Network pharmacology and molecular docking analysis on molecular targets and mechanisms of Angelicae Pubescentis Radix in the treatment of rheumatoid arthritis

yanni yang
  shaanxi university of traditional chinese medicine

yirixiati aihaiti
  xi'an jiaotong university affiliated honghui hospital

peng xu (✉ sousou369@163.com)
  xi'an jiaotong university affiliated honghui hospital  https://orcid.org/0000-0003-2487-9163

haishi zheng
  xi'an jiaotong university affiliated honghui hospital

Research Article

Keywords: rheumatoid arthritis, Angelicae Pubescentis Radix, Network pharmacology, molecular docking

DOI: https://doi.org/10.21203/rs.3.rs-445750/v1

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

**Purpose:** To explore the potential target proteins underlying the effect of Angelicae Pubescentis Radix (APR) on rheumatoid arthritis (RA) using a network pharmacology and molecular docking approach.

**Methods:** First, the active components and target proteins of APR and RA related disease targets were obtained from the TCMSP, Gene Card, OMIM, DisGeNET and STRING databases. Then the active ingredient target in the RA network diagram was drawn using Cytoscape 3.7.1 software. Protein-protein interaction analysis, Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were carried out using the STRING and David databases. The crystal structures of RA related targets were retrieved from the RCSB PDB database. Finally, the potential active compounds and their related targets were validated using molecular docking technology.

**Results:** Five active components of Angelicae Pubescentis Radix (APR) were screened out, including ammidin, isoimperatorin, beta-sitosterol, O-acetylcolubianetin and angelicone and 80 key targets including MAPK8, EGFR, PTGS2, CASPASE3, MTOR, SRC, KDR, MAPK1, NOS3 and MAPK14, etc were obtained. GO enrichment analysis showed that 222 biological processes, 34 cell components and 72 molecular functions were identified; KEGG analysis showed that the targets of APR in the treatment of RA were significantly enriched in pathways in cancer, the PI3K-Akt signaling pathway, Proteoglycans in cancer, osteoclast differentiation, neuroactive ligand-receptor interaction, tuberculosis, TNF signaling pathway, serotonergic synapse, Rap1 signaling pathway, cAMP signaling pathway. The results of molecular docking showed that ammidin, isoimperatorin, beta-sitosterol, O-acetylcolubianetin and angelicone had strong affinity for PTGS2, EGFR and MAPK8.

**Conclusion:** APR treats RA through the characteristics of multi-component, multi-target and multi-pathway regulation.

1 Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune disease that not only causes synovial inflammation, joint swelling, bone and cartilage, and persistent damage to surrounding tissues, but also causes fractures, deformity, cardiovascular disease, mental disease, and even death [1–3]. Research shows that approximately 0.5%-1% of people in the world suffer from RA, which is one of the main causes of disability [4]. RA is one of the most common chronic diseases, which not only causes great pain to patients, but also brings great burden to society [5]. RA treatment is mainly relieves pain and reduces inflammation [6]. Disease-modifying anti-rheumatic drugs (DMARDs) are one of the most commonly used drugs for the treatment of RA [7]. However, long-term clinical practice has found that long-term use of DMARDs has various adverse effects (e.g., tinnitus, hearing loss, gastrointestinal discomfort and so on) [8]. The therapeutic effects of phytochemical drugs on the treatment and prevention of various diseases have aroused widespread interest. Chinese herbal medicine treatment of rheumatoid arthritis has
a history of thousands of years, and the curative effect is good. Because of their nature characteristics, Chinese herbal medicines have fewer side effects in the treatment of RA [9–11]. APR mainly comes from the root of Angelica sinensis, which has the effect of dispelling wind and removing dampness, relieving pain, and is a commonly used drug in the treatment of rheumatoid arthritis [12, 13]. In recent years, with the rapid development of bioinformatics, the comprehensive methods in pharmacology and biological networks have been widely used in the field of life sciences [14]. Network pharmacology is considered a promising method for understanding herbal formulas and predicting potential new drugs or targets for specific diseases [15, 16]. Molecular docking is an important approach in structural molecular biology and computer-aided drug design. Molecular docking is an important approach in structural molecular biology and computer-aided drug design [17–19]. This study intends to study the main targets and pathways of APR in the treatment of RA through network pharmacology and molecular docking, in order to platform provide reference for further research and verification of its related mechanisms. The experimental roadmap is shown in Fig. 1.

