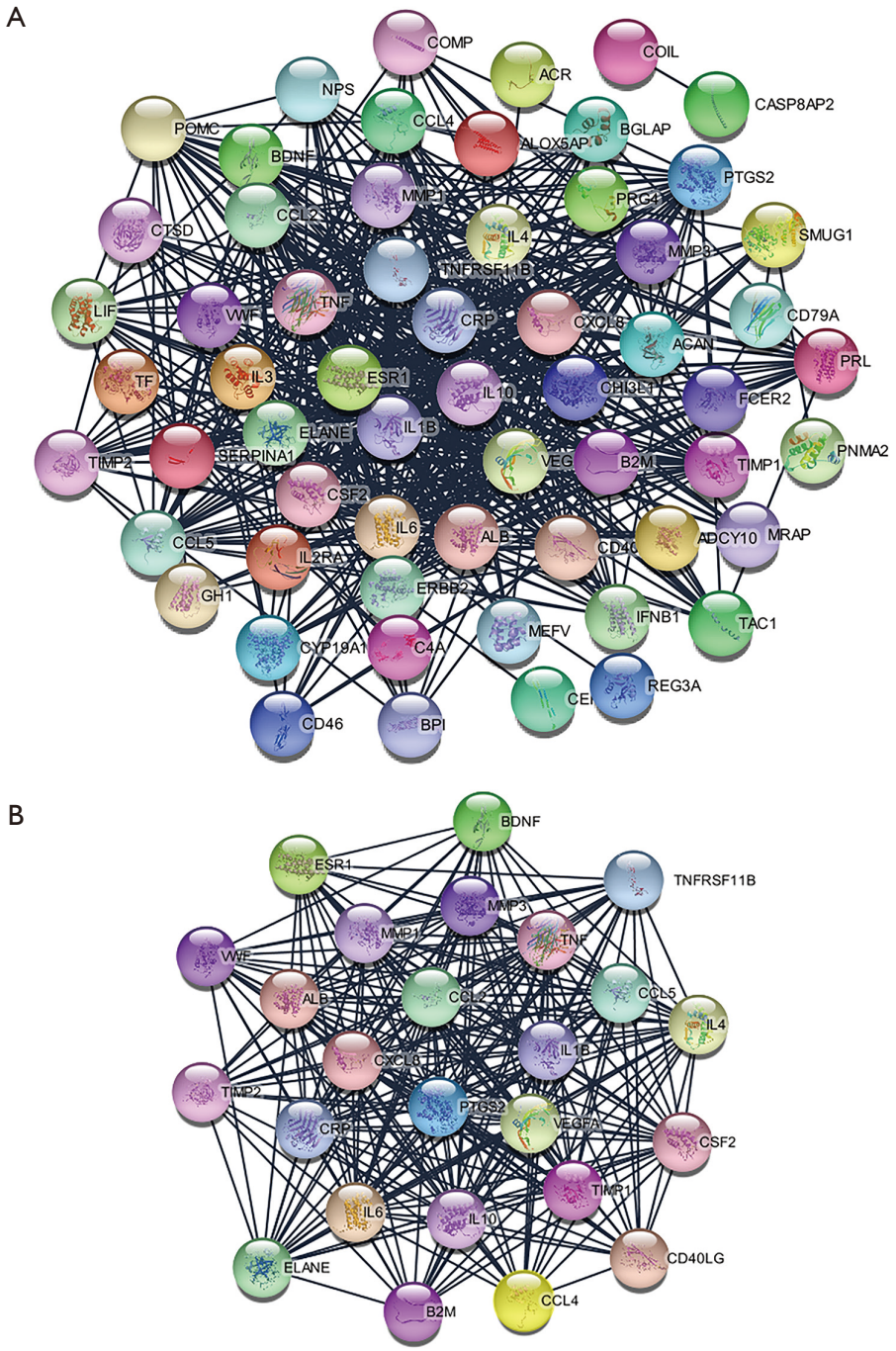

Figure 4 The PPI networks construction and significant gene module analysis. (A) The entire PPI networks of 68 common genes; (B) the significant gene module containing 25 genes. PPI, protein-protein interaction.

