Investigation of the potential mechanism of the Shugan Xiaozhi decoction for the treatment of nonalcoholic fatty liver disease based on network pharmacology and molecular docking

Yufeng Xing (✉ yufeng000729@163.com)  
Shenzhen Traditional Chinese Medicine Hospital  https://orcid.org/0000-0001-6762-3788

Rong Yang  
Macau University of Science and Technology

Huili Yang  
Guangzhou University of Chinese Medicine

Dansheng Jiang  
Guangzhou University of Chinese Medicine

Linyi Xu  
Guangzhou University of Chinese Medicine

Lian Feng  
Guangzhou University of Chinese Medicine

Research

Keywords: Shugan Xiaozhi decoction, Nonalcoholic fatty liver disease (NAFLD), Network pharmacology, Molecular docking, Inflammation

Posted Date: February 2nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1294807/v1

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

Background: Nonalcoholic fatty liver disease (NAFLD) is a metabolic-related disease with its incidence increasing annually. Shugan Xiaozhi (SGXZ) decoction, a composite traditional Chinese medicinal prescription, was demonstrated to exert a therapeutic effect on NAFLD. However, the underlying mechanisms were incompletely elucidated. In this study, the potential bioactive ingredients and mechanism of the SGXZ decoction against NAFLD were explored via network pharmacology and molecular docking.

Methods: Compounds in SGXZ decoction were identified and collected from literature studies, and the corresponding targets were predicted through SEA database. Potential targets related to NAFLD were overlapped by using DisGeNET and GeneCards database. The compound-target-disease and PPI network were then constructed based on the putative intersection targets of SGXZ decoction and NAFLD to recognize the key compounds and targets. Functional enrichment analysis was performed to elucidate the biological process and signaling pathway that SGXZ decoction treated NAFLD. Finally, molecular docking combined with homology modeling was applied to further verify the binding between key active compounds and targets.

Results: A total of 31 active compounds and 220 corresponding targets in SGXZ decoction were collected. Moreover, 1544 overlapped targets of NAFLD were obtained, of which 78 common targets were intersected with targets of SGXZ decoction. Key compounds and targets were recognized from compound-target-disease and PPI network. Functional enrichment analysis revealed that multiple biological pathways were annotated including insulin resistance, HIF-1, PI3K-Akt, and MAPK signaling pathways. Molecular docking confirmed that gallic acid, chlorogenic acid and isochlorogenic acid A could combine firmly with all of the key targets.

Conclusion: SGXZ decoction could produce a promising effect for its multi-component, multi-target and multi-biological process in the treatment of NAFLD. The present study provides the novel insight into a comprehensive understanding for further study.

1. Introduction

Nonalcoholic fatty liver disease (NAFLD) is a progressive condition ranging from simple NAFL to nonalcoholic steatohepatitis (NASH), liver fibrosis, liver cirrhosis and even worse, the hepatocellular carcinoma (HCC) [1–2]. It is a metabolic disease representing the hepatic manifestation of a systemic metabolic disorder [3], and is demonstrated to be associated with obesity-related disorders and diabetes [4–5]. With the increase epidemic of obesity and metabolic-related comorbidities, global incidence of NAFLD is estimated to be 25% and keeps rising continually [6]. NAFLD has been one of the most leading courses of chronic hepatic disease, affecting approximately 1.7 billion individuals worldwide [7]. Moreover, it is one of the most common indication for liver transplantation and NAFLD-HCC is now the fastest growing demand of liver transplantation in the USA [8–9]. Lifestyle modification including healthy
diet and increased physical activity is recommended as the first-line treatment in NAFLD management [10]. Nevertheless, their effectiveness is limited in NAFLD patients due to the low readiness to change lifestyle, especially in terms of increasing physical activity [11]. Another challenge for the lifestyle intervention has been the occurrence of weight regain [12–13]. Up to date, there are no FDA-approved pharma therapies for the treatment of NAFLD at present [14]. Insulin sensitizer (rosiglitazone, pioglitazone and metformin), antioxidants (vitamin E), anti-inflammatory and lipid lowering drugs (atorvastatin and simvastatin) have been in practice for the treatment of NAFLD [15–16]. However, usage of rosiglitazone could increase the risk of cardiovascular causes [17], while long-term use of vitamin E [18–19] was reported to increase the risk of prostate cancer and might increase all-cause mortality. Moreover, the adverse outcome of pioglitazone [20–21], atorvastatin [22] and simvastatin [23] including weight gain, edema, heart failure, cytotoxicity and hepatotoxicity, should not be neglected. Therefore, it is an urgent need to develop pharmacological strategies for the treatment of NAFLD.

The complicated pathophysiological mechanism of NAFLD restricts the efficacy of the application of a single agent. Traditional Chinese Medicine (TCM) has been commonly used to treat hepatic disease and metabolic disorders-related obesity and type 2 diabetes mellitus [24–26]. The therapeutic effect of TCM on NAFLD mainly focused on the holistic regulations including lipid metabolism modulation, antioxidative stress, anti-inflammation, and gut microbiota modulation [27–29]. Shugan Xiaozhi (SGXZ) decoction is a composite traditional Chinese prescription composed of twelve herbs including pleuri Radix, Paeoniae radix alba (stir-baked), Aurantii Fructus Immaturus, Glycyrrhizae Radixet Rhizoma, Artemisiae Scopariae Herba, Gardeniae Fructus, Poria, Alismatis Rhizoma, Crataegi Fructus, Cassiae Semen, Nelumbinis Folium, and Pumex in a ratio of 2 : 1 : 3 : 1 : 6 : 2 : 4 : 6 : 6 : 6: 6: 6: 6. Previous clinical studies demonstrated that the treatment of SGXZ could partially protected and restored the liver functions in NASH patients, which was associated with significant reduction of aminotransferases, total cholesterol and triacylglycerol, and repairment of intestinal mucosal barrier [30–31]. Consistently, researches indicated that SGXZ decoction exerted hepatoprotective effect through regulating fatty acid β-oxidation, relieving intestinal microecological disorder and repairing intestinal mucosal barrier, thereby ameliorating NAFLD in rat model [32–33]. However, therapeutic mechanisms of SGXZ decoction in the treatment of NAFLD have not been fully clarified, and the active ingredients and key targets of SGXZ remain to be further investigated.