2 Materials And Methods

2.1 Data Preparation

2.1.1 Components of RAP

Traditional Chinese medicine systems pharmacology platform (TCMSP) is a traditional Chinese medicine system, pharmacology database and analysis platform integrating network information such as pharmacochemistry, pharmacokinetics and network target prediction [20]. In this study, TCMSP database was used to screen the active components of APR according to the ADME properties of drug molecules [21]. The oral bioavailability (OB) ≥ 30% and drug likeness(DL) ≥ 0.18 were selected as the screening conditions, and the active component database of APR was established. Getting the 2D structure of RAP from PubChem database. The mol2 file of the active ingredient of APR was downloaded and converted into pdbqt format by Open Babel GUI software for molecular docking experient.

2.1.2 Prediction of APR target

The target of action of all components of traditional Chinese medicine was uploaded to UniProt (https://www.uniprot.org/) database for target gene name conversion Gene information; construct component and target network with Cytoscape 3.7.1 software [22–25].

2.1.3 Prediction of RA related pathogenesis targets

Genecards (https://www.genecards.org/) database provides all the information about annotation and prediction of human genes. It automatically integrates gene centric data from about 150 network sources, including genome, transcriptomics, proteomics, genetics, clinical and functional information [22–25]. OMIM(https://www.OMIM.org/) is a database of human genes and genetic disorders. It mainly focuses on genetic diseases, including text information and related reference information, sequence records,
maps and other related databases [27]. DisGeNET (http://www.disgenet.org) Data from public databases, GWAS catalogue, animal models and scientific literature were integrated to collect a large number of mutations and genes related to human diseases [28, 29]. Using “rheumatoid arthritis” as the key word, we searched for the pathogenesis targets related to RA in GeneCards, OMIM and DisGeNET databases. The results of the above three databases were integrated, and the repetitive targets were eliminated to get the targets related to RA.

2.1.4 Prediction of potential therapeutic targets of APR for RA

The compound targets and disease targets were introduced into Venny 2.1.0 (https://bioinfogp.cnb.csic.es/tools/venny/) online analysis tool to obtain the intersection targets.

2.2 Network constructure

STRING (https://www.string-db.org/) is a database for searching known and predicted protein-protein interactions. The database can be applied to 2031 species, including 9.6 million proteins and 13.8 million protein-protein interactions [30–32]. The interaction relationship between these target proteins was observed, the intersection targets were imported into the STRING database, the organization was set as "Homo aspens", the protein interaction information of the string was saved in TSV format and imported into the Cytoscape software for visual analysis. The network diagrams of "APR-active components-key targets-RA" was drew respectively. The PPI network is constructed [33–35]. The network analyzer function is used to analyze the topological properties, and the targets whose degree value and betweenness value are above the average value are selected, which are considered as the key targets.

2.3 Enrichment analysis

David(https://david.ncifcrf.gov/) is a biological information database, which integrates biological data and analysis tools, provides systematic and comprehensive biological function annotation information for large-scale gene or protein list, and helps users to extract biological information from it [36]. The indexes with degree value and betweenness greater than the average value were selected as the closed key indexes and input into David database for GO analysis and KEGG analysis.

2.4 Active components-targets docking

The 2D and 3D structures of each active ingredients were downloaded from PubChem online database (https://pubchem.ncbi.nlm.nih.gov/) [37]. 3D structure of active ingredients were uploaded to chemofice software to minimize the energy, saved in MOL2 format. The 3D structures of PTGS2, EGFR and MAPK8 ligands, which ranked the top three in PPI network were obtained from PDB database (http://www.rcsb.org/) by limiting the organism to “Homo sapiens”. PyMOL software was used for hydrogenation, dehydration and original ligands extraction. Finally Vina was run for molecular docking. It is considered that the binding energy $\leq -5KJ / mol$ indicates better binding activity [38].

3 Results
3.1 Chemical information for the active compounds of APR

Active ingredients of APR were obtained from TCMSP database. According to the ADME properties of drug molecules, five active components were screened out based on OB ≥ 30% and DL ≥ 0.18. Canonical SMILES and 2D structures of selected ingredients were obtained from PubChem (Table 1).