However, they did not identify specific drugs of potential therapeutic value. We used a bioinformatics approach to identify 19 drugs that may have improved outcomes in RA patients with severe joint effusion. These 19 drugs target the following eight genes: *PTGS2* (6 drugs), *TNF* (4 drugs), *VEGFA* (3 drugs), *IL1B* (2 drugs), *CCL2*, *VWF*, *IL6*, and *ESR1* (1 drug each). To some extent, these genes may play an essential role in the pathogenesis of RA.

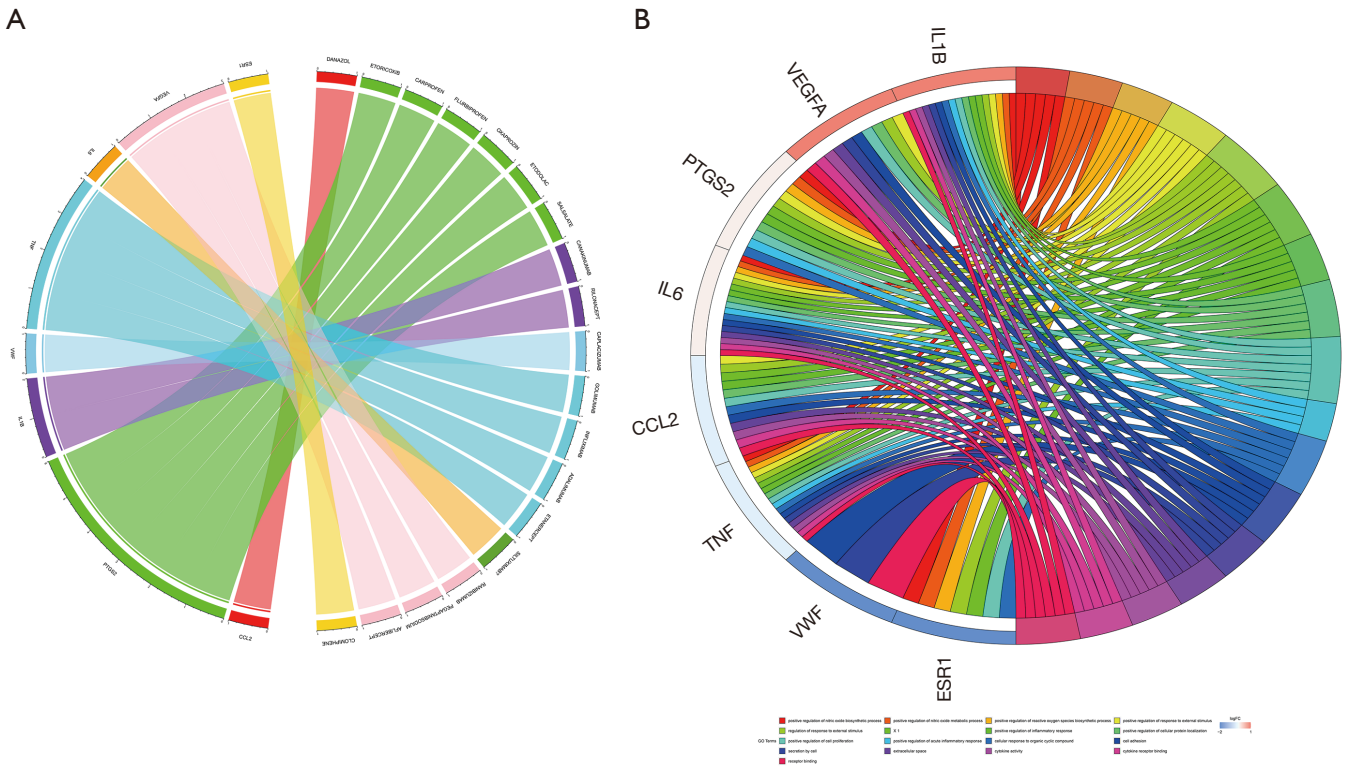

Figure 5 The drugs targeted to genes and functional enrichment analysis of final genes. (A) Chord plot for the connection between 19 drugs and 8 genes; (B) chord plot for functional enrichments of the final 8 genes.