Based on the integration of system biology, bioinformatic and pharmacology [34–36], network pharmacology is currently used to explore the potential pharmacological effect and underlying mechanisms of a drug on a disease [37–38]. In particular, it concentrates on the elucidation of complex biological relationship among the drugs, targets, pathways and diseases from a systemic and holistic perspective [39]. The wholeness, relevance and systematic nature of network pharmacology are in line with the concept and treatment theory of TCM, providing a novel method and powerful tool to decipher the therapeutic mechanisms of TCM [40]. In addition, molecular docking is a computational method based on the analysis of the binding pose and affinity between small molecule and macromolecular target, which is widely utilized to predict and identify the potentially active compounds [41–42]. In our study, network pharmacology combined with molecular docking was applied to explore the
pharmacological and molecular mechanisms involved in the treatment of SGXZ decoction on NAFLD. First, chemical compounds in SGXZ decoction were collected and recognized through literature study and their corresponding targets were then obtained from Similarity ensemble approach (SEA) database. Second, the potential targets involved in the pathology of NAFLD were predicted from DisGeNET and GeneCards, and compared with the predicted targets of SGXZ decoction to intersect and identify the common targets. Subsequently, the network of “compound-target-disease” and protein-protein interaction (PPI) network were constructed to identify the potentially active ingredients and key targets. Furthermore, functional enrichment including Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) was analyzed via DAVID to explore the potential mechanism of SGXZ decoction for the treatment of NAFLD. Last, active ingredients were screened against the key targets by performing molecular docking to identify the compounds treating NAFLD. The overall flowchart of this study is described in Figure 1.

2. Materials And Methods

2.1 Collection of compounds and the putative targets of SGXZ decoction

The active compounds in SGXZ decoction were identified and collected through literature studies. Then, 2D or 3D conformations of the compounds were downloaded either from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) or through sketching in ChemDraw software. Moreover, the simplified molecular input line entry system (SMILS) numbers (Supplementary Table 1) of them were also acquired from PubChem or the conversion by ChemDraw. To predict the putative gene target of SGXZ decoction, the Simplified molecular input line entry system (SMILES) of each chemical compound was submitted into Similarity ensemble approach (SEA) database (https://sea.bkslab.org/), which relates proteins based on the set-wise chemical similarity among their ligands [43]. The predicted protein targets of SGXZ decoction were listed in Supplementary File 1.

2.2 Identification of potential targets of NAFLD

The “Non-alcohol Fatty Liver Disease” or the “Non-alcoholic Fatty Liver Disease” was used as the key words to predict the potential NAFLD-related gene targets by means of GeneCards (https://www.genecards.org/) and DisGeNET (https://www.disgenet.org/home/) disease database. DisGeNET is a discovery platform containing one of the largest publicly available collections of genes associated to human diseases [44]. Based on all annotated and predicted human genes, GeneCards provides comprehensive, user-friendly information via searchable, integrative database [45]. After removing the repetitive targets from the two databases, the NAFLD-related protein targets were retained for further study and the detailed information of the targets were provided in Supplementary File 2.
2.3 Construction of compound-target-disease network

The SGXZ decoction-target network was first constructed based on the compounds and their predicted targets of SGXZ using Cytoscape software (Version 3.7.1). Venny diagram 2.1 version (https://bioinfogp.cnb.csic.es/tools/venny/) was used to intersect the protein targets between predicted targets of SGXZ decoction and potential targets of NAFLD. Then, the compound-target-disease network was established based on the intersected common targets and their corresponding compounds. Moreover, the degree value of nodes in compound-target-disease network was analyzed with NetworkAnalyzer plug-in in Cytoscape software. Degree indicates the total number of nodes connected to this node in the network, and nodes whose degree value were higher played more essential regulatory role in the network. The compounds and targets that contributed significantly to the construction of compound-target-disease network were recognized and selected as key active compounds and targets according to their higher degree value compared to the average degree value.

2.4 Construction of protein-protein interaction (PPI) network

PPI network can elucidate the interaction relationship between targets, which further helps to identify the key nodes and targets in the network. The intersected common targets were submitted to STRING database (https://string-db.org/, ver. 11.5) [46]. The species was limited to “homo sapiens” and the confidence of interaction score was set ≥ 0.400 as default, which ensured a reasonable interaction between targets. Results with tabular text output (TSV) format of PPI network in STRING were exported and then better visualized through Cytoscape software. The nodes with highest degree value were considered as the key targets which mediated the interaction of the network and participated in the treatment of SGXZ decoction on NAFLD.

2.5 Enrichment analysis

Functional annotation of the common targets in PPI network was analyzed through “Database for Annotation, Visualization and Integrated Discovery” (DAVID) database (https://david.ncifcrf.gov/) to identify enriched biological themes of Gene ontology (GO) and visualize genes on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway maps. DAVID can provide a comprehensive set of functional annotation tools for investigators to understand biological meaning behind large list of genes and is widely used in bioinformatics [47-48]. Gene list of the common target proteins was uploaded into DAVID and the species was limited to “homo sapiens”. The functional annotation of biological processes (BPs), molecular functions (MCs), cellular components (CCs) and KEGG pathways with \( P \) value<0.05 were then performed for further analysis. Accordingly, target-pathway network was subsequently constructed using Cytoscape software.
2.6 Homology modeling and molecular docking

3D conformational SDF format of the selected key chemical compounds were downloaded from PubChem. The crystal structures of the key protein targets were retrieved in RCSB Protein Data Bank database (PDB, https://www.rcsb.org/). For targets without released crystal structures, homology modeling was utilized to establish the structure based on the protein template. The construction was completed in SWISS-MODEL website (https://swissmodel.expasy.org/), which is a fully automated protein structure homology-modelling server [49]. Then, the quality of model structures was evaluated via “Ramachandran plot” in SAVES v6.0 website (https://saves.mbi.ucla.edu/). The target protein with neither crystal structure nor the template was excluded in molecular docking. Before docking, the compounds were charged with CHARMm force filed and minimized for optimization in Discovery Studio 2016 (DS) software. The target proteins were also loaded in DS, and waters and other redundant chains were firstly deleted. Then excess structures of the protein were removed, incomplete amino acid residues were supplemented, and hydrogen were added to optimize the conformation. CDOCKER module in DS was applied to calculate the binding energy between key active ingredients and key target proteins. CDOCKER is a grid-based molecular docking method on the basis of CHARMm force field, which can produce high-precision docking results. For docking parameters of CDOCKER, the size of docking box was set to 20Å×20Å×20Å, “Top Hits” was set to 10 and “Pose Cluster Radius” was set to 0.5. Moreover, the docking site of each target protein was obtained from previous research or the active cavities detected in DS (Supplementary Table 2). Only one top hit with best docking pose of each complex was reported and saved for further analysis.