Table 1

<table>
<thead>
<tr>
<th>Mol ID</th>
<th>Molecule Name</th>
<th>OB(%)</th>
<th>DL</th>
<th>MW</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOL001941</td>
<td>ammidin</td>
<td>34.55</td>
<td>0.22</td>
<td>270.3</td>
<td></td>
</tr>
<tr>
<td>MOL001942</td>
<td>isoimperatorin</td>
<td>45.46</td>
<td>0.23</td>
<td>270.3</td>
<td></td>
</tr>
<tr>
<td>MOL000358</td>
<td>beta-sitosterol</td>
<td>36.91</td>
<td>0.75</td>
<td>414.79</td>
<td></td>
</tr>
<tr>
<td>MOL003608</td>
<td>o-acetylcolumbianetin</td>
<td>60.04</td>
<td>0.26</td>
<td>288.32</td>
<td></td>
</tr>
<tr>
<td>MOL004780</td>
<td>angelicone</td>
<td>30.99</td>
<td>0.19</td>
<td>288.32</td>
<td></td>
</tr>
</tbody>
</table>

3.2 Screening of active compounds and potential targets

265 potential targets of active ingredients were obtained from TCMSP database. 1197 RA-related genes were obtained from GeneCards, OMIM, DisGeNET and PubChem databases. 80 overlapped genes were selected as potential targets of APR against RA(Fig. 2).

3.3 Construction of “APR- active ingredients -key target -RA” network

The network topology map of "RAP - active ingredients - key targets - RA" was drawn by Cytoscape 3.7.1 software (Fig. 3). The top ten targets of highest degree were selected as the core targets, which were MAPK8, EGFR, PTGS2, Caspase3, mTOR, SRC, KDR, mapk1, NOS3, MAPK14, respectively. These 10 targets may played a key role in the therapeitoc effect of APR against RA.

3.3 The construction of PPI network

The information of 80 targets mapped by drugs and diseases is imported into STRING database to obtain the interaction between targets. The visualization processing shows that PPI network contains 80 nodes and 549 edges. The larger the node, the greater the degree. 34 targets exceeded the average. Top 10 targets with the highest degree were selected (Fig. 4).

3.4 GO enrichment analysis and KEGG pathway analysis
Potential targets were uploaded to DAVID database. As a result, a total of 328 GO items with P < 0.05 were obtained, including 222 biological process entries, 34 cell component entries, and 72 molecular function entries. In biological processes, the target genes were mostly involved in the regulation of signal transduction, positive regulation of transcription from RNA polymerase II promoter, protein phosphorylation, oxidation-reduction process and inflammatory response. In the cell components, the target genes potentially play roles in the plasma membrane, cytosol, nucleus and cytoplasm; in molecular functions, the target genes may play roles in the process of protein binding, ATP binding, enzyme binding, protein homodimerization activity and zinc ion binding. The top 10 items with the highest -log(p) value of each component were visualized (Fig. 5). For KEGG pathway analysis, a total of 80 pathways were screened out (P < 0.05). The main pathways include cancer related pathways, PI3K-Akt signaling pathway, TNF signaling pathway and Osteoclast differentiation (Fig. 6).

In summary, potential target genes were enriched in several pathways known to participate in pathogenesis of RA, including PI3K-Akt signaling pathway, TNF signaling pathway, Jak-STAT signaling pathway, NF-kappa B signaling pathway, apoptosis and mTOR signaling pathway. Significant pathways associated with pathogenesis of RA were combined as an integrated APR pathway model, which may provide evidence for exploration of drug targets.

The potential signaling pathway is the possible mechanism and interaction of RA. Enrichment factor index represents the ratio of the number of pathway related target genes, which represents the number of annotated genes in a specific pathway. The higher the enrichment factor score, the higher the enrichment level. The size of the dot represents the number of target genes in its representative pathway, and the p value with scale is highlighted with different colors, as shown at the top of the figure.

3.5 Molecular docking analysis

In this study, the possible interactions between the three central genes and their corresponding compounds were verified by molecular docking. At the same time, the target and active compounds were further filtered by the docking affinity value reported by autodock Vina. The higher the absolute value of docking affinity, the stronger the binding ability between the compound and the active site of the targets. In the docking results, most of the binding complexes had high binding affinity. It is considered to have better binding ability when the binding energy is less than −5.0 kcal/mol. The modes of top 10 binding complexes are displayed in Fig. 7. The binding energies of various compounds are shown in Table 2.
### Table 2