Prostaglandin-endoperoxide synthase 2 (*PTGS2*), also known as *COX2*, can significantly promote the production of prostaglandins in synovial tissues of RA patients (28). MicroRNA-101-3p can act as an inhibitor to target *COX2* and reduce the proliferation of FLS cells in a mouse model of RA, thereby alleviating joint effusion (29). In short-term studies, tumor necrosis factor (*TNF*) inhibitors have shown better efficacy in patients with advanced disease (3). But a meta-analysis of 1,264 patients in 16 randomized controlled trials showed that about 47 percent of patients had a relapse after discontinuation of the TNF-α inhibitor (30). In Mexican women, the lower risk of RA may be related to vascular endothelial growth factor A (*VEGFA*) polymorphism, and among the elderly in China, the *VEGFA* rs699947 CA genotype is associated with a lower risk of RA, too (31,32). Using single-cell sequencing, the team of Zhang *et al.* revealed the presence of a large number of IL1B+ pro-inflammatory monocytes in the synovial tissue of RA patients, which may be one of the driving cell populations of joint inflammation (33). Chemokine (C-C motif) ligand 2 (*CCL2*), also known as *MCP1*, plays a significant role in inducing monocyte infiltration and recruitment of peripheral leukocytes in synovial tissue of rheumatoid arthritis (34,35). By promoting the proliferation and invasive capacity of pro-inflammatory factors such as *CCL2* and *IL6*, RA patients' fibroblast-like synoviocytes can be highly sensitive to C-reactive protein (36). IL6-induced WNT5A activates canonical WNT signaling for the autocrine proliferation of human synovial fibroblasts in RA (37). Besides, a 24-week randomized controlled trial showed a better clinical benefit [achieving American College of Rheumatology (ACR) 20 response] of IL6 receptor antibodies over methotrexate (monotherapy) (38). The increase of von Willebrand factor (*VWF*) activity may be related to the increased risk of atherosclerosis in RA patients without cardiovascular risk factors (39,40). Estrogen receptor 1 (*ESR1*) encodes an estrogen receptor, a ligand-activated transcription factor composed of several domains. And *ESR1* gene polymorphism may affect the efficacy of drugs and the age of onset of RA (41-43).

Some of the drugs we have found are not commonly used to treat rheumatoid arthritis. The primary indication

Table 2 The specified information of drugs and its target genes

Number	Drug	Gene	Interaction type	Score*	Drug class	PMID
1	DANAZOL	*CCL2*	Inhibitor	1.67	Anti-endometriosis agent, antineoplastic agent	11056242
2	ETORICOXIB	*PTGS2*	Inhibitor	6.47	Anti-inflammatory agent, selective cyclooxygenase 2 inhibitors	17573128
3	CARPROFEN	*PTGS2*	Inhibitor	2.77	Anti-inflammatory agents, non-steroidal	15939622
4	FLURBIPROFEN	*PTGS2*	Inhibitor	1.08	Anti-inflammatory agents, non-steroidal	10091674
5	OXAPROZIN	*PTGS2*	Inhibitor	1.69	Anti-inflammatory agents, non-steroidal	19338579
6	ETODOLAC	*PTGS2*	Inhibitor	1.62	Anti-inflammatory agents, non-steroidal	12824918
7	SALSALATE	*PTGS2*	Inhibitor	1.62	Anti-Inflammatory agents, non-steroidal	9711054
8	CANAKINUMAB	*IL1B*	Inhibitor, binder, antibody	8.46	Therapeutic antibodies, DMARD	19169963
9	RILONACEPT	*IL1B*	Binder, inhibitor	2.01	Anti-inflammatory agent, DMARD	23319019
10	CAPLACIZUMAB	*VWF*	Inhibitor	14.2	Antithrombotic, antibody fragment	Not available
11	GOLIMUMAB	*TNF*	Inhibitor, antibody	3.87	Immunosuppressive agents	21079302
12	INFLIXIMAB	*TNF*	Inhibitor	1.13	Immunosuppressive agents, DMARD	16720636
13	ADALIMUMAB	*TNF*	Antibody, inhibitor	1.03	Anti-inflammatory agent, DMARD	12044041
14	ETANERCEPT	*TNF*	Inhibitor, antibody	0.7	Immunomodulatory agents, antirheumatic agents	16720636
15	SILTUXIMAB	*IL6*	Antagonist, antibody, inhibitor	8.61	Therapeutic antibodies, agents reducing cytokine levels	8823310
16	RANIBIZUMAB	*VEGFA*	Inhibitor	6.51	Treatment for AMD	18046235
17	PEGAPTANIB SODIUM	*VEGFA*	Antagonist	2.37	Treatment for AMD	23953100
18	AFLIBERCEPT	*VEGFA*	Antibody, binder, inhibitor	1.89	Treatment for AMD	22813448
19	CLOMIPHENE	*ESR1*	Antagonist	0.79	Fertility agents	19761360