3. Results

3.1 Active compounds of SGXZ decoction and the corresponding targets

31 active compounds of SGXZ decoction (SX01~SX31) were identified and collected from the previous studies with 2D chemical structures displayed in Fig. 2. Totally, 220 targets of SGXZ decoction were screened after removing duplicates.

3.2 Potential targets of NAFLD and the compound-target-disease interaction

Potential targets of non-alcohol Fatty Liver Disease (NAFLD) were independently screened in DisGeNET and GeneCards database. After the overlapped targets were deleted, 1544 potential protein targets of NAFLD from the two databases were retained. Venny diagram showed that 78 common targets between SGXZ and NAFLD related targets were intersected, which represented the potential targets of SGXZ in the treatment of NAFLD (Fig. 3A).
The network of SGXZ decoction with its compounds and corresponding targets was first constructed to reveal the multi-target effect of SGXZ decoction (Fig. 3B). The network comprised of 252 nodes and 686 edges with 31 compounds and 220 corresponding targets. To further exhibited the directly regulatory effect of SGXZ decoction treating NAFLD, the network of compounds with common intersection targets was constructed (Fig. 3C), which consists of 109 nodes (31 compounds and 78 common targets) and 232 edges, suggesting a multi-component involved multi-targets of SGXZ decoction in the treatment of NAFLD. In addition, the top 10 compounds and common protein targets were obtained (Fig. 3D), of which the detailed information were listed in Table 1 and 2. Gallic acid possessed the highest degree value (degree=20), followed by kaempferol-3-O-galactoside (degree=19), licoricone (degree=17), isochlorogenic acid A (degree=16), chlorogenic acid (degree=12), liquiritin (degree=12), gardenoside (degree=11), naringin (degree=10), neohesperidin (degree=10) and poncirin (degree=9). These active compounds were recognized in SGXZ decoction to exert the pharmacologic effect on NAFLD. Moreover, as listed in Table 2, the top 10 targets with highest degree value included FGF1 (degree=15), VEGFA (degree=15), IL-2 (degree=13), LGALS3 (degree=11), SLC5A1 (degree=11), F3 (degree=10), ALDH1B1 (degree=7), CES2 (degree=7), IL-6 (degree=7) and P4HB (degree=7), which could be potentially acted on by the 10 key active compounds in SGXZ decoction.

Table 1

Key active compounds of SGXZ decoction in treatment of NAFLD

<table>
<thead>
<tr>
<th>Number</th>
<th>Ingredient</th>
<th>Molecular Formula</th>
<th>Pubchem_CID</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>SX01</td>
<td>Gallic acid</td>
<td>C$_7$H$_6$O$_5$</td>
<td>370</td>
<td>20</td>
</tr>
<tr>
<td>SX02</td>
<td>Naringin</td>
<td>C$<em>{27}$H$</em>{32}$O$_{14}$</td>
<td>442428</td>
<td>10</td>
</tr>
<tr>
<td>SX03</td>
<td>Neohesperidin</td>
<td>C$<em>{28}$H$</em>{34}$O$_{15}$</td>
<td>442439</td>
<td>10</td>
</tr>
<tr>
<td>SX04</td>
<td>Poncirin</td>
<td>C$<em>{28}$H$</em>{34}$O$_{14}$</td>
<td>442456</td>
<td>9</td>
</tr>
<tr>
<td>SX05</td>
<td>Gardenoside</td>
<td>C$<em>{17}$H$</em>{24}$O$_{11}$</td>
<td>24721095</td>
<td>11</td>
</tr>
<tr>
<td>SX07</td>
<td>Chlorogenic acid</td>
<td>C$<em>{16}$H$</em>{18}$O$_{9}$</td>
<td>1794427</td>
<td>12</td>
</tr>
<tr>
<td>SX11</td>
<td>Liquiritin</td>
<td>C$<em>{21}$H$</em>{22}$O$_{9}$</td>
<td>503737</td>
<td>12</td>
</tr>
<tr>
<td>SX13</td>
<td>Isochlorogenic acid A</td>
<td>C$<em>{25}$H$</em>{24}$O$_{12}$</td>
<td>6474310</td>
<td>16</td>
</tr>
<tr>
<td>SX27</td>
<td>Kaempferol-3-O-galactoside</td>
<td>C$<em>{21}$H$</em>{20}$O$_{11}$</td>
<td>5282149</td>
<td>19</td>
</tr>
<tr>
<td>SX31</td>
<td>Licoricone</td>
<td>C$<em>{22}$H$</em>{22}$O$_{6}$</td>
<td>5319013</td>
<td>17</td>
</tr>
</tbody>
</table>
### Table 2