**Binding energy of compounds**

<table>
<thead>
<tr>
<th>Target</th>
<th>PDB ID</th>
<th>Compound</th>
<th>Docking affinity (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR</td>
<td>6ICG</td>
<td>Ammidin</td>
<td>5.61</td>
</tr>
<tr>
<td>EGFR</td>
<td>6ICG</td>
<td>beta-sitosterol</td>
<td>6.81</td>
</tr>
<tr>
<td>EGFR</td>
<td>6ICG</td>
<td>O-Acetylcolumbianetin</td>
<td>5.61</td>
</tr>
<tr>
<td>EGFR</td>
<td>6ICG</td>
<td>Angelicone</td>
<td>5.56</td>
</tr>
<tr>
<td>MAPK8</td>
<td>4G1W</td>
<td>Ammidin</td>
<td>6.41</td>
</tr>
<tr>
<td>MAPK8</td>
<td>4G1W</td>
<td>isoimperatorin</td>
<td>6.61</td>
</tr>
<tr>
<td>MAPK8</td>
<td>4G1W</td>
<td>beta-sitosterol</td>
<td>6.34</td>
</tr>
<tr>
<td>MAPK8</td>
<td>4G1W</td>
<td>O-Acetylcolumbianetin</td>
<td>5.98</td>
</tr>
<tr>
<td>MAPK8</td>
<td>4G1W</td>
<td>Angelicone</td>
<td>5.62</td>
</tr>
<tr>
<td>PTGS2</td>
<td>5F19</td>
<td>Ammidin</td>
<td>5.31</td>
</tr>
</tbody>
</table>

### 4. Discussion

Rheumatoid arthritis (RA) is a chronic, inflammatory, and systemic autoimmune disease. Its clinical manifestation is multiple arthritis of the small joints of the hands and feet, which can lead to joint pain, injury and even loss of function, and seriously affects the quality of life of patients [39]. At present, the pathogenesis of rheumatoid arthritis is not very clear, but most studies believe that T lymphocytes and macrophages release a large number of inflammatory cytokines to cause bone and joint damage in rheumatoid arthritis [40].

Traditional Chinese medicine (TCM) believes that APR has the effect of dispelling cold and dampness, dredging channels, and relieving pain, which can effectively treat rheumatoid arthritis. Modern pharmacological studies have shown that the effective components of APR can inhibit the expression of tumor necrosis factor-α (TNF-α), interleukin (IL), and high-sensitivity C-reactive protein (hs-CRP) [12]. The flavonoids, glycosides, organic acids and other compounds in *Angelica pubescens* may be used to treat rheumatoid arthritis by inhibiting the activity of COX-2 [41]. We used the TCMSP database to screen out five major compounds, including, ammidin, isoimperatorin, beta-sitosterol, O-acetylcolumbianetin and angelicone. Studies have found that isoimperatorin has anti-inflammatory, analgesic, antispasmodic, and anticancer effects, and regulates cell function through the PI3K-Akt signaling pathway [42]. Beta-sitosterol is widely found in many plants and has many biological functions [43]. Beta-sitosterol, a plant-derived bioactive compound, has a chemical structure similar to that of mammalian cells [44]. The antitumor effects of beta-sitosterol have been widely investigated. Beta-sitosterol markedly decreased the proliferative index by downregulating β-catenin and PCNA in colon cancer cells [45]. Beta-sitosterol exerts anti-rheumatoid effects by modulating macrophage polarization in collagen-induced arthritis rats [46]. Thus, we speculated that the active ingredients in APR may exert anti-inflammatory effects in RA.
In recent years, with the rapid development of bioinformatics analysis and network pharmacology, the advantages of this kind of speculation and simulation of pharmacokinetics and drug toxicity are more prominent. In the present study, we utilized the TCMSP database to investigate the ADME-related characteristics of APR. Five active components in APR with good pharmacokinetic characteristics and drug properties were selected that might play a therapeutic role in RA by regulating the key target proteins. By constructing a PPI network of intersected target genes, we obtained MAPK8, EGFR, PTGS2, Caspase3, mTOR, SRC, KDR, MAPK1, NOS3 and MAPk14 as core targets. Among the core targets, PTGS2, EGFR and MAPK8 nodes were the top three nodes with the highest degree, (28, 28, and 27 degrees, respectively). The PTGS family is a key enzyme in prostaglandin biosynthesis and acts both as a dioxygenase and as a peroxidase. PTGS2 is closely related to inflammatory factors in synovial tissue [47, 48]. PTGS2 secretion is positively correlated with the pathological damage of synovial tissue, and also affected the expression of apoptosis-related factors, Caspase-3, Bcl-2, and Bax [49]. MAPK8 is a member of the MAP kinase family, which is involved in variety of cellular process. MAPK8 and EGFR play important roles in tumor growth [50]. Similar to tumor growth, RA synoviocytes also express EGFR, MAPK8, and their ligands. Binding to EGF induces receptor dimerization and phosphorylation, which causes downstream signal transduction cascades, including the MAPK and PI3K/Akt signaling pathways.