*, the score is based on the evidence of an interaction, meaning that it doesn't changed from search to search. DMARD, disease-modifying antirheumatic drug; AMD, age-related macular degeneration.

for danazol is endometriosis, but it has also been studied for the treatment of autoimmune diseases such as autoimmune progesterone dermatitis, systemic lupus erythematosus, and autoimmune thrombocytopenia associated with rheumatic diseases (44-46). Caplacizumab is a drug used to treat acquired thrombotic thrombocytopenic purpura (47). More interestingly, some drugs (ranibizumab, aflibercept and pegaptanib sodium) that target *VEGFA* for the treatment of age-related macular degeneration may also be useful for rheumatoid patients with joint effusion. *VEGF* can induce migration of vascular endothelial cells and plays a key role in angiogenesis (48,49), which is essential in the pathological process of both RA and AMD. Therefore, our findings make it possible to use existing drugs for the treatment of RA.

In addition, DMARDs and *COX2* inhibitors are the more common drugs currently used in RA treatment. However, in many cases we do not know how to choose the type of them. Applying the drugs, we have found (*Table 1*) may give better results. As an *IL1B* inhibitor, Canakinumab, which has the highest interaction score in these drugs, deserves to be more explored for its therapeutic value in RA patients with joint effusion. Although NASIDs are no longer first-line drugs for rheumatoid therapy, they still have a reliable effect on the improvement of pain and stiffness (3). Interestingly, our study shows that joint effusion, one of the symptoms

Table 3 The functional enrichments of the final 8 genes

Category	ID	Term	Genes	P value
BP_FAT	GO:0045429	Positive regulation of nitric oxide biosynthetic process	*IL6, IL1B, PTGS2, ESR1, TNF*	1.34E-09
BP_FAT	GO:1904407	Positive regulation of nitric oxide metabolic process	*IL6, IL1B, PTGS2, ESR1, TNF*	1.34E-09
BP_FAT	GO:1903428	Positive regulation of reactive oxygen species biosynthetic process	*IL6, IL1B, PTGS2, ESR1, TNF*	2.50E-09
BP_FAT	GO:0032103	Positive regulation of response to external stimulus	*IL6, IL1B, CCL2, PTGS2, TNF, VEGFA*	2.09E-08
BP_FAT	GO:0032101	Regulation of response to external stimulus	*IL6, IL1B, CCL2, PTGS2, ESR1, TNF, VEGFA*	4.36E-08
BP_FAT	GO:0050727	Regulation of inflammatory response	*IL6, IL1B, CCL2, PTGS2, ESR1, TNF*	4.71E-08
BP_FAT	GO:0050729	Positive regulation of inflammatory response	*IL6, IL1B, CCL2, PTGS2, TNF*	7.97E-08
BP_FAT	GO:1903829	Positive regulation of cellular protein localization	*IL6, IL1B, CCL2, PTGS2, TNF, VEGFA*	9.71E-08
BP_FAT	GO:0008284	Positive regulation of cell proliferation	*IL6, IL1B, CCL2, PTGS2, ESR1, TNF, VEGFA*	1.16E-07
BP_FAT	GO:0002675	Positive regulation of acute inflammatory response	*IL6, IL1B, PTGS2, TNF*	1.33E-07
CC_FAT	GO:0005615	Extracellular space	*IL6, IL1B, CCL2, TNF, VEGFA*	0.00265906
MF_FAT	GO:0005125	Cytokine activity	*IL6, IL1B, CCL2, TNF, VEGFA*	1.39E-06
MF_FAT	GO:0005126	Cytokine receptor binding	*IL6, IL1B, CCL2, TNF, VEGFA*	3.18E-06
MF_FAT	GO:0005102	Receptor binding	*IL6, VWF, IL1B, CCL2, TNF, VEGFA*	1.33E-04