Key targets of SGXZ decoction in the treatment of NAFLD

<table>
<thead>
<tr>
<th>UniprotKB</th>
<th>Target</th>
<th>Protein name</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>P05230</td>
<td>FGF1</td>
<td>Fibroblast growth factor 1</td>
<td>15</td>
</tr>
<tr>
<td>P15692</td>
<td>VEGFA</td>
<td>Vascular endothelial growth factor A</td>
<td>15</td>
</tr>
<tr>
<td>P60568</td>
<td>IL-2</td>
<td>Interleukin-2</td>
<td>13</td>
</tr>
<tr>
<td>P17931</td>
<td>LGALS3</td>
<td>Galectin-3</td>
<td>11</td>
</tr>
<tr>
<td>P13866</td>
<td>SLC5A1</td>
<td>Sodium/glucose cotransporter 1</td>
<td>11</td>
</tr>
<tr>
<td>P13726</td>
<td>F3</td>
<td>Tissue factor</td>
<td>10</td>
</tr>
<tr>
<td>P30837</td>
<td>ALDH1B1</td>
<td>Aldehyde dehydrogenase X, mitochondrial</td>
<td>7</td>
</tr>
<tr>
<td>O00748</td>
<td>CES2</td>
<td>Cocaine esterase</td>
<td>7</td>
</tr>
<tr>
<td>P05231</td>
<td>IL-6</td>
<td>Interleukin-6</td>
<td>7</td>
</tr>
<tr>
<td>P07237</td>
<td>P4HB</td>
<td>Protein disulfide-isomerase</td>
<td>7</td>
</tr>
</tbody>
</table>

#### 3.3 Outcome of PPI network

The PPI network was constructed from the common 78 targets (Fig. 4). Due to the lack of any interactions with other targets, Chymotrypsin-C (CTRC) and Alpha-ketoglutarate-dependent dioxygenase (FTO) were removed from the network. Therefore, the network contained 76 nodes and 376 edges with the average degree value of 9.64. The target proteins with higher degree value than average degree value were Interleukin-6 (IL-6), Vascular endothelial growth factor A (VEGFA), Estrogen receptor (ESR1), Hypoxia-inducible factor 1-alpha (HIF1A), Matrix metalloproteinase-9 (MMP9), Amyloid-beta precursor protein (APP), cAMP responsive element binding protein 1 (CREB1), Heat shock protein family A member 5 (HSPA5), Transcription factor p65 (RELA) and Interleukin-2 (IL-2). The 10 targets were indispensable for constructing interaction between protein and protein in the network, and were considered as the putative key target of SGXZ decoction for the treatment of NAFLD.

#### 3.4 GO enrichment analysis

Gene ontology (GO) functionally annotated targets in PPI network into three main aspects containing molecular function (MF), cellular component (CC) and biological process (BP). In total, 185 GO entries including 45 of MF, 25 of CC and 115 of BP were acquired based on the $P$ value ($P < 0.05$). For BP, the targets were mainly concentrated on the oxidation-reduction process, positive regulation of angiogenesis, negative regulation of apoptotic process, regulation of insulin-like factor receptor signaling pathway and
so on. For CC, the targets were mainly responsible for the extracellular space, extracellular exosome, cell surface, extracellular region, mitochondrion, extracellular matrix and so on. In regard to MF, the intersection targets were distributed in enzyme binding, insulin-like growth factor binding, electron carrier activity, oxidoreductase activity, receptor binding, protein heterodimerization activity, iron ion binding, glycosphingolipid binding and protein binding. The bubble plot and histogram of the most significant enriched GO terms were exhibited in Figure 5.

### 3.5 KEGG enrichment analysis

The 43 annotated pathways were totally obtained based on the targets in PPI network, and the top 20 of them with smallest significance value were shown in Figure 6A and Table 3. The bubble plot of the pathways suggested a concentration on the insulin resistance, amino acid metabolism, cancer-related pathways, inflammation-related pathways including MAPK signaling pathway, HIF-1 signaling pathway, TNF signaling pathway and PI3K-Akt pathway. Moreover, the common targets also concentrated on the endocrine participated biological process such as estrone signaling pathway and immunological pathway of T cell receptor. To reveal the network interaction of the pathways and the involved targets, target-pathway network was constructed and analyzed which consisted of 58 nodes and 130 edges as shown in Figure 6B, indicating a complicated interaction among them. Pathways with largest degree including pathways in cancer, PI3K-Akt, MAPK signaling pathway, proteoglycan in cancer and HIF-1 pathway, suggesting their significant role in the treatment of SGXZ decoction on NAFLD.