In the process of gene ontology (GO) enrichment analysis in this study, we found that the biological process of APR in the treatment of RA mainly includes the regulation of signal transduction, positive regulation of RNA polymerase II promoter transcription, protein phosphorylation, redox processes, and inflammatory reactions. In cell components, target genes may play roles in the plasma membrane, cytoplasm, nucleus and cytoplasm. In molecular function, target genes may participate in protein binding, ATP binding, enzyme binding, protein homodimer activity, zinc ion binding and other processes. Through the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment, we identified some signaling pathways that are known to be involve in the pathogenesis of RA. By integrating these pathways and screening target genes, the potential APR pathway model was established(Fig. 8). Among these enriched pathways, the PI3K/Akt signaling pathway was confirmed to regulate chondrocyte and synovial cell proliferation, angiogenesis, apoptosis, and autophagy in the pathogenesis of RA [51, 52]. HIF-1 is a key factor in the hypoxia response of fibroblasts to hypoxia [51, 52]. High expression of HIF-1 can induces migration and invasion of fibroblasts. HIF-1 is also associated with synovitis. Some studies have shown that high expression of HIF-1 can cause the proliferation of inflammatory T cells and produce a large number of inflammatory factors [54]. The JAk-STAT signaling pathway is closely related to RA pain. Continuous activation of the JAk-STAT signaling pathway leads to the secretion of inflammatory cytokines in the synovium and participates in the inflammatory response [55]. JAK inhibitors effectively alleviate the pain in patients [56]. The NF-kappa B signaling pathway is ubiquitous in many kinds of cells. In synovial cells, NF - kappa B enters the nucleus after activation, regulates the transcription of TNF-α, IL-1β, IL-6, and other cytokines, and participates in the pathogenesis of RA. Thus, we speculate that APR may exert therapeutic effects against RA through multi signaling pathways.
To further explore the therapeutic effect of APR on RA, we performed molecular docking. The results showed that most of the active components in APR have a strong binding force to the top three key targets with binding affinities $<-5$ kcal/mol, which indicates that the therapeutic effects of APR against RA were possibly through the interactions with related targets.

**Conclusions**

In this study, we preliminarily confirmed the potential pharmacological mechanism of APR in the treatment of RA using the method of network pharmacology and molecular docking. This research method is based on "disease-gene target-network" interactions and reveals the network characteristics of TCM more comprehensively. Compared with single-target drugs, multi-target drugs have better clinical efficacy and fewer adverse reactions. TCM and its compounds have the advantage of multi-component, multi-target, and multi-link interventions, which can platform new ideas for the future research and the development of new drugs.

**Declarations**

**AUTHOR CONTRIBUTIONS**

YNY and YA designed the experiments; YNY, YA and HSZ performed Network pharmacology and molecular docking analysis. YA and YNY wrote the manuscript. All authors read and approved the final version and agreed to be accountable for all aspects of the work.

**FUNDING**

The present study was supported by a grant from the National Natural Science Foundation of China (grant no. 81772410 and grant no.8207091212).

**References**


[52] Xia P, Xu XY, PI3K/Akt/mTOR signaling pathway in cancer stem cells: from basic research to clinical application, AM J CANCER RES. 5 (2015) 1602-9,


Figures
Figure 1

Roadmap of this study.
Figure 2

Venny diagram of RA and RAP intersected target.
Figure 3

Compound-putative target network of "APR - active ingredients - key targets - diseases", Note: The square node indicates the disease; the round node indicates the RAP; the triangle node indicates the active ingredient; the diamond node indicates the key target.
Figure 4

PPI network of APR and RA intersection targets; PPI network diagram of top ten targets.
Figure 5

The selected genes were enriched by Gene Ontology (GO). The GO analysis had shown 222 entries on biological processes, 34 entries on cell components, and 72 entries on molecular functions with $P<0.05$. The top 10 entries with the most significant $P$ value are shown.
Figure 6

The first 20 potential KEGG pathways of the target genes screened in RA were enriched.
Figure 7

Molecular docking diagram of active components and targets. (a) MAPK8-beta-sitosterol,(b) MAPK8-Ammidin,(c) MAPK-isoimperatorin,(d) MAPK-O-Acetylcolumbianetin,(e) MAPK8-Angelicone,(f) EGFR-Ammidin,(g) EGFR-O-Acetylcolumbianetin,(h) EGFR-beta_sitosterol,(i) PTGS2-Ammidin
Figure 8

The Potential signaling pathways of APR