of rheumatoid arthritis, can be relieved by selective *COX2* inhibitors. Given the adverse effect profile of these drugs and their inability to slow disease progression, if necessary, the combination of *COX2* inhibitors with proton-pump inhibitors, DMARDs, etc. should be considered. The levels of nitric oxide and inducible nitric oxide synthase (INOS), which act as mediators of the inflammatory response, are significantly increased in the synovial fluid of RA patients. By interfering with the synthesis of IL12 and B-cell activating factors (BAF), nitric oxide can disrupt the normal physiological function of T and B cells (50,51). Also, studies have shown that nitric oxide plays a vital role in vascular endothelial cell dysfunction in RA (52). Serum cytokine activity in RA patients is strongly correlated with treatment effect, and multiplexed cytokine testing may be a new way of evaluation (53). Activation of the MAPK, NF-κB, JAK-STAT, and Toll signaling pathways can lead to the development of RA. In contrast, activation of the classical Wnt pathway can induce osteogenesis and promote the repair of damaged joints, which has a positive effect on the treatment of RA (54-56).

In our study, a bioinformatics strategy was used to search for drugs suitable for RA patients with joint effusion. Among the 19 drugs identified, the initial indications are not only RA but also AMD, tumor, and other diseases. As a result of our analysis, it suggests opportunities for further research in the clinical setting and give us a chance to explore new indications of these available drugs. Although our study provides new insight into drug discovery in rheumatoid arthritis, more basic experiments and clinical evaluations are needed to confirm our results in the future.

Conclusions

The final eight genes (*PTGS2*, *TNF*, *VEGFA*, *IL1B*, *CCL2*, *VWF*, *IL6*, and *ESR1*) and 19 drugs may provide significant therapeutic value for RA patients with joint effusion.

Acknowledgments

Funding: This work was supported by National Key Research and Development Program of China (No. 2020YFC2004900).

Footnote

Reporting Checklist: The authors have completed the MDAR checklist. Available at http://dx.doi.org/10.21037/apm-20-

2631b

Peer Review File: Available at http://dx.doi.org/10.21037/apm-20-2631b

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/apm-20-2631b). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. We just re-analyzed the open accessed datasets (http://pubmed2ensembl.ls.manchester.ac.uk), and no ethical approval was required by the local ethics committees. The study conformed to the provisions of the Declaration of Helsinki (as revised in 2013).

References

1. Hunter TM, Boytsov NN, Zhang X, et al. Prevalence of rheumatoid arthritis in the United States adult population in healthcare claims databases, 2004-2014. Rheumatol Int 2017;37:1551-7.
2. Rudan I, Sidhu S, Papana A, et al. Prevalence of rheumatoid arthritis in low- and middle-income countries: A systematic review and analysis. J Glob Health 2015;5:010409.
3. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet 2010;376:1094-108.
4. Bergstra SA, Allaart CF, Ramiro S, et al. Sex-associated Treatment Differences and Their Outcomes in Rheumatoid Arthritis: Results from the METEOR Register. J Rheumatol 2018;45:1361-6.
5. van Vollenhoven RF. Sex differences in rheumatoid arthritis: more than meets the eye. BMC Med 2009;7:12.
6. An J, Nyarko E, Hamad MA. Prevalence of comorbidities and their associations with health-related quality of life and healthcare expenditures in patients with rheumatoid arthritis. Clin Rheumatol 2019;38:2717-26.
7. Fazal SA, Khan M, Nishi SE, et al. A Clinical Update and Global Economic Burden of Rheumatoid Arthritis. Endocr Metab Immune Disord Drug Targets 2018;18:98-109.
8. Mortality GBD, Causes of Death C. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;385:117-71.
9. Klareskog L, Padyukov L, Lorentzen J, et al. Mechanisms of disease: Genetic susceptibility and environmental triggers in the development of rheumatoid arthritis. Nat Clin Pract Rheumatol 2006;2:425-33.
10. Padyukov L, Silva C, Stolt P, et al. A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 2004;50:3085-92.
11. Feldmann M, Brennan FM, Maini RN. Rheumatoid arthritis. Cell 1996;85:307-10.
12. Choy EH, Isenberg DA, Garrood T, et al. Therapeutic benefit of blocking interleukin-6 activity with an anti-interleukin-6 receptor monoclonal antibody in rheumatoid arthritis: a randomized, double-blind, placebo-controlled, dose-escalation trial. Arthritis Rheum 2002;46:3143-50.
13. Singh JA, Saag KG, Bridges SL Jr, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. Arthritis Rheumatol 2016;68:1-26.
14. van Everdingen AA, Jacobs JW, Siewertsz Van Reesema DR, et al. Low-dose prednisone therapy for patients with early active rheumatoid arthritis: clinical efficacy, disease-modifying properties, and side effects: a randomized, double-blind, placebo-controlled clinical trial. Ann Intern Med 2002;136:1-12.
15. Alonso-Ruiz A, Pijoan JI, Ansuategui E, et al. Tumor necrosis factor alpha drugs in rheumatoid arthritis: systematic review and metaanalysis of efficacy and safety. BMC Musculoskelet Disord 2008;9:52.
16. Sethi MK, O'Dell JR. Combination conventional DMARDs compared to biologicals: what is the evidence? Curr Opin Rheumatol 2015;27:183-8.
17. Moosavinasab S, Patterson J, Strouse R, et al. 'RE:fine drugs': an interactive dashboard to access drug repurposing opportunities. Database (Oxford) 2016;2016.
18. Berg EL. Systems biology in drug discovery and development. Drug Discov Today 2014;19:113-25.