#### Table 3

Annotation of KEGG pathways and the involved potential targets
<table>
<thead>
<tr>
<th>KEGG ID</th>
<th>Term</th>
<th>Count</th>
<th>P value</th>
<th>Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa04931</td>
<td>Insulin resistance</td>
<td>8</td>
<td>7.75E-05</td>
<td>RPS6KA3, PTPN1, IL-6, CREB1, PRKCE, PTPN11, RELA, NFKB1</td>
</tr>
<tr>
<td>hsa04066</td>
<td>HIF-1 signaling pathway</td>
<td>7</td>
<td>3.17E-04</td>
<td>IL-6, SERPINE1, PRKCA, HIF1A, RELA, NFKB1, VEGFA</td>
</tr>
<tr>
<td>hsa00380</td>
<td>Tryptophan metabolism</td>
<td>5</td>
<td>5.77E-04</td>
<td>MAOB, ALDH2, ALDH1B1, MAOA, CYP1A1</td>
</tr>
<tr>
<td>hsa04726</td>
<td>Serotonergic synapse</td>
<td>7</td>
<td>6.91E-04</td>
<td>APP, CYP2C8, MAOB, MAOA, ALOX5, PRKCA, ALOX12</td>
</tr>
<tr>
<td>hsa00340</td>
<td>Histidine metabolism</td>
<td>4</td>
<td>0.001191</td>
<td>MAOB, ALDH2, ALDH1B1, MAOA</td>
</tr>
<tr>
<td>hsa05030</td>
<td>Cocaine addiction</td>
<td>5</td>
<td>0.001253</td>
<td>CREB1, MAOB, MAOA, RELA, NFKB1</td>
</tr>
<tr>
<td>hsa05200</td>
<td>Pathways in cancer</td>
<td>12</td>
<td>0.001302</td>
<td>IL-6, MMP1, MMP2, RARA, PRKCA, IKBKG, FGF1, HIF1A, MMP9, RELA, NFKB1</td>
</tr>
<tr>
<td>hsa00330</td>
<td>Arginine and proline metabolism</td>
<td>5</td>
<td>0.001352</td>
<td>MAOB, P4HA1, ALDH2, ALDH1B1, MAOA</td>
</tr>
<tr>
<td>hsa05134</td>
<td>Legionellosis</td>
<td>5</td>
<td>0.001803</td>
<td>IL-6, RELA, NFKB1, HSPD1, HSPA1A</td>
</tr>
<tr>
<td>hsa05161</td>
<td>Hepatitis B</td>
<td>7</td>
<td>0.002742</td>
<td>IL-6, CREB1, PRKCA, IKBKG, MMP9, RELA, NFKB1</td>
</tr>
<tr>
<td>hsa04010</td>
<td>MAPK signaling pathway</td>
<td>9</td>
<td>0.002986</td>
<td>RPS6KA3, NR4A1, PRKCA, IKBKG, MAPT, FGF1, RELA, NFKB1, HSPD1, HSPA1A</td>
</tr>
<tr>
<td>hsa05205</td>
<td>Proteoglycans in cancer</td>
<td>8</td>
<td>0.003139</td>
<td>MMP2, PTPN6, PTPN11, PRKCA, HIF1A, ESR1, MMP9, VEGFA</td>
</tr>
<tr>
<td>hsa05142</td>
<td>Chagas disease (American trypanosomiasis)</td>
<td>6</td>
<td>0.003296</td>
<td>IL-6, SERPINE1, IKBKG, RELA, NFKB1, IL-2</td>
</tr>
<tr>
<td>hsa04668</td>
<td>TNF signaling pathway</td>
<td>6</td>
<td>0.003726</td>
<td>IL-6, CREB1, IKBKG, MMP9, RELA, NFKB1</td>
</tr>
<tr>
<td>hsa05202</td>
<td>Transcriptional misregulation in cancer</td>
<td>7</td>
<td>0.005495</td>
<td>IL-6, BCL2A1, IGFBP3, RARA, MMP9, RELA, NFKB1</td>
</tr>
<tr>
<td>hsa04151</td>
<td>PI3K-Akt signaling pathway</td>
<td>10</td>
<td>0.005885</td>
<td>NR4A1, IL-6, CREB1, PRKCA, IKBKG, FGF1, RELA, NFKB1, IL-2, VEGFA</td>
</tr>
<tr>
<td>hsa05219</td>
<td>Bladder cancer</td>
<td>4</td>
<td>0.007229</td>
<td>MMP1, MMP2, MMP9, VEGFA</td>
</tr>
<tr>
<td>hsa05215</td>
<td>Prostate cancer</td>
<td>5</td>
<td>0.010389</td>
<td>CREB1, SRD5A2, IKBKG, RELA, NFKB1</td>
</tr>
<tr>
<td>hsa04915</td>
<td>Estrogen signaling pathway</td>
<td>5</td>
<td>0.015491</td>
<td>CREB1, MMP2, ESR1, MMP9, HSPA1A</td>
</tr>
<tr>
<td>hsa04660</td>
<td>T cell receptor signaling</td>
<td>5</td>
<td>0.016021</td>
<td>PTPN6, IKBKG, RELA, NFKB1, IL-2</td>
</tr>
</tbody>
</table>
3.6 Verification of the binding between key compounds and targets

Molecular docking was employed to verify the binding between the key compounds with key targets, revealing the potential therapeutic effect of SGXZ decoction acting on these targets. Crystal structures of the targets were downloaded from PDB database except SLC5A1, CES2 and ALDH1B1 for their lack of 3D structures. Homology modeling was used to build the structures of SLC5A1 and ALDH1B1 according to previous studies [50-51] (Supplementary Figure 1). However, the crystal structure of CREB1 could be obtained neither from the PDB database nor through the homology modeling since none of templates with good quality existed. Besides, 3 duplicated key targets between compound-target-disease network and PPI network were removed. Thus, a total of 10 key compounds were docked into 16 key targets. After docking, 160 pair of compound-target complexes were retained according to their best binding affinity, and the heat map of their docking energy was shown in Figure 7. Moreover, the binding energy of each pair of compound-target complex was provided in Supplementary Table 3. Binding energy with negative value indicated that the ligand molecule was able to combine with the receptor target proteins. The lower and more negative binding energy suggested a better binding affinity of active compounds with target proteins. Almost half of the key compound could combine with most of the targets. Among them, gallic acid (SX01), chlorogenic acid (SX07) and isochlorogenic acid A (SX13) could bind into all the key targets with the lowest binding energy. The accurate values of binding energy between these three key active compounds and key targets were shown in Table 4. Isochlorogenic acid A had the most negative binding energy with HSPA5 (-53.91 kcal/mol), followed by the RELA (-51.73 kcal/mol), MMP9 (-48.88 kcal/mol), SLC5A1 (-34.57 kcal/mol), HIF1A (-39.94 kcal/mol) and FGF1 (-38.21 kcal/mol). Chlorogenic acid also showed great binding affinity with HSPA5 (-43.27 kcal/mol). Gallic acid exhibited with the most negative binding energy with ESR1 (-35.57 kcal/mol), followed by HSPA5 (-33.25 kcal/mol) and RELA (-32.32 kcal/mol). Residues that produce hydrogen bonds shared by all 3 key active compounds with target were also listed in the Table 4. Moreover, the interaction between the best poses of three compounds and targets were displayed in Figure 8. The results of molecular docking revealed that gallic acid, chlorogenic acid and isochlorogenic acid A might be the most potentially active ingredients of SGXZ decoction treating NAFLD.