19. Mosca E, Bertoli G, Piscitelli E, et al. Identification of functionally related genes using data mining and data integration: a breast cancer case study. BMC Bioinformatics 2009;10 Suppl 12:S8.
20. Yu S, Tranchevent LC, De Moor B, et al. Gene prioritization and clustering by multi-view text mining. BMC Bioinformatics 2010;11:28.
21. Baran J, Gerner M, Haeussler M, et al. pubmed2ensembl: a resource for mining the biological literature on genes. PLoS One 2011;6:e24716.
22. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27-30.
23. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019;47:D607-13.
24. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015;43:D447-52.
25. Su G, Morris JH, Demchak B, et al. Biological network exploration with Cytoscape 3. Curr Protoc Bioinformatics 2014;47:8.13.1-24.
26. Cotto KC, Wagner AH, Feng YY, et al. DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. Nucleic Acids Res 2018;46:D1068-73.
27. Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 2014;506:376-81.
28. Spangler RS. Cyclooxygenase 1 and 2 in rheumatic disease: implications for nonsteroidal anti-inflammatory drug therapy. Semin Arthritis Rheum 1996;26:435-46.
29. Wei Q, Lv F, Zhang H, et al. MicroRNA-101-3p inhibits fibroblast-like synoviocyte proliferation and inflammation in rheumatoid arthritis by targeting PTGS2. Biosci Rep 2020;40.
30. Mangoni AA, Al Okaily F, Almoallim H, et al. Relapse rates after elective discontinuation of anti-TNF therapy in rheumatoid arthritis: a meta-analysis and review of literature. BMC Rheumatol 2019;3:10.
31. Ramirez-Bello J, Cadena-Sandoval D, Fragoso JM, et al. The VEGFA -1154G/A polymorphism is associated with reduced risk of rheumatoid arthritis but not with systemic lupus erythematosus in Mexican women. J Gene Med 2018;20:e3024.
32. Zhang Y, Qiu H, Zhang H, et al. Vascular endothelial growth factor A (VEGFA) polymorphisms in Chinese patients with rheumatoid arthritis. Scand J Rheumatol 2013;42:344-8.
33. Zhang F, Wei K, Slowikowski K, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat Immunol 2019;20:928-42.
34. Tylaska LA, Boring L, Weng W, et al. Ccr2 regulates the level of MCP-1/CCL2 in vitro and at inflammatory sites and controls T cell activation in response to alloantigen. Cytokine 2002;18:184-90.
35. Zhang L, Yu M, Deng J, et al. Chemokine Signaling Pathway Involved in CCL2 Expression in Patients with Rheumatoid Arthritis. Yonsei Med J 2015;56:1134-42.
36. Fang Z, Lv J, Wang J, et al. C-Reactive Protein Promotes the Activation of Fibroblast-Like Synoviocytes From Patients With Rheumatoid Arthritis. Front Immunol 2020;11:958.
37. Katoh M, Katoh M. STAT3-induced WNT5A signaling loop in embryonic stem cells, adult normal tissues, chronic persistent inflammation, rheumatoid arthritis and cancer (Review). Int J Mol Med 2007;19:273-8.
38. Jones G, Sebba A, Gu J, et al. Comparison of tocilizumab monotherapy versus methotrexate monotherapy in patients with moderate to severe rheumatoid arthritis: the AMBITION study. Ann Rheum Dis 2010;69:88-96.
39. Ristic GG, Subota V, Lepic T, et al. Subclinical Atherosclerosis in Patients with Rheumatoid Arthritis and Low Cardiovascular Risk: The Role of von Willebrand Factor Activity. PLoS One 2015;10:e0130462.
40. Daza L, Aguirre M, Jimenez M, et al. Common carotid intima-media thickness and von Willebrand factor serum levels in rheumatoid arthritis female patients without cardiovascular risk factors. Clin Rheumatol 2007;26:533-7.
41. Pawlik A, Dziedziejko V, Kurzawski M, et al. Effect of ESR1 and ESR2 gene polymorphisms on rheumatoid arthritis treatment with methotrexate. Pharmacol Rep 2012;64:185-90.
42. Dziedziejko V, Kurzawski M, Safranow K, et al. Oestrogen receptor polymorphisms in female patients with rheumatoid arthritis. Scand J Rheumatol 2011;40:329-33.
43. Dziedziejko V, Kurzawski M, Safranow K, et al. The effect of ESR1 and ESR2 gene polymorphisms on the outcome of rheumatoid arthritis treatment with leflunomide. Pharmacogenomics 2011;12:41-7.
44. Zhang M, Tang X, Zhou H, et al. Case of autoimmune progesterone dermatitis presenting as necrotic migratory erythema successfully controlled by danazol. J Dermatol 2020;47:178-80.
45. Letchumanan P, Thumboo J. Danazol in the treatment