Table 4

<p>| Binding energy between key active compounds and targets (kcal/mol) |</p>
<table>
<thead>
<tr>
<th>Target proteins</th>
<th>Ingredients</th>
<th>Average binding energy</th>
<th>Hydrogen Bond Interactions (Amino acid)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FGF1</td>
<td>SX01: -27.07</td>
<td>SX07: -22.06</td>
<td>SX13: -38.21</td>
</tr>
<tr>
<td>VEGFA</td>
<td>SX01: -21.22</td>
<td>SX07: -18.94</td>
<td>SX13: -29.06</td>
</tr>
<tr>
<td>LGALS3</td>
<td>SX01: -22.81</td>
<td>SX07: -17.71</td>
<td>SX13: -34.94</td>
</tr>
<tr>
<td>SLC5A1</td>
<td>SX01: -27.92</td>
<td>SX07: -31.68</td>
<td>SX13: -44.10</td>
</tr>
<tr>
<td>F3</td>
<td>SX01: -24.94</td>
<td>SX07: -5.43</td>
<td>SX13: -19.86</td>
</tr>
<tr>
<td>CES2</td>
<td>SX01: -27.76</td>
<td>SX07: -23.35</td>
<td>SX13: -32.97</td>
</tr>
<tr>
<td>IL-6</td>
<td>SX01: -26.37</td>
<td>SX07: -17.42</td>
<td>SX13: -30.38</td>
</tr>
<tr>
<td>P4HB</td>
<td>SX01: -19.07</td>
<td>SX07: -12.82</td>
<td>SX13: -33.19</td>
</tr>
<tr>
<td>ESR1</td>
<td>SX01: -35.57</td>
<td>SX07: -24.86</td>
<td>SX13: -3.86</td>
</tr>
<tr>
<td>MMP9</td>
<td>SX01: -23.77</td>
<td>SX07: -25.52</td>
<td>SX13: -48.88</td>
</tr>
<tr>
<td>HSPA5</td>
<td>SX01: -33.25</td>
<td>SX07: -43.27</td>
<td>SX13: -53.91</td>
</tr>
<tr>
<td>RELA</td>
<td>SX01: -32.32</td>
<td>SX07: -30.32</td>
<td>SX13: -51.73</td>
</tr>
</tbody>
</table>

N/A, none hydrogen bond interaction residues shared by SX01, SX07 and SX13 with IL-6

### 4. Discussion

NAFLD is a metabolic related syndrome characterized as dysfunctional hepatic lipid metabolism and insulin resistance [52–53], which has been considered the fastest growing cause of HCC [54] and strongly associated with the increasing risks of type 2 diabetes, cardiovascular disease, chronic kidney disease and hypertension [55–57]. Currently, no pharmacological therapies were approved by FDA in the treatment of NAFLD and seeking an effective agent is urgently desiderated. TCM has been proved to treat NAFLD with the unique advantages of holistic concept and differentiation treatment, as well as mechanisms of multi-target and multi-channel action [27–58–59](34344394; 31558860; 33639194). Shugan Xiaozhi (SGXZ) decoction was proposed and designed with insights gained from two classic prescriptions of TCM - “Sini decoction” and “Yinchenhao decoction”, and showed a good efficacy for the treatment of NAFLD [32–60]. In addition, both of Sini decoction and Yinchenhao decoction exerted
therapeutic effect for NAFLD and NASH [61–63]. Nevertheless, the underlying therapeutic mechanisms of SGXZ decoction remained to be further elucidated.

In the present study, active compounds of SGXZ decoction with corresponding targets were firstly identified, and the overlapping targets between SGXZ decoction and NAFLD were considered to be the main targets these compounds acted on. Moreover, compounds including series of phenolic acids and flavonoids were selected as the key bioactive ingredients for the treatment of SGXZ decoction on NAFLD based on their contributions in the compound-target-disease network (Figure 2, Table 1). Previous study indicated that the intake of phenolic acids alleviated hepatic steatosis, reduced fibrosis and the insulin resistance in NAFLD patients [64]. Gallic acid, a simple polyphenol, was reported to reduce lipid accumulation that is related to β-oxidation and ketogenesis [65]. Specifically, the hepatoprotective effect of gallic acid attributed to the repression of inflammatory signaling pathways including nuclear factor-κB (NF-κB)/TNF-α/IL-6 and ROS/NF-κB /TNF-α in NAFLD rats [66–67]. Chlorogenic acid, a natural polyphenol extracted from Artemisiae Scopariae (Yin-Chen in Chinese) [68] and Lonicera japonica (Jin-Yin-Hua) [69], could ameliorate HFD-induced hepatic steatosis and inflammation via inhibition of TNF-α and IL-6 in liver, which was associated with regulation of gut microbiota and an increase of Glucagon-like peptide-1 secretion [70–71]. Isochlorogenic acid A was suggested to possess properties of hepatoprotective and anti-hepatitis B through suppressing oxidation [72]. Moreover, it was indicated that isochlorogenic acid A exerted a protective effect on liver fibrosis through inhibiting inflammation via HMGB1/TLR4/NF-kB signaling pathways [73]. In addition, flavonoids are natural products widely distributed in plants, exhibiting not only antidiabetic and hypoglycemic activities but also anti-inflammatory and antioxidant properties, and were recognized to have protective effect in the treatment of NAFLD [74–76]. As a kind of flavonoid, liquiritin could ameliorate cyclophosphamide-induced liver injury and inflammation by inhibiting the elevated MMP-9 expression, hepatic infiltration of neutrophils, myeloperoxidase activity, IL-6 mRNA expression and NF-κB activation [77]. Naringin is a flavanone glycoside isolated from Aurantii Fructus Immaturus (Zhi-Shi) [78] and has been proved to improve lipid metabolism disorders through reducing hepatic lipid accumulation in tissue-engineered NAFLD model [79]. Neohesperidin could induce the PGC-1α expression through activating AMP-activated protein kinase (AMPK), increasing hepatic mitochondrial biogenesis and fatty acid oxidation in NAFLD mice [80]. To sum up, three phenolic acids (gallic acid, chlorogenic acid, isochlorogenic acid A) and three flavonoids (liquiritin, naringin, neohesperidin) with hepatoprotective effect could serve as the potentially main active ingredients of SGXZ decoction for the treatment of NAFLD.

Together with key targets in compound-target-disease network (Figure 3, Table 2), 10 targets with highest degree in the PPI network were as well recognized as core targets of SGXZ decoction treating NAFLD (Figure 4). The target protein level of VEGFA, FGF1, IL-2, LGALS3, SLC5A1 and IL-6 were excessively expressed in compound-target-disease network. Moreover, IL-6 possessed the highest degree in PPI network, followed by VEGFA, ESR1, HIF1A, MMP9, HSPA5, RELA and IL-2. Particularly, IL-6, IL-2 and VEGFA contributed significantly to the construction of both compound-target-disease and PPI networks, indicating their critical role among the targets. IL-6 has been considered to cause inflammatory properties and a vital factor associated with the pathology of metabolic disorders [81]. Numerous studies have
highlighted that increased IL-6 level promoted hepatic insulin resistance [82] and impaired the lipid metabolism [83]. IL-6 deficiency ameliorated the hepatic inflammation and injury in NASH mice fed with methionine and choline-deficient diet [84]. Similarly, upregulated IL-2 potentially leads to the increased insulin resistance as well as several other metabolic inflammatory markers in obese population [85].

VEGFA is a major proangiogenic cytokine regulating angiogenesis. Increased VEGFA was reported in liver of animal model and serum of NASH patients [86–87], which accelerates angiogenesis and therefore drives hepatic inflammation and fibrosis in NAFLD [87].

Hepatic steatosis and lipid accumulation could produce a hypoxic proangiogenic microenvironment, while the response to this condition was mediated by hypoxia-inducible factors (HIFs) [88–89]. Hypoxia upregulated the HIF1A expression in liver, leading to the aggravated steatosis by suppression of fatty acid (FFA) β-oxidation and by induction of FFA uptake and inflammatory factors [89–90]. Furthermore, HIF1A promoted liver fibrosis in NAFLD by activating PTEN/ NF-κB p65 signaling pathway [91]. RELA, also known as p65, is one of the five members in NF-κB family and is a pivotal transcription factor regulating inflammatory molecules [92]. Inhibition of NF-κB signaling alleviated hepatic lipid accumulation and hepatic inflammation in NAFLD [93–94]. Moreover, abnormal lipid deposition as well as insulin resistance in NAFLD often leads to endoplasmic reticulum (ER) stress, which further triggers the unfolded protein response and thereby causes inflammation in hepatocytes [95]. GRP78 (HSPA5) is a chaperone heat shock protein playing the central role in maintaining ER proteostasis under excessive stress [96]. Decreased expression of ER stress molecule GRP78 contributed to reducing fasting glucose and lipid profile, and attenuating NAFLD [97–99]. LGALS3 has been shown to participate in glucose intolerance and lipid metabolism disorders, which is an essential regulator of insulin resistance, fibrosis and inflammation cytokines including TNF-α, IL-6 and IL-1β [100–101]. FGF1 exerts a protective role in series of metabolic disorders [102]. Investigations revealed that FGF1 could reduce blood glucose and ameliorate hepatic steatosis, inflammation and fibrosis through modulation of oxidative stress and ER stress [103–104]. The onset of hepatic inflammation caused the fibrogenesis in NAFLD, which was manifested by deposited extracellular matrix (ECM) proteins including collagens, elastin and fibronectin [105–106]. MMP-9 performs the vital role in modulating and degrading gelatins, collagens and other ECM compounds [107–108]. A decreased MMP-9 was associated with more advanced fibrosis and serum liver injury indices (AST, GGT) in NAFLD patients, while increased MMP-9 activity could precede the clearance of fibrotic matrix [109–110]. SLC5A1 encodes the sodium glucose cotransporter 1 (SGLT1), and inhibition of SGLT1 not only contributed to regulate hepatic glucose metabolism but also mitigate development and progression of NAFLD [111–112]. To conclude, it was assumed that SGXZ decoction might perform comprehensive regulations of anti-inflammation, anti-lipid deposition, anti-insulin resistance, anti-fibrosis, and anti-ER stress on NAFLD through these key targets.

After that, GO functional enrichment analysis revealed that the targets in PPI network mainly involved oxidation, positive regulation of angiogenesis, metabolic process, hypoxia, extracellular matrix and insulin-like factor binding, and other biological processes (Figure 5). The results of GO analysis coincided with the indispensable contribution of aforementioned key targets among all targets in PPI network. In addition, KEGG enrichment analyses indicated that the action pathway mainly included insulin resistance,
amino acid metabolism, cancer-related pathways, inflammation-related pathways, estrone signaling pathway and T cell receptor pathway (Figure 6A, Table 3). As shown in Figure. 6B, the ‘target-pathway’ interaction network suggested that bioactive compounds in SGXZ decoction performed therapeutic role in regulation of NAFLD through multi-targets and multi-pathways. Through literature retrieval, six signaling pathways including PI3K-Akt, MAPK, insulin resistance, HIF-1, Tryptophan metabolism and Estrogen signaling pathway were recognized as core pathways involved in the treatment of SGXZ decoction on NAFLD. The six core pathways and their interaction were cited from KEGG database and displayed in Figure 9. Among the core pathways, PI3K-Akt, MAPK and HIF-1 signaling pathway were associated with inflammation that is the driving force for the development and evolution of NAFLD [113–114].

Researches provided evidence that regulation of PI3K-Akt signaling mediators could improve hepatocyte damage, hepatic gluconeogenesis and lipid disorder [115–117]. Moreover, inhibiting apoptosis and promoting autophagy via the reactive oxygen species (ROS)/MAPK pathway significantly decreased total cholesterol and triglyceride levels of both plasma and liver in high fat diet-fed mice [118]. In addition, increased ROS formation, triglyceride and lipid accumulation in hepatocyte leaded to excessive oxidative stress and inflammatory responses that further produced a hypoxic microenvironment [88–119] (33272734, 33804956). HIF-1 signaling acts as the key response to hypoxia. Loss of HIF-1α (HIF1A) activates the oxidative stress, promoted the lipid deposition and the secretion of pro-inflammatory IL-6 and TNF-α, suggesting its regulatory role in the progression of NAFLD [119–120] (33272734, 30242180). Indeed, increased FFAs and lipogenesis in hepatocytes, release of proinflammatory cytokines from adipose tissue such as IL-6 and TNF-α, together with altered gut microbiota gave rise to the insulin resistance, which predisposes to the development of NAFLD and progression to NASH [121–123]. Tryptophan is a kind of amino acid that could produce indoles by bacterial enzyme trytophanase A in intestinal epithelial cell. Indoles promote intestinal barrier function and could translocate from the intestinal to the liver to modulate hepatic lipid metabolism and inflammation to protect against NAFLD [124]. Moreover, the hepatic metabolism and inflammation in the onset and progression of NAFLD showed a sex difference, which was potentially due to different estrogen signaling activity [125–126]. Estrogen deficiency increased risk for liver fibrosis, and postmenopausal women were more likely to develop NAFLD than men [127–128]. In general, the above six core pathways and their interaction were correlated to NAFLD and might be acted on by SGXZ decoction in the treatment of NAFLD.