of systemic lupus erythematosus: a qualitative systematic review. Semin Arthritis Rheum 2011;40:298-306.
46. Blanco R, Martinez-Taboada VM, Rodriguez-Valverde V, et al. Successful therapy with danazol in refractory autoimmune thrombocytopenia associated with rheumatic diseases. Br J Rheumatol 1997;36:1095-9.
47. Peyvandi F, Scully M, Kremer Hovinga JA, et al. Caplacizumab for Acquired Thrombotic Thrombocytopenic Purpura. N Engl J Med 2016;374:511-22.
48. Ferrara N. Role of vascular endothelial growth factor in regulation of physiological angiogenesis. Am J Physiol Cell Physiol 2001;280:C1358-66.
49. Shibuya M. Vascular endothelial growth factor-dependent and -independent regulation of angiogenesis. BMB Rep 2008;41:278-86.
50. Liu Y, Wei W, Hong C, et al. Calreticulin induced endothelial ICAM-1 up-regulation associated with tristetraprolin expression alteration through PI3K/Akt/eNOS/p38 MAPK signaling pathway in rheumatoid arthritis. Mol Immunol 2019;107:10-20.
51. Farrell AJ, Blake DR, Palmer RM, et al. Increased concentrations of nitrite in synovial fluid and serum samples suggest increased nitric oxide synthesis in rheumatic diseases. Ann Rheum Dis 1992;51:1219-22.
52. Garg N, Syngle A, Krishan P. Nitric Oxide: Link between Inflammation and Endothelial Dysfunction in Rheumatoid Arthritis. Int J Angiol 2017;26:165-9.
53. Alex P, Szodoray P, Knowlton N, et al. Multiplex serum cytokine monitoring as a prognostic tool in rheumatoid arthritis. Clin Exp Rheumatol 2007;25:584-92.
54. Cici D, Corrado A, Rotondo C, et al. Wnt Signaling and Biological Therapy in Rheumatoid Arthritis and Spondyloarthritis. Int J Mol Sci 2019;20:5552.
55. Schett G, Zwerina J, Firestein G. The p38 mitogen-activated protein kinase (MAPK) pathway in rheumatoid arthritis. Ann Rheum Dis 2008;67:909-16.
56. Kim HR, Cho ML, Kim KW, et al. Up-regulation of IL-23p19 expression in rheumatoid arthritis synovial fibroblasts by IL-17 through PI3-kinase-, NF-kappaB- and p38 MAPK-dependent signalling pathways. Rheumatology (Oxford) 2007;46:57-64.

Cite this article as: Li Z, Fu J, Cao Y, Xu C, Han X, Zhang W, Song Z, Chen J. Drug discovery in rheumatoid arthritis with joint effusion identified by text mining and biomedical databases. Ann Palliat Med 2021;10(5):5218-5230. doi: 10.21037/apm-20-2631b