In the end, molecular docking (Table 4, Figure 8) showed that most of the key compounds could be docked into the key protein targets. Moreover, bioactive compounds of isochlorogenic acid A, chlorogenic acid and gallic acid could bind best with all of the key targets, especially with HSPA5, RELA, MMP9, SLC5A1, HIF1A and FGF1. The results of molecular docking simulations provided a putative pharmacological activity for the treatment of SGXZ decoction on NAFLD. Our present study applied network pharmacology combined with molecular docking to predict the key active ingredients and pathways of SGXZ decoction in NAFLD treatment, which still need to be verified by future experimental validations.
5. Conclusions

In this study, network pharmacology approach combined with molecular docking method were performed to elucidate the underlying mechanisms, including the putative active ingredients, key targets and pivotal signaling pathways, of SGXZ decoction in the treatment of NAFLD. Among the predicted key compounds, three bioactive ingredients of SGXZ, isochlorogenic acid A, chlorogenic acid and gallic acid might exert the essential pharmacological effect. Furthermore, VEGFA, IL-6, IL-2, HSPA5, RELA, MMP9, SLC5A1, HIF1A and FGF1 were assumed to be the key targets acted on by those bioactive ingredients. In addition, signaling pathways of inflammation, oxidative stress, insulin resistance, intestinal barrier function, and lipid metabolism were mainly involved. The results provided an insight into the therapeutic strategies of NAFLD and evidence for future research. Nonetheless, more experimental approaches should be carried out to validate the findings.

Abbreviations

TCM: Traditional Chinese medicine; NAFLD: Nonalcoholic fatty liver disease; SGXZ: Shugan Xiaozhi decoction; SMILES: Simplified molecular input line entry system; SEA: Similarity ensemble approach; PDB: RCSB PROTEIN DATA BANK; PPI: Protein-protein interaction; STRING: Search Tool for the Retrieval of Interacting Genes/Proteins; DS: Discovery studio software 2016; GO: Gene ontology; BP: Biological process; CC: Cell components; MF: Molecular functions; KEGG: Kyoto Encyclopedia of Genes and Genomes; FFA: Free fatty acid

Declarations

Acknowledgements

Not applicable.

Authors' contributions

YR and XYF conceived and designed the study. YR collected and acquired the data. YHL, JDS, XLY and FL performed the data analysis, RY wrote the first version of manuscript, YHL made the critical revision, and XYF finalized the manuscript. All authors are responsible for reviewing and evaluating the data. All authors read and approved the final manuscript.

Funding

This work is financially supported by the Shenzhen Natural Science Foundation (grant No. JCYJ20190812161605538) and Sanming Project of Medicine in Shenzhen (grant No. SZSM201612074).
Availability of supporting data

All data are available in the manuscript and they are showed in figures, tables and supplementary files.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflicts of interest in relation to this work.

References


52. Kuchay MS, Choudhary NS, Mishra SK. Pathophysiological mechanisms underlying MAFLD. Diabetes Metab Syndr. 2020;14:1875–87.


125. Della Torre S. Non-alcoholic Fatty Liver Disease as a Canonical Example of Metabolic Inflammatory-Based Liver Disease Showing a Sex-Specific Prevalence: Relevance of Estrogen Signaling. Front Endocrinol (Lausanne). 2020;11:572490.


**Figures**
Figure 1

Schematic diagram of the whole study design.

Figure 2

Homology modeling and Molecular docking
2D structures of the 31 identified compounds in SGXZ decoction.

Figure 3

(A) Intersection targets between SGXZ decoction and NAFLD related targets via Venny diagram, (B) Compound-target network of SGXZ decoction. SGXZ decoction, chemical compounds and targets were colored as red, pink and purple respectively, and the 78 intersection targets in the network were displayed in blue, (C) Interaction network of compound-target-disease, the size of nodes corresponds to the value of the degree, (D) Top 10 putative compounds and targets with highest degree value in compound-target-disease network.

Figure 4
(A) PPI network of potential targets of SGXZ decoction treating NAFLD. The size of the node with shade of the color indicated the corresponding degree value, (B) The identified top 10 putative target with degree value higher than the average according to the PPI network.

Figure 5

Representative GO enrichment analysis based on the targets in PPI network. Bubble plot of the top 10 terms of BP, CC and MF. The bigger the dot, the more genes are enriched in the term, while the redder indicates the smaller $P$ value.
Figure 6

Results of KEGG pathway enrichment analysis based on the targets in PPI network. (A) The bubble plot of top 20 pathway enrichment analysis, (B) Target-pathway interaction network suggested the underlying mechanisms of SGXZ decoction for the treatment on NAFLD. Blue circle nodes indicated the targets and orange diamond represents the pathways. The gray edges linked the interaction between targets and pathways, and the node size was proportional to the degree value.
Figure 7

Heat map of the docking energy (kcal/mol) between key active ingredients and key targets. The redder the color, the lower the binding energy and the stronger the binding ability.
Figure 8

**Interaction mode between SX01, SX07, SX13 and targets.** Target proteins were displayed as solid ribbon and surface mode, and SX01, SX07 and SX13 were displayed in sticks and colored in yellow, pink and blue respectively. The residues interacting with the compounds were shown as sticks with their name labeled in red.

Figure 9

**The predicted six core pathways and their interaction involved in the therapeutics mechanisms of SGXZ decoction treating NAFLD.**

Supplementary Files

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

- SupplementaryFigure1.tif
- SupplementaryFile1.xlsx
- SupplementaryFile2.xlsx
- SupplementaryTable1.docx
- SupplementaryTable2.docx
- SupplementaryTable3.docx